Capítulo 5

Espacios Vectoriales Euclídeos. Métodos de los mínimos cuadrados

En este tema iniciamos el estudio de los conceptos geométricos de distancia y perpendicularidad en $\mathbb{K}^n$. Empezaremos con las definiciones de estos conceptos. Posteriormente, analizaremos la existencia de bases formadas por vectores mutuamente perpendiculares lo que nos llevará al teorema de la mejor aproximación.

5.1. Producto escalar o interno. Norma y distancia euclídea. Ortogonalidad

Definición 5.1.1. *Sea V un espacio vectorial sobre $\mathbb{K}$. Un* ***producto escalar o interno sobre*** *V es una aplicación tal que a cada par de elementos $\boldsymbol{x}$ e $\boldsymbol{y}$ de V le asigna un número $\langle \boldsymbol{x}, \boldsymbol{y}\rangle \in \mathbb{K}$ que satisface las siguientes propiedades:*

- $\langle \boldsymbol{x}, \boldsymbol{y}+\boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{x}, \boldsymbol{z}\rangle$,
- $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x}\rangle}$,
- $\langle a\boldsymbol{x}, \boldsymbol{z}\rangle = \overline{a}\langle \boldsymbol{x}, \boldsymbol{z}\rangle$,
- $\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$ *y* $\langle \boldsymbol{x}, \boldsymbol{x}\rangle = 0$ *si y sólo si* $\boldsymbol{x} = \boldsymbol{0}$,

para cualesquiera $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$, $a \in \mathbb{K}$.

De las propiedades anteriores se deduce que $\langle \boldsymbol{x} + \mathbf{y}, \mathbf{z}\rangle = \langle \boldsymbol{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle$, $\langle \boldsymbol{x}, a\mathbf{z}\rangle = a\langle \boldsymbol{x}, \mathbf{z}\rangle$. Nótese que si $\mathbb{K}$=$\mathbb{R}$, entonces $\langle \boldsymbol{x}, \mathbf{y}\rangle = \langle \mathbf{y}, \boldsymbol{x}\rangle$.

Ejemplos.
(1) En $\mathbb{K}^n$, dados dos vectores $\boldsymbol{x}$ e $\mathbf{y}$, definimos un producto interno como

$$\langle \boldsymbol{x}, \mathbf{y}\rangle = \overline{x_1}y_1 + \overline{x_2}y_2 + \cdots + \overline{x_n}y_n = \boldsymbol{x}^*\mathbf{y}.$$

Obsérvese que si $\mathbb{K} = \mathbb{R}$, entonces

$$\langle \boldsymbol{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \boldsymbol{x}^t \mathbf{y}.$$

(2) En el espacio vectorial de las funciones continuas definidas sobre un intervalo cerrado $[a, b]$ con valores en $\mathbb{R}$, dadas f y g, definimos un producto interno como

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

Definición 5.1.2. *Un espacio vectorial con un producto interno se llama **espacio vectorial euclídeo**.*

5.1.1. Norma y distancia euclídea

En un espacio euclídeo usaremos la notación:

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

el número $||\boldsymbol{x}||$ se llama **norma euclídea** (o simplemente norma) de $\boldsymbol{x}$.

Lema 5.1.1 (Desigualdad de Cauchy-Schwarz.)**.** *Sea V un espacio vectorial euclídeo y $\boldsymbol{x}, \boldsymbol{y} \in V$, entonces se verifica:*

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq ||\boldsymbol{x}|| \, ||\boldsymbol{y}||,$$

además, la igualdad se da si y sólo si $\boldsymbol{x}$ e $\boldsymbol{y}$ son proporcionales.

Lema 5.1.2 (ley del paralelogramo.)**.** *Sea V un espacio vectorial euclídeo, se cumple:*

$$||\boldsymbol{x} + \boldsymbol{y}||^2 + ||\boldsymbol{x} - \boldsymbol{y}||^2 = 2\left(||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2\right).$$

cualquiera que sean $\boldsymbol{x}, \boldsymbol{y} \in V$.

