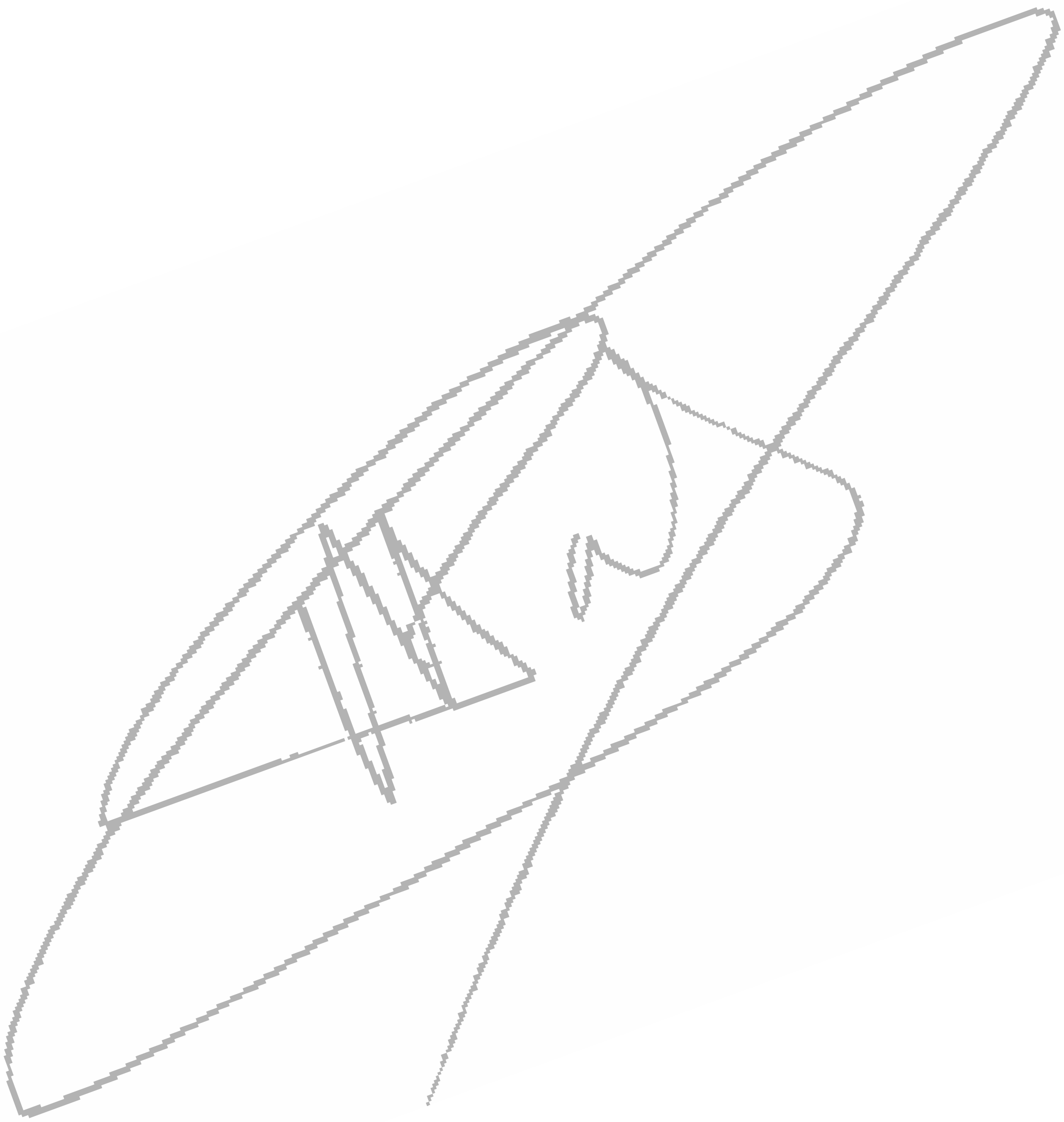
