Tema 7: Estructura euclídea de R^n

1 Producto escalar

Definición. 1.1 Dados $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ definimos el producto escalar de $\mathbf{x}$ e $\mathbf{y}$ como el número

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Proposición. 1.2 Dados $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, se verifica:

1) $\mathbf{u} \cdot \mathbf{u} \geq 0$
2) $\mathbf{u} \cdot \mathbf{u} = 0 \quad \Longleftrightarrow \quad \mathbf{u} = 0$
3) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
4) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
5) $\mathbf{u} \cdot (\alpha \mathbf{v}) = \alpha(\mathbf{u} \cdot \mathbf{v})$

Definición. 1.3 Dado $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ definimos la norma de $\mathbf{u}$ como,

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \left(\sum_{i=1}^{u} u_i^2 \right)^{\frac{1}{2}}.$$

Proposición. 1.4 Dado $\mathbf{u} \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, se verifica:

1) $\|\mathbf{u}\| \geq 0$
2) $\|\mathbf{u}\| = 0 \quad \Longleftrightarrow \quad \mathbf{u} = 0$
3) $\|\alpha \mathbf{u}\| = |\alpha| \|\mathbf{u}\|$
4) $\dfrac{1}{\|\mathbf{u}\|} \mathbf{u}$ es un vector unitario, es decir, de norma 1.

Proposición. 1.5 (Desigualdad de Cauchy-Schwartz)

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \quad |\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

es decir, si $\mathbf{u} = (u_1, \ldots, u_n)$ y $\mathbf{v} = (v_1, \ldots, v_n)$,

$$\left| \sum_{i=1}^{n} u_i v_i \right| \leq \left(\sum_{i=1}^{n} u_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} v_i^2 \right)^{\frac{1}{2}}.$$

Además, dicha desigualdad es una igualdad si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son proporcionales.

Teorema. 1.6 (Desigualdad triangular)

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \quad \|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Además, dicha desigualdad es una igualdad si y sólo si existe $\alpha \in \mathbb{R}$, $\alpha \geq 0$ tal que $\mathbf{u} = \alpha \mathbf{v}$.

Corolario. 1.7 (Teorema de Pitágoras)

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \quad \mathbf{u} \cdot \mathbf{v} = 0 \quad \Longleftrightarrow \quad \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

2 Ortogonalidad

Definición. 2.1 Dados $\mathbf{u} \neq \mathbf{0}, \mathbf{v} \neq \mathbf{0} \in \mathbb{R}^n$, definimos el ángulo entre $\mathbf{u}$ y $\mathbf{v}$ como el número $\alpha = \widehat{(\mathbf{u}, \mathbf{v})}$ dado por:

$$\cos\alpha = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$$

Definición. 2.2 Dados $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, decimos que $\mathbf{u}$ y $\mathbf{v}$ son ortogonales, y lo denotamos por $\mathbf{u} \perp \mathbf{v}$, si $\mathbf{u} \cdot \mathbf{v} = 0$.

Lema. 2.3 Dados $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$,

$$\mathbf{u} \perp \mathbf{v} \quad \Longleftrightarrow \quad \widehat{(\mathbf{u}, \mathbf{v})} = \frac{\pi}{2}.$$

Definición. 2.4 Dados $\mathbf{u} = (u_1, \dots, u_n), \mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$, definimos la distancia entre $\mathbf{u}$ y $\mathbf{v}$ como el número

$$dist(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \left(\sum_{i=1}^{n} (u_i - v_i)^2\right)^{\frac{1}{2}}.$$

3 Bases ortonormales. Método de Gram-Schmidt

Definición. 3.1 Un conjunto $\{\mathbf{a}_1, \dots \mathbf{a}_r\} \subseteq \mathbb{R}^n$ es un sistema ortogonal si

$$\forall i, j = 1, \dots r, \; i \neq j, \quad \mathbf{a}_i \cdot \mathbf{a}_j = 0.$$

Definición. 3.2 Un conjunto $\{\mathbf{a}_1, \dots \mathbf{a}_r\} \subseteq \mathbb{R}^n$ es un sistema ortonormal si

$$\forall i, j = 1, \dots, r \quad \mathbf{a}_i \cdot \mathbf{a}_j = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases}$$

En otras palabras, el conjunto dado es ortonormal si es ortogonal y todos sus elementos tienen norma 1.

Proposición. 3.3 Si $\{\mathbf{a}_1, \dots \mathbf{a}_r\} \subseteq \mathbb{R}^n$ es un sistema ortonormal entonces es linealmente independiente.