Proposición 5.1.3. *Sea V un espacio euclídeo sobre $\mathbb{K}$. La **norma** definida anteriormente en V verifica las siguientes propiedades:*

- $||\boldsymbol{x}|| = 0$ *si y sólo si* $\boldsymbol{x} = \boldsymbol{0}$,
- $||a\boldsymbol{x}|| = |a| \cdot ||\boldsymbol{x}||$, *y*
- $||\boldsymbol{x} + \boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$ *(desigualdad triangular).*

para cualesquiera $\boldsymbol{x}, \boldsymbol{y} \in V$ y $a \in \mathbb{K}$.

Si V es un espacio euclídeo, entonces $||\boldsymbol{x} - \mathbf{y}||$ representa la distancia entre los vectores $\boldsymbol{x}$ e $\mathbf{y}$, $\mathrm{d}(\boldsymbol{x}, \mathbf{y}) = ||\boldsymbol{x} - \mathbf{y}||$.

Proposición 5.1.4. *Sea V un espacio euclídeo sobre $\mathbb{K}$. La **distancia** definida anteriormente en V verifica las siguientes propiedades:*

- $d(\boldsymbol{x}, \boldsymbol{x}) = 0$.
- $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$.
- $d(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z}) = d(\boldsymbol{x}, \boldsymbol{y})$.
- $d(\boldsymbol{x}, \boldsymbol{y}) \leq d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})$ *(desigualdad triangular).*

para cualesquiera $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.

Definición 5.1.3. *Un espacio vectorial en el que esté definida una distancia se llama* ***espacio vectorial métrico****.*

5.1.2. Ortogonalidad

Definición 5.1.4. *Sea V un espacio vectorial euclídeo. Diremos que dos vectores $\boldsymbol{x}, \boldsymbol{y} \in V$ son* ***ortogonales****, o perpendiculares, si $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. En ese caso, escribiremos $\boldsymbol{x} \perp \boldsymbol{y}$.*

Dados dos vectores no nulos $\boldsymbol{x}, \mathbf{y} \in \mathbb{R}^n$, sea t el único ángulo (en radianes) del intervalo $[0, \pi]$ tal que

$$\cos(t) = \frac{\langle \boldsymbol{x}, \mathbf{y} \rangle}{||\boldsymbol{x}||\,||\mathbf{y}||} \in [-1, 1],$$

entonces diremos que t es el **ángulo** que forman $\boldsymbol{x}$ e $\mathbf{y}$.

Teorema 5.1.5 (Pitágoras)**.** *Sea V un espacio vectorial euclídeo. Si $\boldsymbol{x}$ e $\boldsymbol{y}$ son vectores ortogonales en V, entonces*

$$||\boldsymbol{x}+\boldsymbol{y}||^2 = ||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2.$$

En general, si $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r$ son ortogonales dos a dos, entonces

$$||\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_r||^2 = ||\boldsymbol{x}_1||^2 + ||\boldsymbol{x}_2||^2 + \cdots + ||\boldsymbol{x}_r||^2.$$

Observación. En cualquier espacio vectorial euclídeo real es cierto el recíproco. Sin embargo, en $\mathbb{C}^n$ eso no ocurre (basta tomar en $\mathbb{C}^2$ los vectores $\boldsymbol{x} = (0, 1)^t$ e $\mathbf{y} = (0, i)^t$).

5.2. Bases ortonormales: Método de ortogonalización de Gram–Schmidt

Definición 5.2.1. *Sea V un espacio vectorial euclídeo. Diremos que un conjunto A de vectores en V es un* ***sistema ortogonal*** *si los vectores de A son ortogonales dos a dos. Si además, todos tienen norma euclídea igual a 1, diremos que A es un* ***sistema ortonormal.***

Proposición 5.2.1. *Sea V un espacio vectorial euclídeo. Todo sistema ortogonal de vectores no nulos en V es un conjunto linealmente independiente.*

Es fácil encontrar ejemplos que muestran que el recíproco de la proposición anterior no es cierto.

Definición 5.2.2. *Sea V un espacio vectorial euclídeo. Diremos que una base B de un subespacio de V es ortonormal (ortogonal) si B es un sistema ortonormal (ortogonal).*

Teorema 5.2.2 (Método de ortogonalización de Gram-Schmidt.)**.** *Sea V un espacio vectorial euclídeo. Sea $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r\}$ un conjunto de vectores linealmente independiente en V. Entonces existe un sistema ortogonal $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r\}$ tal que:*

$$\langle \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p \rangle = \langle \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_p \rangle \qquad \textit{para} \;\; p = 1, 2, \ldots, r.$$

Los vectores $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_p$, pueden hallarse usando el método de Gram–Schmidt que consiste en:

(1) Tomar $\boldsymbol{y}_1 = \boldsymbol{x}_1$.