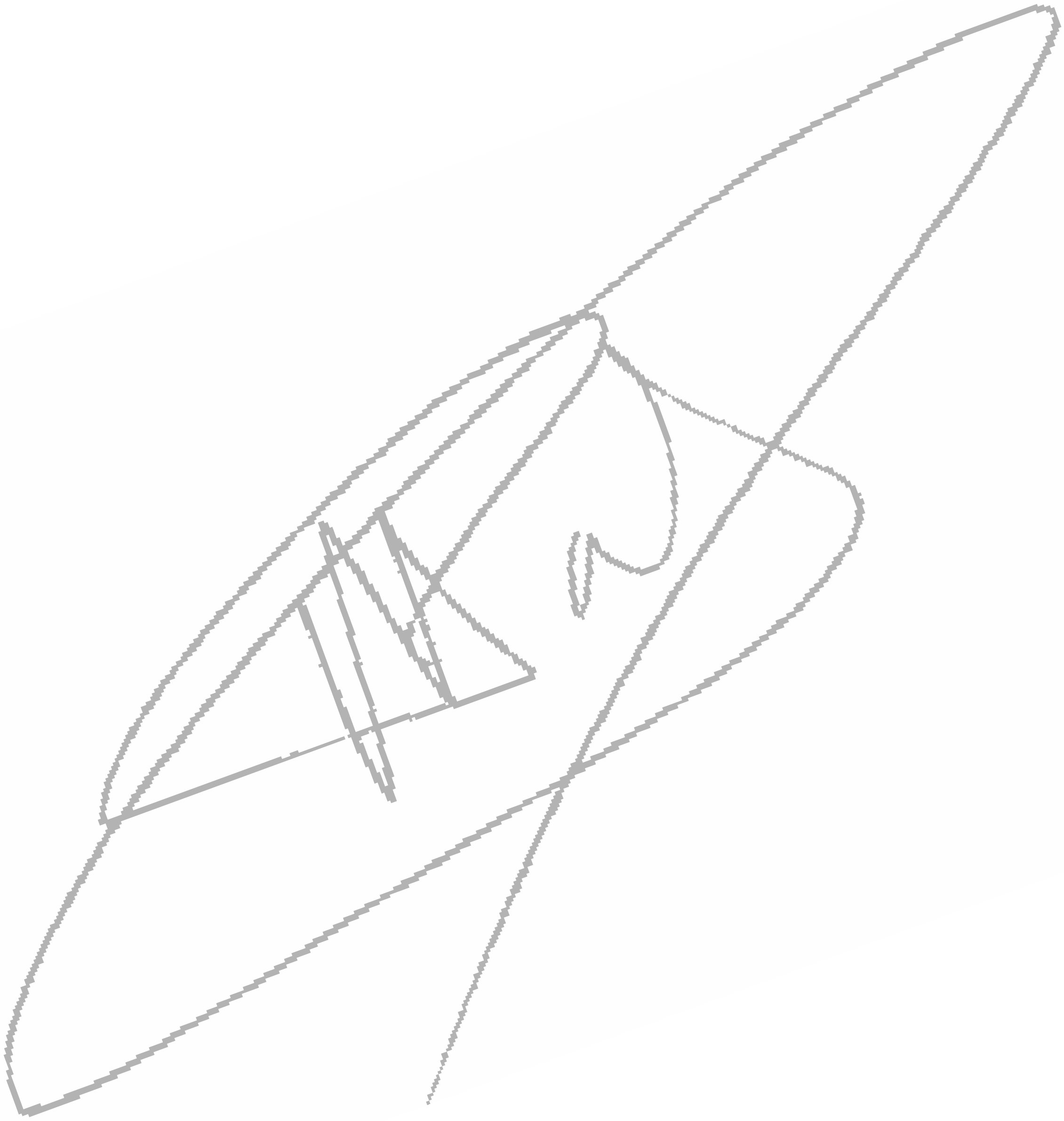
8. Espacios con producto escalar.