Nota. 3.4 La proposición es válida también para sistemas ortogonales, con tal de ser todos los vectores distintos de $\mathbf{0}$.

Definición. 3.5 Sea $W \subseteq \mathbb{R}^n$ una v.l. y $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ una base de W. Diremos que $\mathcal{B}$ es una base ortogonal de W, abreviadamente B.O.G., si $\mathcal{B}$ es un sistema ortogonal. Análogamente, diremos que $\mathcal{B}$ una base ortonormal de W, abreviado B.O.N., si es un sistema ortonormal.

Proposición. 3.6 Sea $W \subseteq \mathbb{R}^n$ una v.l. y $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ una base de W. Sea $A = (a_{ij}) \in \mathcal{M}_r$ la matriz simétrica definida por

$$\forall i, j = 1, \dots r, \quad a_{ij} = \mathbf{u}_i \cdot \mathbf{u}_j.$$

Entonces, dados $\mathbf{u}, \mathbf{v} \in W$ con $\mathbf{u}_{\mathcal{B}} = (\alpha_1, \dots, \alpha_r)$ $\mathbf{v}_{\mathcal{B}} = (\beta_1, \dots, \beta_r)$ se tiene

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i,j=1}^{r} a_{ij}\alpha_i\beta_j = \mathbf{u}_{\mathcal{B}}^t A \mathbf{v}_{\mathcal{B}}.$$

En particular, si $\mathcal{B}$ es B.O.N., tenemos

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{r} \alpha_i \beta_i = \mathbf{u}_{\mathcal{B}}^t \mathbf{v}_{\mathcal{B}}.$$

A continuación veremos un método para obtener una base ortonormal a partir de una base cualquiera.

Método de ortonormalización de Gram-Schmidt:

Sea $W \subseteq \mathbb{R}^n$ una v.l. y $\mathcal{B} = \{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ una base de W. El método de Gram-Schmidt construye una B.O.N. de W, $\mathcal{B}' = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$, tal que

$$\begin{array}{l} L(\{\mathbf{a}_1\}) = L(\{\mathbf{u}_1\}) \\ L(\{\mathbf{a}_1, \mathbf{a}_2\}) = L(\{\mathbf{u}_1, \mathbf{u}_2\}) \\ \vdots \\ L(\{\mathbf{a}_1, \dots, \mathbf{a}_r\}) = L(\{\mathbf{u}_1, \dots, \mathbf{u}_r\}). \end{array}$$

Proposición. 3.7 Dada $W \subseteq \mathbb{R}^n$ una variedad lineal y una base $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$, construimos una base ortogonal $\{\mathbf{z}_1, \dots, \mathbf{z}_r\}$, tal que $L(\mathbf{a}_1) = L(\mathbf{z}_1), \dots, L(\mathbf{a}_1, \dots, \mathbf{a}_i) = L(\mathbf{z}_1, \dots, \mathbf{z}_i)$

Para obtener una BON basta con tomar:

$$\mathbf{u}_k = \frac{1}{\|\mathbf{z}_k\|} \mathbf{z}_k, k = 1, \dots, r$$

4 Cambio de bases ortonormales. Matrices ortogonales

Definición. 4.1 Diremos que $A \in \mathcal{M}_n$ es ortogonal si $A^t = A^{-1}$, es decir,

$$A^t A = A A^t = I.$$

Proposición. 4.2 Sea $A = [\mathbf{a}_1 | \dots | \mathbf{a}_n] \in \mathcal{M}_n$ ortogonal. Entonces,

$$\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$$

es una base ortonormal de $\mathbb{R}^n$.

Proposición. 4.3 Si $A \in \mathcal{M}_n$ es una matriz ortogonal sus filas forman una base ortonormal de $\mathbb{R}^n$.

Teorema. 4.4 Sean $\mathcal{B}$ y $\mathcal{B}'$ dos bases ortonormales de $\mathbb{R}^n$ y $A \in \mathcal{M}_n$ la matriz de cambio de base,

$$\mathbf{x}_{\mathcal{B}} = A \mathbf{x}_{\mathcal{B}'}.$$

Entonces, A es ortogonal.

5 Complemento ortogonal de una variedad

Definición. 5.1 Sean W y W' dos subconjuntos de $\mathbb{R}^n$. Diremos que W y W' son ortogonales, y lo denotamos por $W \perp W'$, si:

$$\forall \mathbf{x} \in W, \ \forall \mathbf{y} \in W', \quad \mathbf{x} \cdot \mathbf{y} = 0.$$

Proposición. 5.2 Se verifica:

1) Si $\mathbf{x} \in \mathbb{R}^n$ y $W = L(\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}) \subseteq \mathbb{R}^n$ entonces,

$$\{\mathbf{x}\} \perp W \qquad \Longleftrightarrow \qquad \forall i = 1, \ldots, m, \quad \mathbf{x} \cdot \mathbf{a}_i = 0.$$

2) Si $H,\ H' \subseteq \mathbb{R}^n, \quad H \perp H' \quad \Longleftrightarrow \quad L(H) \perp L(H')$.