(2) Para $p = 2, 3, \ldots, r$ tomar $\boldsymbol{y}_p = \boldsymbol{x}_p - a_1\boldsymbol{y}_1 - a_2\boldsymbol{y}_2 - \cdots - a_{p-1}\boldsymbol{y}_{p-1}$ donde los escalares a_i se escogen con la condición $\boldsymbol{y}_p \perp \boldsymbol{y}_i$, $i = 1, 2, \ldots, p-1$.

Corolario 5.2.3. *El sistema $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r\}$ del teorema anterior puede escogerse ortonormal.*

Corolario 5.2.4. *Todo espacio vectorial euclídeo de dimensión finita no trivial admite bases ortonormales.*

Observación. Si aplicamos el método de Gram-Schmidt a un conjunto que no es linealmente independiente, obtenemos el vector cero cuando llegamos a un vector que es combinación lineal de los anteriores.

5.3. Espacios fundamentales de una matriz

Definición 5.3.1 (Espacios fundamentales de una matriz)**.** *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Se definen:*

$\mathcal{F}(A)$ ***El espacio fila de la matriz A*** *como la variedad lineal de $\mathbb{K}^n$ engendrada por los vectores filas de la matriz A, o sea $\mathcal{F}(A) =< \boldsymbol{f}_1,\ \boldsymbol{f}_2,\ \ldots,\ \boldsymbol{f}_m >$.*

$\mathcal{R}(A)$ ***El espacio columna de la matriz A*** *como la variedad lineal de $\mathbb{K}^m$ engendrada por los vectores columnas de la matriz A, o sea $\mathcal{R}(A) =< \boldsymbol{c}_1,\ \boldsymbol{c}_2,\ \ldots,\ \boldsymbol{c}_n >$.*

Nótese que $\mathcal{R}(A) = \{\boldsymbol{x}_1\boldsymbol{c}_1 + \boldsymbol{x}_2\boldsymbol{c}_2 + \cdots + \boldsymbol{x}_n\boldsymbol{c}_n, \}\forall(\boldsymbol{x}_1, \ldots,\ \boldsymbol{x}_n) \in \mathbb{K}^n =$

$$= \begin{pmatrix} \vdots & \vdots & & \vdots \\ c_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix},\ \forall(\boldsymbol{x}_1, \ldots,\ \boldsymbol{x}_n) \in \mathbb{K}^n\} \ = \ \{A\boldsymbol{x} \ / \ \forall \boldsymbol{x} \in \mathbb{K}^n\}.$$

$\mathcal{N}(A)$ ***El espacio nulo de la matriz A*** *como la variedad lineal de $\mathbb{K}^n$ definida por*

$$\mathcal{N}(A) = \{\boldsymbol{x} \in \mathbb{K}^n \ / \ A\boldsymbol{x} = \boldsymbol{0}\}.$$

5.4. Subespacios ortogonales. Proyecciones ortogonales sobre subespacios

Definición 5.4.1 (Subespacio ortogonal)**.** *Sea F un subespacio de un espacio vectorial euclídeo $\mathbb{E}$. Se llama ortogonal de F al conjunto de los vectores que son ortogonales a todos los vectores de F, es decir,*

$$F^{\perp} = \{\boldsymbol{x} \in \mathbb{E} / \ \boldsymbol{x} \cdot \boldsymbol{y} = 0, \ \ \forall \ \boldsymbol{y} \in F\}$$

Propiedades.-

a) $F^{\perp}$ es un subespacio de $\mathbb{E}$.

b) $(F+G)^{\perp} = F^{\perp} \cap G^{\perp} \quad \forall \ F, G \subset \mathbb{E}$

c) $F = (F^{\perp})^{\perp}$

d) Si $F \subset G \Rightarrow G^{\perp} \subset F^{\perp}$

e) $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$

Teorema 5.4.1. *Sea $\mathbb{E}$ un espacio euclídeo, $dim(\mathbb{E}) = n$. Sea F un subespacio de $\mathbb{E}$, entonces $\mathbb{E} = F \oplus F^{\perp}$, es decir:*

$$F \cap F^{\perp} = \{\vec{0}\} \qquad \mathbb{E} = F + F^{\perp}$$

Los subespacios fundamentales de una matriz satisfacen determinadas relaciones de ortogonalidad, veamos cuales son.