Manuel Palacios
Departamento de Matemática Aplicada
Centro Politécnico Superior
Universidad de Zaragoza

Otoño 2010



Contents



Bibliography

[1] Merino, L. y Santos, E.: Álgebra lineal con métodos elementales. *Thomson*. 2006.

[2] Palacios, M.: Enunciados de ejercicios de Matemáticas II, grado Ing. Tecn. Industr. http://pcmap.unizar.es/∼mpala. 2010

[3] Villa, A. de la: Problemas de Álgebra. *CLAGSA*. 1988.

[4] Arvesu, J.; Marcellán, F. y Sánchez J.: Problemas resueltos de Álgebra lineal. Paso a paso. *Thomson*. 2005.

[5] Burden, R. L. and Faires, J. D.: Análisis Numérico. (Fifth ed.) *PWS-Kent Publ. Co.*, 2006.

[6] Rojo, J.: Álgebra lineal. *Vector Ediciones, Madrid*. 1989.

[7] Rojo, J. y Martín, I.: Ejercicios y problemas de Álgebra lineal. *Editorial AC, Madrid*. 1989

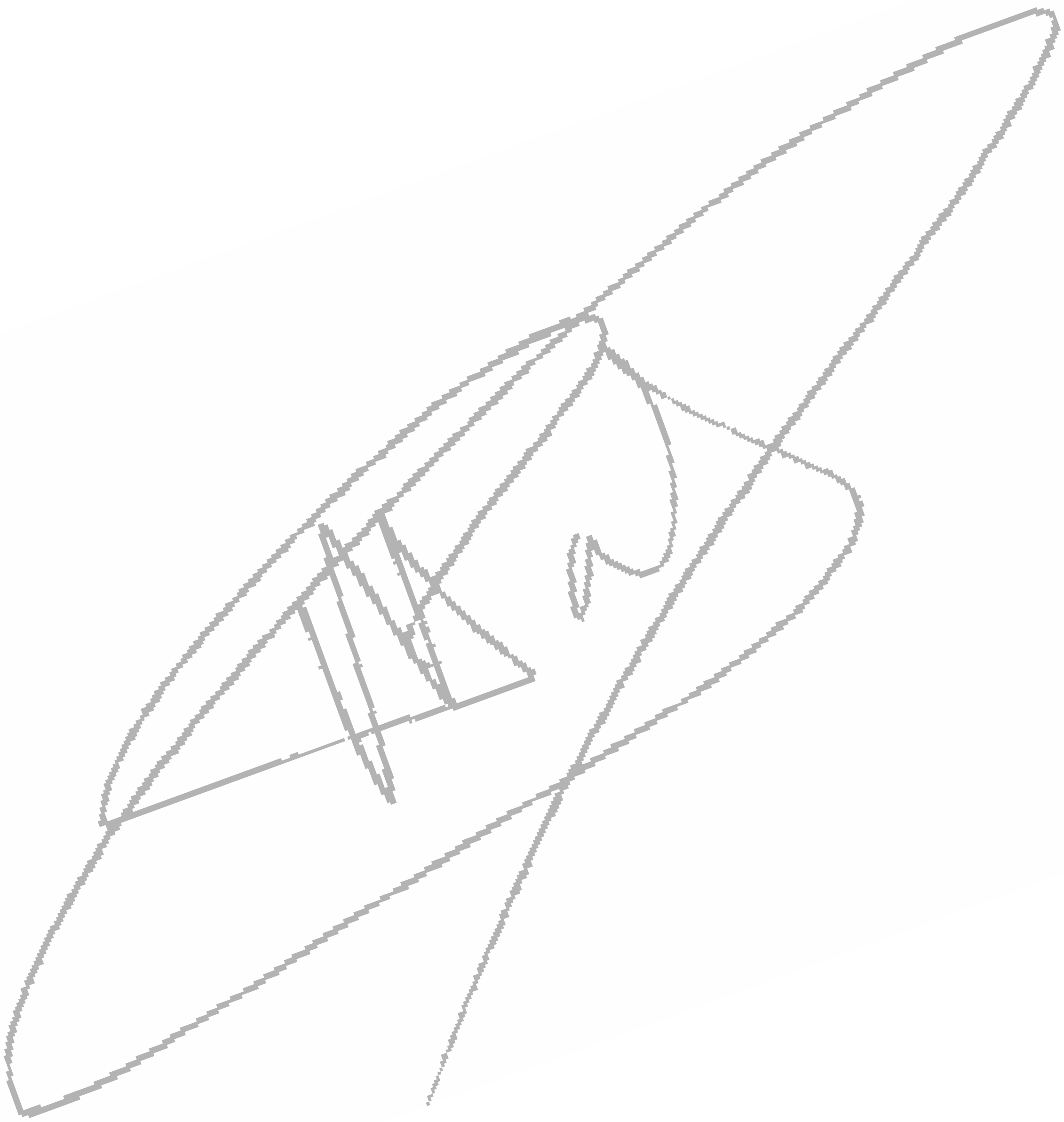
[8] Burgos, J. de: Álgebra lineal y geometría cartesiana. *McGraw-Hill*. 2000.

[9] Burgos, J. de: Álgebra lineal. *McGraw-Hill*. 1993.

[10] Diego, B. de et al. : Problemas de Álgebra lineal (2^a ed.). *Deimos*. 1986.

[11] Griffel, D. G.: Linear Algebra and its Applications. (Vol. 1 & 2) *Ellis Horwood*. 1989.

[12] Hernández, E.: Álgebra y geometría. *Addison-Wesley, Universidad Autónoma de Madrid*. 1994.



Chapter 8

Espacios con producto escalar.