3) Si L y W son v.l. de $\mathbb{R}^n$,

$$L \perp W \qquad \Longrightarrow \qquad L \cap W = \{\mathbf{0}\}.$$

Definición. 5.3 Dado $H \subseteq \mathbb{R}^n$, definimos el complemento ortogonal de H como

$$H^{\perp} = \{\mathbf{x} \in \mathbb{R}^n : \ \forall \mathbf{y} \in H,\ \mathbf{x} \cdot \mathbf{y} = 0\}.$$

Proposición. 5.4 Sean $H, G \subseteq \mathbb{R}^n$. Se verifica:

1) $H^{\perp}$ es una v.l. de $\mathbb{R}^n$.

2) $H \perp H^{\perp}$ y $H \subseteq (H^{\perp})^{\perp}$.

3) $H^{\perp}$ es el mayor subespacio de $\mathbb{R}^n$ ortogonal a H.

4) $H \subseteq G \quad \Longrightarrow \quad G^{\perp} \subseteq H^{\perp}$.

Teorema. 5.5 Si $L \subseteq \mathbb{R}^n$ es una v.l. entonces

$$\mathbb{R}^n = L \oplus L^{\perp}.$$

En particular, $dim(L^{\perp}) = n - dim(L)$ y $(L^{\perp})^{\perp} = L$.

Proposición. 5.6 Dada una v.l. $L \subseteq \mathbb{R}^n$, $L^{\perp}$ es la única variedad lineal de $\mathbb{R}^n$ tal que:

$$\mathbb{R}^n = L \oplus L^{\perp} \quad \text{y} \quad L \perp L^{\perp}.$$

En particular, si L es una variedad lineal $(L^{\perp})^{\perp} = L$.

6 Proyección ortogonal

Definición. 6.1 Dada $L \subseteq \mathbb{R}^n$ v.l. y $\mathbf{x} \in \mathbb{R}^n$, puesto que $\mathbb{R}^n = L \oplus L^{\perp}$, existen unos únicos $\mathbf{x}_1 \in L$ y $\mathbf{x}_2 \in L^{\perp}$, tales que

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2.$$

$\mathbf{x}_1$ se denomina proyección ortogonal de $\mathbf{x}$ sobre L.

Teorema. 6.2 Sea $L \subseteq \mathbb{R}^n$ una v.l. y $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}_1 \in L$. Son equivalentes:

1) $\mathbf{x}_1$ es la proyección ortogonal de $\mathbf{x}$ sobre L.

2) $\forall \mathbf{y} \in L, \quad \mathbf{x} \cdot \mathbf{y} = \mathbf{x}_1 \cdot \mathbf{y}$.

3) $\forall \mathbf{y} \in L,\ \|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x} - \mathbf{x}_1\|$.
La igualdad se da si sólo si $\mathbf{y} = \mathbf{x}_1$.

Nota. 6.3 *Dados* $\mathbf{x} \in \mathbb{R}^n$ *y una v.l.* $L \subseteq \mathbb{R}^n$, *sea* $\mathbf{x}_1$ *la proyección ortogonal de* $\mathbf{x}$ *sobre* L, *entonces, para todo* $\mathbf{y} \in L$ *tenemos*

$$dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x} - \mathbf{x}_1\| = dist(\mathbf{x}, \mathbf{x}_1),$$

es decir, **la proyección ortogonal de $\mathbf{x}$ sobre L es el vector de L más próximo a $\mathbf{x}$**.