Teorema 5.4.2. *Se verifica:*

1. $\mathcal{N}(A) = (\mathcal{F}(A))^{\perp} \qquad \mathcal{F}(A) = (\mathcal{N}(A))^{\perp}$

2. $\mathcal{N}(A^t) = (\mathcal{R}(A))^{\perp} \qquad \mathcal{R}(A) = (\mathcal{N}(A^t))^{\perp}$

5.4.1. Proyecciones ortogonales sobre subespacios

Es bien conocido que para hallar la mínima distancia desde un punto a un plano en el espacio tridimensional, hay que proyectar perpendicularmente dicho punto sobre el plano. En un espacio vectorial euclídeo se verifica un teorema análogo que veremos a continuación.

Definición 5.4.2. *Sean V un espacio vectorial euclídeo, $\boldsymbol{x}$ un vector de V y E un subespacio de V. Se define la **proyección ortogonal de $\boldsymbol{x}$ sobre** E como aquel vector de E, si existe, denotado por $Proy_E(\boldsymbol{x})$, tal que $\boldsymbol{x} - Proy_E(\boldsymbol{x})$ es ortogonal a todos los vectores de E.*

Observación: Este concepto está bien definido porque cuando exista sólo habrá un vector que cumpla estas condiciones.

Definición 5.4.3. *Sea V un espacio vectorial euclídeo. Sea E un subespacio de V de dimensión finita con base ortonormal $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r\}$. Dado un vector $\boldsymbol{x} \in V$ definimos su* ***desarrollo de Fourier*** *con respecto a $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r\}$ como el vector de E dado por*

$$\vec{v} = \mathcal{F}(\boldsymbol{x}) = \sum_{i=1}^{r} \langle \boldsymbol{x}_i, \boldsymbol{x} \rangle \boldsymbol{x}_i$$

El coeficiente $\langle \boldsymbol{x}_i, \boldsymbol{x} \rangle$ recibe el nombre de i-ésimo coeficiente de Fourier de $\boldsymbol{x}$ con respecto al sistema ortonormal dado.

Teorema 5.4.3 (Teorema de la mejor aproximación.)**.** *Sea V espacio vectorial euclídeo. Sea E un subespacio finito dimensional de V con base ortonormal $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r\}$. Dado un vector $\boldsymbol{x} \in V$, sea $\vec{v} = \mathcal{F}(\boldsymbol{x})$ su desarrollo de Fourier correspondiente a la base dada. Entonces:*

- *$\boldsymbol{x} - \vec{v}$ es ortogonal a todos los vectores de E, es decir, el desarrollo de Fourier de $\boldsymbol{x}$ coincide con la Proyección ortogonal de $\boldsymbol{x}$ sobre E.*

$$Proy_E \boldsymbol{x} = \mathcal{F}(\boldsymbol{x}).$$

- *El vector $\vec{v}$ es el vector de E más cercano a $\boldsymbol{x}$, es decir:*

$$||\boldsymbol{x} - \vec{v}|| = \text{mín}\{||\boldsymbol{x} - \boldsymbol{y}|| \,:\, \boldsymbol{y} \in E\},$$

y, además, si $\vec{w}$ es otro vector en E distinto de v, entonces $||\boldsymbol{x} - \vec{v}|| < ||\boldsymbol{x} - \boldsymbol{w}||$.

En particular, $\boldsymbol{x} \in E$ si y sólo si $\boldsymbol{x} = \sum_{i=1}^{r} \langle \boldsymbol{x}_i, \boldsymbol{x} \rangle \boldsymbol{x}_i$.