En este capítulo introducimos una forma de medir en un espacio vectorial, es decir, definimos el concepto general de medida o métrica y el de ángulo, lo que nos permitirá generalizar también los conceptos de área y de volumen. El ejemplo que nos motiva es el espacio euclídeo ordinario con el producto escalar ordinario. En lo sucesivo, consideraremos un espacio vectorial V sobre $\mathbb{R}$.

8.1 Definiciones y consecuencias.

Definición 8.1.1 *Llamaremos* **producto escalar** *sobre V a toda forma $F : V \times V \longrightarrow \mathbb{R}$ que sea bilineal simétrica definida positiva.*

Definición 8.1.2 *Llamaremos* **espacio (vectorial) euclídeo** *a cualquier espacio vectorial V sobre $\mathbb{R}$ de dimensión finita sobre el que se ha definido un producto escalar, es decir, al par $(V,\ F)$.*

Ejemplo 8.1.3 *En el espacio vectorial $\mathcal{M}_{\mathbb{R}}(n)$ de las matrices cuadradas de orden n sobre $\mathbb{R}$ la aplicación:*

$$F(A, B) = tr(A\, B^T)$$

es un producto escalar.

Ejercicio 8.1.4 *Probar que la aplicación $\mathbb{R}_2[x] \times \mathbb{R}_2[x] \longrightarrow \mathbb{R}$ definida por*

$$F(p(x), q(x)) = \int_a^b w(x)\, p(x)\, q(x)\, dx, \ \ con\ w(x) \geq 0, a = -1, b = 1$$

es un producto escalar sobre $\mathbb{R}_2[x]$

Definición 8.1.5 *Llamaremos* **matriz coordenada** *de un producto escalar respecto de una base a la matriz coordenada de la aplicación F que lo define.*

Ejemplo 8.1.6 *Las matrices coordenadas del producto escalar del ejercicio 8.1.4 respecto de la base canónica y respecto de la base $\{1, x, -1 + 3\, x^2\}$ en el caso $w(x) = 1$ son, respectivamente:*

$$A = \begin{bmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 8/5 \end{bmatrix}$$

que se corresponden con las expresiones coordenadas siguientes:

$$F(x,y) = X^T A Y = 2\,x_1\,y_1 + 2/3\,x_1\,y_3 + 2/3\,x_2\,y_2 + 2/3\,x_3\,y_1 + 2/5\,x_3\,y_3$$
$$F(x,y) = \bar{X}^T B \bar{Y} = 2\,\bar{x}_1\,\bar{y}_1 + 2/3\,\bar{x}_2\,\bar{y}_2 + 8/5\,\bar{x}_3\,\bar{y}_3$$

Como se ve, una es diagonal y otra no según sea la base elegida.

8.1.1 Norma, distancia, ángulo

Recordando la definición de longitud de un vector libre, damos la siguiente

Definición 8.1.7 *Llamaremos* **norma (euclídea)** *(o longitud) de un vector x de un espacio euclídeo, y la denotaremos por $||x||$, a la aplicación:*

$$||\cdot|| : V \longrightarrow \mathbb{R}^+\{0\}, \text{ definida por } ||\,x\,|| = +\sqrt{F(x,x)}$$

Notemos que:

- $\forall x \in V,$ existe $||\,x\,||$, ya que $F(x,x) \geq 0$
- El único vector con $||\,x\,|| = 0$ es $x = 0$

Definición 8.1.8 *Si $||\,x\,|| = 1$ diremos que x es* **unitario** *o que está normalizado.*

Propiedad 8.1.9 *(Desigualdad de Schwarz)*

$$|F(x,y)| \leq ||x||\,||y||$$

Dem.: Basta tener en cuenta:

$$0 \leq F(t\,x+y, t\,x+y) = t^2\,F(x,x) + 2\,t\,F(x,y) + F(y,y) = t^2\,||x||^2 + 2\,t\,F(x,y) + ||y||^2 \Longrightarrow$$
$$\Longrightarrow \Delta = F(x,y)^2 - ||x||^2\,||y||^2 \leq 0. \ \blacksquare$$

Observemos que la igual en la desigualdad anterior se verifica si x e y son linealmente dependientes.