Definición. 6.4 Dada una v.l. $L \subseteq \mathbb{R}^n$ y $\mathbf{x} \in \mathbb{R}^n$ decimos que $\mathbf{u} \in L$ da la distancia mínima de $\mathbf{x}$ a L si:

$$\forall \mathbf{v} \in L, \quad \|\mathbf{x} - \mathbf{u}\| \leq \|\mathbf{x} - \mathbf{v}\|.$$

Definimos la distancia de $\mathbf{x}$ a L como

$$dist(\mathbf{x}, L) = \|\mathbf{x} - \mathbf{u}\| = \min\{\|\mathbf{x} - \mathbf{v}\| : \ \mathbf{v} \in L\}.$$

Nota. 6.5 *Acabamos de probar que la proyección ortogonal es* **el vector** *(es único) que da la distancia mínima.*

7 Pseudosoluciones

Definición. 7.1 Dado un S. E. L. $(S) : A\mathbf{x} = \mathbf{b}$, con $A \in \mathcal{M}_{m\times n}$, diremos que $\mathbf{x}_0 \in \mathbb{R}^n$ es una pseudosolución de (S) si:

$$\forall \mathbf{u} \in \mathbb{R}^n, \quad \|A\mathbf{u} - \mathbf{b}\| \geq \|A\mathbf{x}_0 - \mathbf{b}\|.$$

Proposición. 7.2 Sea $(S) : A\mathbf{x} = \mathbf{b}$, con $A \in \mathcal{M}_{m\times n}$ y $\mathbf{b} \in \mathbb{R}^n$. Se verifica:

1) Si (S) es compatible, toda solución de (S) es una pseudosolución.
2) (S) es compatible $\iff$ $\mathbf{b} \in C(A)$.
3) $\mathbf{u}_0 \in \mathbb{R}^n$ es una pseudosolución de (S) si y sólo si $A\mathbf{u}_0$ es la proyección ortogonal de $\mathbf{b}$ sobre $C(A)$.

Definición. 7.3 Dado $(S) : A\mathbf{x} = \mathbf{b}$, llamaremos sistema normal asociado a (S), al sistema:

$$(S') : A^tA\mathbf{x} = A^t\mathbf{b}.$$

Proposición. 7.4 Dado $(S) : A\mathbf{x} = \mathbf{b}$, $\mathbf{u}_0 \in \mathbb{R}^n$ es una pseudosolución de (S) si y sólo si es una solución del sistema normal asociado, es decir,

$$\mathbf{u}_0 \text{ es pseudosolución} \quad \iff \quad A^tA\mathbf{u}_0 = A^t\mathbf{b}.$$

En particular, deducimos que el sistema normal asociado siempre es compatible.

Puesto que las pseudosoluciones de un sistema son la soluciones del sistema normal asociado (que siempre es compatible) en general, pueden existir más de una pseudosolución. Los siguientes resultados aclaran cuando es así.

Lema. 7.5 Dada $A \in \mathcal{M}_{m\times n}$ se verifica:

$$r(A) = r(A^tA).$$

Proposición. 7.6 Dados $A \in \mathcal{M}_{m\times n}$, $(S) : A\mathbf{x} = \mathbf{b}$ y $(S') : A^tA\mathbf{x} = A^t\mathbf{b}$ el sistema normal asociado, se tiene:

- Si $r(A) = n$ entonces (S') es compatible determinado y, por tanto, (S) posee una única pseudosolución
$$\mathbf{x}_0 = (A^t A)^{-1} A^t \mathbf{b}.$$
- Si $r(A) < n$ entonces (S') es compatible indeterminado y, por tanto, existen infinitas pseudosoluciones.

De este modo, en ciertas ocasiones existen infinitas pseudosoluciones para un S.E.L. dado, por lo que podemos preguntarnos cual de ellas es, en cierto sentido, la mejor. A continuación veremos un criterio para elegir dicha pseudosolución.

Definición. 7.7 Dado $(S) : A\mathbf{x} = \mathbf{b}$ con $A \in \mathcal{M}_{m\times n}$ llamaremos pseudosolución óptima de (S) a la pseudosolución, $\mathbf{x}_0$, de norma mínima, es decir,
$$\|\mathbf{x}_0\| = \text{mínimo}(\{\|\mathbf{x}\| : \ A^t A\mathbf{x} = A^t\mathbf{b}\}).$$

Proposición. 7.8 Dada $A \in \mathcal{M}_{m\times n}$ se verifica
$$N(A) = F(A)^{\perp}.$$
En particular, $\mathbb{R}^n = F(A) \oplus N(A)$.

Teorema. 7.9 La pseudosolución óptima (es decir, de norma mínima) del sistema $(S) : A\mathbf{x} = \mathbf{b}$, es la única psedosolución $\mathbf{x}_0$ de (S) tal que
$$\mathbf{x}_0 \in F(A).$$
Más aún, $\mathbf{x}_0$ es la proyección ortogonal sobre $F(A)$ de una pseudosolución cualquiera de (S).

8 Cálculo de la proyección ortogonal

Sea $L \subseteq \mathbb{R}^n$ una variedad lineal con $dim(L) = r$ y $\mathbf{b} \in \mathbb{R}^n$. Queremos hallar la proyección ortogonal de $\mathbf{b}$ sobre L. Para ello disponemos de dos métodos.