Teorema 5.4.4. *Sea $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r\}$ un sistema ortonormal en V, $\boldsymbol{x} \in V$ y*

$$\vec{v} = \sum_{i=1}^{r} \langle \boldsymbol{x}_i, \boldsymbol{x} \rangle \boldsymbol{x}_i$$

su desarrollo de Fourier con respecto al sistema dado. Entonces:

- $||\vec{v}||^2 + ||\boldsymbol{x} - \vec{v}||^2 = ||\boldsymbol{x}||^2 = \sum_{i=1}^{r} |\langle \boldsymbol{x}_i, \boldsymbol{x} \rangle|^2 + ||\boldsymbol{x} - \vec{v}||^2$.
- $||\vec{v}||^2 = \sum_{i=1}^{r} |\langle \boldsymbol{x}_i, \boldsymbol{x} \rangle|^2 \leq ||\boldsymbol{x}||^2$ *(Desigualdad de Bessel).*
- $\sum_{i=1}^{r} |\langle \boldsymbol{x}_i, \boldsymbol{x} \rangle|^2 = ||\boldsymbol{x}||^2$ *si y sólo si* $\boldsymbol{x} \in \langle \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r \rangle$ *(Igualdad de Parseval).*

Nota: Obsérvese que la construcción del método de ortonormalización de Gram-Schmidt implica que

$$\mathbf{y}_p = \boldsymbol{x}_p - Pro\mathbf{y}_{\langle \mathbf{y}_1, \cdots, \mathbf{y}_{p-1} \rangle}(\boldsymbol{x}_p),$$

y por tanto

$$||\mathbf{y}_p|| = d\left(\boldsymbol{x}_p, \langle \boldsymbol{x}_1, \cdots, \boldsymbol{x}_{p-1} \rangle\right).$$

5.4.2. Complemento ortogonal de un subespacio vectorial

Definición 5.4.4. *Dado un subespacio E de un espacio vectorial euclídeo V, definimos su* ***ortogonal*** *como*

$$E^{\perp} := \{\boldsymbol{y} \in V \,:\, \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \text{ para todo } \boldsymbol{x} \in E\}.$$

Teorema 5.4.5. *Si E es un subespacio de V (espacio vectorial euclídeo de dimensión finita), entonces $E^{\perp}$ es también un subespacio de V y se cumple $E \bigoplus E^{\perp} = V$ y $E^{\perp\perp} = E$.*

5.5. Aproximación por mínimos cuadrados

Sea A una matriz real $m \times n$ y $\mathbf{b}$ un vector $m \times 1$. Si $m > n$, entonces el sistema $A\boldsymbol{x} = \mathbf{b}$ será, en general, incompatible: hay más ecuaciones que incógnitas. La solución aproximada que buscaremos será aquella que hace más pequeña la distancia entre $A\boldsymbol{x}$ y $\mathbf{b}$; es decir, minimiza el error $E^2 = \|A\boldsymbol{x} - \mathbf{b}\|^2$, que recibe el nombre de **desviación cuadrática**.

Definición 5.5.1 (Solución en el sentido de los mínimos cuadrados)**.** *Diremos que $\boldsymbol{x}_0$ es una* ***solución en el sentido de los mínimos cuadrados*** *para el sistema $A\boldsymbol{x} = \boldsymbol{b}$ si se verifica:*

$$\|A\boldsymbol{x}_0 - \boldsymbol{b}\| \leq \|A\boldsymbol{x} - \boldsymbol{b}\| \qquad \forall\ \boldsymbol{x} \in \mathbb{R}^n$$

ya vimos que el elemento que más se aproxima a un subespacio es su proyección ortogonal sobre él. En consecuencia, el vector $A\boldsymbol{x}_0 - \mathbf{b} \in \mathcal{R}(A)^{\perp}$. Denotaremos por P la proyección ortogonal de $\mathbb{R}^m$ sobre el espacio columna de A, $\mathcal{R}(A)$, es decir, $P\mathbf{b} = A\boldsymbol{x}_0$.