Propiedad 8.1.10 *(Propiedades características)*
1) $||x|| \geq 0$
2) $||x|| = 0 \Longleftrightarrow x = 0$
3) $||t\,x|| = |t|\,||x||$
4) $||x+y|| \leq ||x||\,||y||$ *(desigualdad triangular)*

Dem.:
4) $||x+y||^2 = ||x||^2 + 2\,F(x,y) + ||y||^2 \leq ||x||^2 + 2\,||x||\,||y|| + ||y||^2 = (||x|| + ||y||)^2 \ \blacksquare$

Ejemplo 8.1.11 *En el ejemplo 8.1.6, la norma del vector* $v = 1 + x + x^2$ *es*

$$||v|| = (\int_{-1}^{1} (1 + x + x^2)^2 \, dx)^{1/2} = ((1\,1\,1)\,A\,(1\,1\,1)^T)^{1/2} = (\frac{22}{5})^{1/2}$$

Observemos que también se pueden definir normas no euclídeas, es decir, no asociadas a un producto escalar. Por ejemplo, la norma de Chebyshev o la norma uniforme. (ver capítulo 9).

Ejercicio 8.1.12 *Probar que* $||x+y||^2 + ||x-y||^2 = 2\,||x||^2 + 2\,||y||^2$ *(Prop. del paralelogramo)*

Ejercicio 8.1.13 *Probar que* $||x+y||^2 = ||x||^2 + ||y||^2$*, si F(x, y) = 0 (Teorema de Pitágoras)*

A partir de una norma se puede definir la distancia entre vectores

Definición 8.1.14 *Llamaremos* **distancia** *del vector x al vector y a la norma del vector diferencia, es decir,*

$$d(x, y) = ||x - y||$$

Propiedad 8.1.15 .

1) $d(x, y) \geq 0$

2) $d(x, y) = 0 \Longleftrightarrow x = y$

3) $d(x, y) = d(y, x)$

4) $d(x, y) \leq d(x, z) + d(z, y)$

La distancia en un conjunto permite definir los conjuntos abiertos o cerrados, es decir, la topología. También se pueden definir distancias no euclídeas (que verifiquen las cuatro propiedades anteriores).

Teniendo en cuenta la desigualdad de Schwarz escrita en la forma siguiente:

$$-1 \leq \frac{F(x, y)}{||x||\,||y||} \leq 1$$

Definición 8.1.16 *Llamaremos* **ángulo** *de dos vectores x, y al número real* $ang(x, y) = \alpha \in [-\pi, \pi]$ *tal que*

$$\cos\alpha = \frac{F(x, y)}{||x||\,||y||}$$

Ejemplo 8.1.17 *Con el producto escalar F del ejemplo 8.1.6, el ángulo entre* $v = 1 + x + x^2$ *y* $w = 1 + x$ *es* α *tal que*

$$\cos\alpha = \frac{F(x, y)}{||x||\,||y||} = \frac{10/3}{(56/15)\,(8/3)} = \frac{75}{224}$$

Consecuencia 8.1.18 .

1) $F(x, y) = ||x||\,||y||\,\cos\alpha$

2) El vector proyección de y sobre x es:

$$pro\vec{y}_x y = (||y||\,\cos\alpha)\,\frac{x}{||x||} = \frac{F(x, y)}{||x||}\,\frac{x}{||x||}$$

3) $||x - y||^2 = ||x||^2 + ||y||^2 - 2\,||x||\,||y||\,\cos\alpha$ *(Teor. del coseno)*

8.2 Bases ortonormadas.

Recordando que dos vectores del espacio ordinario son ortogonales o perpendiculares si forman un ángulo de $\pi/2$ radianes, lo que implica que su producto escalar es nulo, y admitiendo, además, que el vector nulo es ortogonal a todo otro vector, podemos dar la siguiente definición.

Definición 8.2.1 *En un espacio vectorial euclídeo V se dice que dos* **vectores** *x, y son* **ortogonales** *si $F(x,y)=0$*

Ejemplo 8.2.2 *Consideremos el espacio vectorial euclídeo de las funciones continuas; aquí, las funciones sen 2pnt y cos 2pmt son ortogonales, ya que*

$$F(\operatorname{sen} 2\pi nt, \cos 2\pi mt) = \int_{-\pi}^{\pi} \operatorname{sen} 2\pi nt \, \cos 2\pi mt \, dt = 0, \forall n, m = 1, 2, 3, ...$$

Definición 8.2.3 *Una* **familia** *de vectores $\{x_1, ..., x_p\}$ se dice* **ortogonal** *si*

$$F(x_i, x_k) = 0, \forall i \neq k$$

Como ejemplos de familias de vectores ortogonales podemos considerar los llamados polinomios ortogonales: Legendre, Chebyshev, Laguerre, Hermite.