1) Sea $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ una B.O.N. de L. Entonces podemos razonar como sigue:
   Puesto que $\mathbf{b} \in \mathbb{R}^n = L \oplus L^{\perp}$, tenemos
   $$\mathbf{b} = \mathbf{b}_0 + \mathbf{v},$$
   siendo $\mathbf{b}_0 \in L$ la proyección ortogonal de $\mathbf{b}$ sobre L y $\mathbf{v} \in L^{\perp}$. Puesto que $\mathcal{B}$ es B.O.N., se tiene
   $$\mathbf{b}_0 = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i$$
   donde $\forall i = 1, \ldots, r, \quad \lambda_i = \mathbf{b}_0 \cdot \mathbf{u}_i = \mathbf{b} \cdot \mathbf{u}_i$. Luego,
   $$\mathbf{b}_0 = \sum_{i=1}^{r} (\mathbf{b} \cdot \mathbf{u}_i)\mathbf{u}_i.$$

2) Supongamos ahora que poseemos una base $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ no necesariamente ortogonal, entonces la matriz
   $$A = [\mathbf{a}_1 | \ldots | \mathbf{a}_r] \in \mathcal{M}_{n\times r}$$
   es tal que $r(A) = r$ y $C(A) = L$, por ser $\mathcal{B}$ una base de L. En consecuencia, $(S) : A\mathbf{x} = \mathbf{b}$, posee una única pseudosolución $\mathbf{x}_0 \in \mathbb{R}^r$. En tales circunstancias, sabemos que $A\mathbf{x}_0$ es la proyección ortogonal de $\mathbf{b}$ sobre $C(A) = L$.

   Resumiendo, podemos calcular la proyección ortogonal de $\mathbf{b}$ sobre L del siguiente modo:

1) Obtener una base $\mathcal{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ de L.
2) Plantear el sistema $(S) : A\mathbf{x} = \mathbf{b}$ con $A = [\mathbf{a}_1|\dots|\mathbf{a}_r] \in \mathcal{M}_{n\times r}$.
3) Obtener una pseudosolución $\mathbf{x}_0$ de (S).
4) $\mathbf{b}_0 = A\mathbf{x}_0$ es la proyección ortogonal de $\mathbf{b}$ sobre L.

9 Ajuste de datos por mínimos cuadrados

Supongamos que tenemos cierta cantidad de datos

$$(t_1, y_1),\ (t_2, y_2), \dots, (t_m, y_m),$$

siendo $t_i \neq t_j$ para $i \neq j$, y queremos hallar la "mejor" relación del tipo $y = at + b$ a la hora de describirlos. Es decir, pretendemos que, para cada t_i, el valor

$$at_i + b$$

sea lo mas parecido posible (en algún sentido) al valor y_i. Una forma de proceder es la siguiente:

Consideremos el S.E.L.

$$(S) : \left\{ \begin{array}{rcl} y_1 & = & t_1 a + b \\ y_2 & = & t_2 a + b \\ \vdots & & \\ y_m & = & t_m a + b \end{array} \right.$$

Matricialmente,

$$(S) : \begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

donde a y b son incognitas.

Normalmente (S) será incompatible y, por tanto, deberemos conformarnos con sus pseudosoluciones, esto es, las soluciones del sistema normal asociado

$$(S') : \begin{bmatrix} \sum_{i=1}^m t_i^2 & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & m \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m t_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix}$$

Una solución (a_0, b_0) de (S') tiene la propiedad de hacer

$$\left\| \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \right\|$$

mínima, en otras palabras

$$E^2 = \sum_{i=1}^m [(a_0 t_i + b_0) - y_i]^2$$

es mínimo.

Geométricamente, $y = a_0 t + b_0$ nos da la recta que mejor se ajusta a los puntos $(t_1, y_1), \dots, (t_m, y_m)$.

Nota. 9.1 *El ajuste de datos por el método de mínimos cuadrados no se restringe sólo a rectas, puesto que, en general, podemos ajustar datos por mínimos cuadrados utilizando para ello parábolas, polinomios de grado fijo u otro tipo de funciones. Por ejemplo, si pretendemos ajustar los datos*

$$(t_1, y_1),\ (t_2, y_2), \ldots, (t_m, y_m),$$

mediante una parábola,

$$y = at^2 + bt + c,$$

bastará considerar el sistema

$$(S) : \begin{bmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ \vdots & \vdots & \vdots \\ t_m^2 & t_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

y razonar como antes, resolviendo el sistema normal asociado (S').