Como $A\boldsymbol{x}_0 - \mathbf{b} \in \mathcal{R}(A)^{\perp}$, y todos los elementos de $\mathcal{R}(A)$ son de la forma $A\mathbf{y}$,$\forall\ \mathbf{y} \in \mathbb{R}^n$, entonces

$$A\mathbf{y} \cdot (A\boldsymbol{x}_0 - \mathbf{b}) = 0 \ \ \forall\ \mathbf{y} \in \mathbb{R}^n$$

Teniendo en cuenta que el producto escalar en $\mathbb{R}^m$ se expresa de la forma $\boldsymbol{x} \cdot \mathbf{y} = \boldsymbol{x}^t \mathbf{y}$ obtenemos la expresión:

$$(A\mathbf{y})^t (A\boldsymbol{x}_0 - \mathbf{b}) = 0$$

de donde se deduce que

$$\mathbf{y}^t (A^t A\boldsymbol{x}_0 - A^t \mathbf{b}) = 0$$

Al ser esta expresión cierta $\forall \mathbf{y} \in \mathbb{R}^n$, entonces

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

Estas ecuaciones, que siempre tienen solución $\boldsymbol{x}_0$, se llaman **Ecuaciones Normales de Gauss**. Una vez visto esto, es inmediato comprobar el siguiente:

Teorema 5.5.1. *Sea A una matriz real $m \times n$ y $\boldsymbol{b} \in \mathbb{R}^m$. Las siguientes condiciones son equivalentes:*

- *$\boldsymbol{x}_0$ es una solución en mínimos cuadrados para el sistema $A\boldsymbol{x} = \boldsymbol{b}$.*

- *$A\boldsymbol{x}_0 = P\boldsymbol{b}$ ($A\boldsymbol{x}_0$ es la proyección de $\boldsymbol{b}$ sobre el espacio columna).*

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

Nótese que:

- Si A^tA es invertible, entonces $\boldsymbol{x}_0 = (A^tA)^{-1}A^t\mathbf{b}$ es la solución única en el sentido de los mínimos cuadrados.

- Si A^tA no es invertible, existen infinitas soluciones en el sentido de los mínimos cuadrados. Ello nos lleva a definir el concepto de **solución óptima**.

5.5.1. Soluciones óptimas

Definición 5.5.2. ***(Solución óptima)*** *Diremos que $\boldsymbol{x}_0$ es una solución óptima en el sentido de los mínimos cuadrados para el sistema $A\boldsymbol{x} = \boldsymbol{b}$ si es una solución en mínimos cuadrados con norma mínima, es decir, $\|\boldsymbol{x}_0\| \leq \|\boldsymbol{x}\|$ para cualquier otra solución en mínimos cuadrados $\boldsymbol{x}$.*

Teorema 5.5.2. *Sea A una matriz real $m \times n$ y $\boldsymbol{b} \in \mathbb{R}^m$. Entonces las siguientes condiciones son equivalentes:*

- *$\boldsymbol{x}_0$ es una solución óptima en mínimos cuadrados para $A\boldsymbol{x} = \boldsymbol{b}$.*

- *$A\boldsymbol{x}_0 = P\boldsymbol{b}$ y $\boldsymbol{x}_0 \in \mathcal{F}(A)$ (o $\mathcal{R}(A^t)$).*

Además, en ese caso, $\boldsymbol{x}_0$ es la única solución óptima.

5.5.2. Aplicaciones. Regresión

Supongamos que tenemos una nube de puntos (x_i, y_i), $i = 1, 2, \ldots, m$ que representa los resultados de un cierto experimento (por ejemplo, x_i es la carga aplicada sobre una cierta estructura e y_i es la deformación que se produce; y_i es la distancia a un satélite que va camino de Marte y x_i el tiempo; x_i los costes de producción en un proceso económico e y_i los volúmenes producidos con los precios y las ganancias). El problema de la regresión lineal consiste en buscar una expresión del tipo $y = ax + b$ que nos permita calcular y conocida x. Si las mediciones no son exactas o bien la función $ax + b$ es sólo una aproximación, es de suponer que, en general, la igualdad $y_i = ax_i + b$ no se va a cumplir para ciertas parejas de datos, cualesquiera que sean a y b. Lo que se hace es calcular a y b de manera que la recta $y = ax + b$ es la que mejor se aproxima a la nube de puntos (x_i, y_i) en el sentido de los mínimos cuadrados.

El sistema a resolver es:

$$\begin{aligned} ax_1 + b &= y_1 \\ ax_2 + b &= y_2 \\ &\vdots \\ ax_m + b &= y_m \end{aligned}$$

que, escrito en forma matricial, queda:

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Si todos los valores de x_i son iguales, la recta es $x = x_1$, que no es de la forma que buscamos.