Ejemplo 8.2.4 *(Polinomios de Chebyshev) Se definen como polinomios $T_n(x)$ ortogonales en el intervalo [1,1] con respecto de la función peso $w(x) = 1/\sqrt{1-x^2}$, es decir, considerando el producto escalar definido en $\mathbb{R}_n[x]$ mediante*

$$F(T_n(x), T_m(x)) = \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} T_n(x)\, T_m(x)\, dx$$

Existe una relación de recurrencia (fórmula de Rodrigues) (también para toda familia de polinomios ortogonales)

$$T_n(x) = 2\,x\,T_{n-1}(x) - T_{n-2}(x), \quad n = 2, 3, ...$$

que permite su obtención sin necesidad de calcular integrales; de cualquiera de las dos formas se puede obtener la familia de los primeros polinomios:

$$T_o(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2\,x^2 - 1, \quad T_3(x) = 4\,x^3 - 3\,x, ...$$

Ejercicio 8.2.5 *Comprobar que la familia anterior es ortogonal.*

Propiedad 8.2.6 *Toda familia ortogonal de vectores no nulos es libre.*

Dem.: Sea la familia $\{a_i\}$ ortogonal con p vectores. Supongamos que $\sum_{i=1}^{p} t_j\, a_j = 0$. Multiplicando escalarmente por a_k resulta:

$$\sum_{i=1}^{p} t_j\, F(a_j, a_k) = t_k\, F(a_k, a_k) = t_k\, ||a_k||^2 = 0 \Rightarrow t_k = 0, \forall k. \;\blacksquare$$

Definición 8.2.7 *Una* **base** $\{e_i\}$ *de un espacio vectorial euclideo se dice* **ortogonal** *si es una familia ortogonal. En particular, una* **base** *se dirá* **ortonormada** *si es ortogonal y sus vectores son unitarios, es decir, si*

$$F(e_i, e_k) = \delta_{i\,k} \quad \textit{(deltas de Krönecker)}$$

Por ejemplo, las familias compuestas por los n+1 primeros polinomios ortogonales son bases ortogonales y, convenientemente normalizadas, también son bases ortonormadas.

Definición 8.2.8 *Las coordenadas de cualquier vector respecto de una base ortonormada se denominan* **coordenadas euclídeas o cartesianas** *y el* **sistema coordenado cartesiano**.

Observación. Recordando que los elementos de la matriz de Gram coordenada del producto escalar que define el espacio vectorial euclídeo en una base $\{e_i\}$ están definidos por $a_{ik} = F(e_i, e_k)$, si la base fuese ortonormada resultaría $a_{ik} = F(e_i, e_k) = \delta_{i\,k}$, es decir, la matriz de Gram A sería la matriz unidad de orden n. En este caso, la forma fundamental (el producto escalar) se expresaría del siguiente modo:

$$F(x, y) = X^T\, I\, Y = X^T\, Y = x_1\, y_1 + ... + x_n\, y_n$$

que corresponde a la clásica definición de producto escalar de vectores libres. Por esta razón, se escribe $F(x, y) = x \cdot y$.

Es evidente que cualquier otro producto escalar (o, en general, cualquier forma bilineal) que esté definido en el mismo espacio vectorial euclideo V y considerando la misma base, no tendrá como matriz asociada la matriz unidad.

Propiedad 8.2.9 *La coordenada i-ésima de un vector x respecto de una base ortonormada* $\{e_i\}$ *está dada por* $x_i = x \cdot e_i$.

Corolario 8.2.10 *Si x es un vector unitario cualquiera,* $x_i = \cos(x, e_i)$.

Definición 8.2.11 *Llamaremos* **cosenos directores** *de un vector unitario a las coordenadas de dicho vector unitario en una base ortonormada.*

Hemos presentado ya varios ejemplos de familias ortogonales que son bases, veremos en las siguientes lineas cómo encontrar una base ortonormada en forma general.

Teorema 8.2.12 *Dado un vector* $a \in V, a \neq 0$, *para cada* $x \in V$, *existen un solo* $t \in \mathbb{R}$ *y un solo vector* $v \in V$ *tales que*

$$x = ta + v \quad y \quad a \cdot v = 0$$

Dem.: Supongamos que exista $v \in V$ tal que $a \cdot v = 0$, entonces:

$$a \cdot x = a \cdot (ta + v) = t(a \cdot a) + a \cdot v = t\, a \cdot a$$

por lo que, debe ser: $t = \dfrac{a \cdot x}{a \cdot a}$

En consecuencia, el vector v está determinado unívocamente por

$$v = x - \frac{a \cdot x}{a \cdot a} a. \quad \blacksquare$$

Teorema 8.2.13 *(Método de ortogonalización de Gram-Schmidt) Una familia ortogonal* $\{a_1, ..., a_p\}$*,* $p \leq n = dimV$*, puede completarse hasta formar una base ortogonal de V.*

Dem.: Por el teorema de la base incompleta, si p < n, existe un vector b_{p+1} de modo que $\{a_1, ..., a_p, b_{p+1}\}$ es libre. Entonces, elegimos el vector a_{p+1} de la forma siguiente:

$$a_{p+1} = b_{p+1} - \frac{b_{p+1} \cdot a_1}{||a_1||} \frac{a_1}{||a_1||} - ... - \frac{b_{p+1} \cdot a_p}{||a_p||} \frac{a_p}{||a_p||}$$

es decir, a_{p+1} es la diferencia entre el vector b_{p+1} y la suma de las proyecciones ortogonales de b_{p+1} sobre los vectores $a_1, ..., a_p$. De esta forma, la familia $\{a_1, ..., a_{p+1}\}$ es ortogonal, como se comprueba fácilmente.

Así, pues, hemos encontrado una familia ortogonal compuesta de p+1 vectores. Si p+1 = n, dicha familia es base; caso contrario, p+1 < n, y repetimos el proceso r veces hasta que p + r = n, hasta encontrar una base ortogonal. ■

Corolario 8.2.14 *Una base ortonormada* $\{v_1, \cdots, v_r\}$ *de un subespacio S de V puede completarse con vectores* $\{v_{r+1}, \cdots, v_n\}$ *hasta formar una base ortonormada de V.*

Definición 8.2.15 *Los vectores* $\{v_{r+1}, \cdots, v_n\}$ *componen una base ortonormada de un subespacio* $S^{\perp}$ *de V denominado* **complemento ortogonal** *de S.*

Ejercicio 8.2.16 $V = \mathbb{R}^2$ *y* $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ *la matriz asociada a cierta aplicación bilineal f. Se pide: i) ¿Es f un producto escalar?; ii) encontrar el producto escalar de los vectores x = (5,2), y = (1,2); iii) encontrar una base en la que la matriz asociada a f sea la unidad.*

Solución:
i) Basta comprobar los axiomas de la definición de producto escalar.
ii) $f(x, y) = x \cdot y = X^T \, A \, X = 49$
iii) Hay que buscar una base ortonormal, para lo cual usaremos el método de ortogonalización de Schmidt a partir de la base canónica. ■

8.3 Matrices ortogonales.

En el espacio euclídeo V, (V, F), consideraremos bases ortonormadas, B_v y $B_{\bar{v}}$. Sea P la matriz del cambio de la base B_v a la $B_{\bar{v}}$, es decir tal que $X = P \, \bar{X}$. El producto escalar se expresará en ambas bases en la forma:

$$F(x, y) = x \cdot y = X^T \, Y = \bar{X}^T \, P^T \, P \, \bar{X}$$

Por lo que, al ser las bases ortonormadas, se debe verificar:

$$P^T \, P = I, \quad \text{es decir,} \quad P^T = P^{-1} \tag{8.1}$$

Definición 8.3.1 *Una matriz real que cumple (8.1) se denomina* **matriz ortogonal**

Propiedad 8.3.2 .

1) P ortogonal $\Longleftrightarrow$ sus columnas (y sus filas) forman una base ortonormada

2) Si P es ortogonal $\Longrightarrow \det P = \pm 1$

Propiedad 8.3.3 *El conjunto $\mathcal{O}(n)$ de las matrices ortogonales de orden n es grupo con respecto al producto de matrices, denominado* **grupo ortogonal**.

Observación: Algunos autores denominan Geometría euclídea al estudio de los invariantes del grupo ortogonal.

Definición 8.3.4 *Una aplicación lineal definida por una matriz ortogonal se denomina* **transformación ortogonal (o isometría)**

Teorema 8.3.5 *Las transformaciones ortogonales conservan la norma de vectores.*

Dem.: En efecto: $h(x) \cdot h(y) = X^T P^T P Y = X^T Y = x \cdot y$. ■

Las isometrías del plano euclídeo son las rotaciones y las simetrías.

8.4 Factorización Q R.

Notemos que al aplicar el método de ortogonalización de Gram-Schmidt a una familia de vectores $\{a_j\}_1^p$, $p \leq n$, se obtiene otra familia $\{v_j\}$ que es ortogonal. Una vez normalizada se obtiene una tercera familia ortonormada $\{q_j\}$ tales que $q_j = \dfrac{v_j}{||v_j||}$ que se puede completar hasta formar una base ortonormada.