Supongamos que, al menos, hay dos valores de x_i que son distintos, el rango de la matriz es 2 y, por tanto, la solución óptima es la única solución de las ecuaciones normales de Gauss:

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

es decir,

$$\begin{pmatrix} \displaystyle\sum_{i=1}^{n} x_i^2 & \displaystyle\sum_{i=1}^{n} x_i \\ \displaystyle\sum_{i=1}^{n} x_i & m \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{i=1}^{n} x_i y_i \\ \displaystyle\sum_{i=1}^{n} y_i \end{pmatrix}$$

resolviendo el sistema, obtenemos los valores óptimos a_0 y b_0:

$$a_0 = \frac{m\displaystyle\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{m\displaystyle\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b_0 = \frac{1}{m}\left(\sum_{i=1}^{n} y_i - a_0 \sum_{i=1}^{n} x_i\right)$$

Observaciones.

Algunos ajustes no lineales pueden reducirse a ajustes lineales. Por ejemplo, $\mathbf{y} = ae^{bx}$ se reduce a un ajuste lineal tomando logaritmos: $\log(\mathbf{y}) = \log(a) + bx$.

Los resultados de la sección anterior pueden aplicarse al problema de la regresión polinómica, $\mathbf{y} = a_0 + a_1 x + \cdots + a_n x^n$. En este caso la matriz que se obtiene es de orden $m \times (n+1)$.

Se pueden resolver también ajustes con más variables, como por ejemplo: $z = ax + by + c$.

La recta $y = ax+b$, así calculada, se llama **recta de regresión de** y **sobre** x. Análogamente se puede buscar la **recta de regresión de** x **sobre** y de la forma $x = \alpha y + \beta$. En general, ambas rectas no tienen por qué ser iguales. En la recta de regresión $y = ax + b$ las distancias se miden proyectando el punto verticalmente sobre la recta, mientras que en la recta $x = \alpha y + \beta$ las distancias se miden proyectando horizontalmente (Ver Figura 8.1).

Otra posibilidad sería considerar la distancia a la recta; o sea, buscar la recta r que hace mínima la suma de los cuadrados de las distancias de cada punto a la recta: $\sum_i d((x_i, y_i), r)^2$. Estas distancias se miden proyectando el punto perpendicularmente sobre la recta de la Figura 8.2.

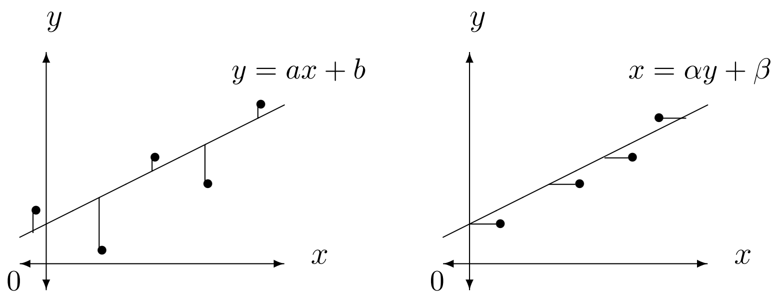

Figura 8.1: Rectas de Regresión

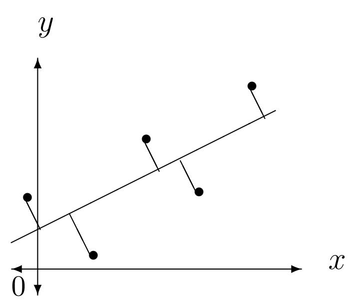

Figura 8.2: Recta de Regresión generalizada

Ejercicios

5.1 Dado el espacio vectorial de los polinomios de grado menor o igual a uno y el producto escalar $p_1(x) \circ p_2(x) = \int_0^1 p_1(x) \cdot p_2(x) dx$
¿Qué ángulo forman los polinomios $x+3$ y $2x+4$?

5.2 Sea V el espacio vectorial de los polinomios en x de grado menor o igual que 2 y una base $B\{\ \mathbf{e}_1 = x^2,\ \mathbf{e}_2 = x,\ \mathbf{e}_3 = 1\}$. Se considera el producto escalar definido por: $\mathbf{e}_i \cdot\ \mathbf{e}_j = \dfrac{1}{i+j+1}$. Se pide:

(a) El ángulo de los vectores 1 y x.

(b) Estudiar para qué valores de a son ortogonales $x+a$ y $x-a$.

5.3 Sea A $\in \mathcal{M}_{m\times n}$. Probar que en $\mathbb{R}^n$ los subespacios $\mathcal{N}(\mathrm{A})$ y $\mathcal{F}(\mathrm{A})$ son complementarios y ortogonales.
Comprobar lo anterior para la matriz $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{pmatrix}$ y descomponer el vector (3,3,3) en suma de un vector de $\mathcal{N}(\mathrm{A})$ y otro de $\mathcal{F}(\mathrm{A})$.

5.4 (a) Hallar las ecuaciones implícitas de $F^\perp$ siendo $F =< (1,0,1),(0,1,-1) >$.

(b) Hallar una base de G siendo las ecuaciones implícitas de $G^\perp$:

$$\begin{cases} x_1 - x_2 + x_3 = 0 \\ x_1 = 0 \end{cases}$$

(c) Hallar una base de $H^\perp$ siendo las ecuaciones implícitas de H: $x_1 - 6x_2 = 0$

5.5 Sean $F^\perp =< (1,-1,2,1),(0,1,-2,1),(1,0,0,2) >$ y $G =< (1,1,-1,0) >$. Hallar:

(a) F y $G^\perp$

(b) $(F+G)^\perp$, $(F\cap G)^\perp$

5.6 Sea $F =< (1,1,1,0),(0,1,-1,0) > \subset \mathbb{R}^4$

(a) Comprobar que $F + F^\perp = \mathbb{R}^4$ y $F \cap F^\perp = \{\mathbf{0}\}$.

(b) Descomponer el vector (0,1,2,0) como suma de un vector de F y otro de $F^\perp$.

(c) Demostrar que dicha descomposición es única.

5.7 Descomponer el vector (1,3,-1,4) en suma de dos vectores $\mathbf{u} + \vec{v}$ siendo $\mathbf{u}$ proporcional a (2,1,0,1) y $\vec{v} \perp \mathbf{u}$.

5.8 Hallar la solución óptima del siguiente sistema

$$\begin{array}{rrl} -x_1 & -2x_2 & = 1 \\ 2x_1 & +4x_2 & = 0 \\ x_1 & +2x_2 & = 0 \\ 3x_1 & +6x_2 & = 0 \end{array}$$

5.9 Tratar de ajustar mediante una recta $y = c + dt$ los puntos $y = 0,\ t = 2$, e $y = 6,\ t = 2$.

5.10 Hallar por el método de los mínimos cuadrados la solución óptima de los sistemas:

$$\text{a)}\ \begin{array}{rrrrl} x_1 & +x_2 & +x_3 & +x_4 & = 2 \\ x_1 & +x_2 & +x_3 & +x_4 & = 3 \\ x_1 & +x_2 & +x_3 & +x_4 & = 4 \end{array} \qquad \text{b)}\ \begin{array}{rrl} x_1 & +x_2 & = 2 \\ x_1 & -x_2 & = 0 \\ 2x_1 & +x_2 & = 2 \end{array}$$

5.11 Para los siguientes puntos, ajustar por el método de los mínimos cuadrados una recta y una parábola , comparando los errores correspondientes:

(a) (-1,-1), (0,1), (1,2) y (2,2)

(b) (-1,0), (0,3), (1,2) y (2,2)

5.12 Una pieza se estira hasta las longitudes de L= 5,6 y 7 metros bajo la aplicación , respectiva, de tres fuerzas de F= 1,2 y 4 toneladas. Suponiendo que se verifica la ley de Hooke $L = a + bF$, encontrar la longitud aproximada de la pieza.

5.13 Dados los cuatro puntos de $\mathbb{R}^3 : (1, 2, \alpha_1), (0, -1, \alpha_2), (1, 2, \alpha_3)$ y $(0, -1, \alpha_4)$,

(a) Encontrar el plano $z = ax + by + c$ que mejor aproxima, en el sentido de los mínimos cuadrados, a dichos puntos.

(b) Hallar las condiciones sobre los α_i que garantizan que el plano pasa por los cuatro puntos.