La ecuaciones que definen los vectores v_j se pueden escribir también en la forma:

$$\begin{array}{lll} v_1 = a_1 = ||v_1||\, q_1 & a_1 = q_1\, R_{11}, & \text{con } R_{11} = ||v_1|| \\ v_2 = a_2 - \dfrac{a_2 \cdot v_1}{v_1 \cdot v_1} v_1, & a_2 = q_1\, R_{12} + q_2\, R_{22}, & \text{con } R_{22} = ||v_2|| \\ v_3 = a_3 - \dfrac{a_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{a_3 \cdot v_2}{v_2 \cdot v_2} v_2, & a_3 = q_1\, R_{13} + q_2\, R_{23} + q_3\, R_{33}, & \text{con } R_{33} = ||v_3|| \\ \dots & & \end{array}$$

es decir,

$$A = [a_1 \dots a_p] = [q_1 \dots q_p] \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1p} \\ 0 & R_{22} & \dots & R_{2p} \\ \dots & & & \\ 0 & 0 & \dots & R_{pp} \end{bmatrix} = Q\, R$$

Definición 8.4.1 *La descomposición anterior de una matriz A como producto de una Q de columnas ortonormadas y una R triangular superior se le denomina* **factorización Q R** *de la matriz A. En particular, si A es matriz cuadrada, Q es una matriz ortogonal.*

Ejercicio 8.4.2 *Obtener la factorización $Q\,R$ de la matriz* $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix}$

Solución: Los vectores a_j son las columnas de A. Las ecuaciones de Gram-Schmidt nos conducen a:

$$Q = \begin{bmatrix} 1/2 & \dfrac{-3}{2\sqrt{5}} & 1/2 \\ 1/2 & \dfrac{-1}{2\sqrt{5}} & -1/2 \\ 1/2 & \dfrac{1}{2\sqrt{5}} & -1/2 \\ 1/2 & \dfrac{3}{2\sqrt{5}} & 1/2 \end{bmatrix}, \quad R = \begin{bmatrix} 2 & 1 & 3 \\ 0 & \sqrt{5} & \sqrt{5} \\ 0 & 0 & 2 \end{bmatrix}$$

8.4.1 Método de Householder

Dado que el método de Gram-Schmidt requiere que los sucesivos vectores construidos sean ortogonales, en el caso de que los cálculos se hagan de manera aproximada (casi siempre) eso no ocurre exactamente, por lo que se producen graves errores. Una forma de evitarlos es utilizar el método de Householder. Este método consiste en realizar operaciones de eliminación (como en el método de Gauss), pero usando ahora matrices de Householder, que anulan todos los elementos por debajo de la diagonal en cada uno de los n-1 sucesivos pasos.

Definición 8.4.3 *Llamaremos* **matriz de Householder** $H(u)$ *correspondiente a un vector unitario u a la matriz*

$$H(u) = I - 2\,u\,u^T, \qquad ||u|| = 1$$

Propiedad 8.4.4 *Toda matriz de Householder verifica:*
1) es simétrica y ortogonal
2) $H(u)\,v = -v$, *si* $v = \lambda\,u$, $\qquad H(u)\,v = v$, *si* $v \cdot u = 0$

Recordemos (8.2.12).

El Algoritmo

$\boxed{\text{Primera etapa}}$ Se busca una matriz $H_1 = I - 2u_1\,u_1^T \quad (u_1 \in \mathbb{R}^m$, y $u_1^T\,u_1 = 1)$ tal que

$$H_1\,A = \begin{pmatrix} \alpha_1 & * & \ldots & * \\ 0 & * & \ldots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \ldots & * \end{pmatrix} \tag{8.2}$$

Si A^1 es la primera columna de A, se debe cumplir que $H_1\,A^1 = \alpha_1\,e_1 = (\alpha, 0, ..., 0)^T$, luego: $|\alpha_1| = ||H_1\,A^1||_2 = ||A^1||_2$. Por definición de H_1, debe ser

$$H_1\,A^1 = A^1 - 2u_1\,u_1^T\,A^1 = \alpha_1\,e_1$$

Como la expresión $u_1^T\,A^1$ es un escalar, resulta que

$$u_1 = C\,v_1, \quad \text{donde } v_1 = A^1 - \alpha_1\,e_1 \tag{8.3}$$

y la constante C se determina de forma que $||u_1||_2 = 1$. Como todavía queda la libertad de elegir el signo de α_1, este se escoge en la forma

$$\alpha_1 = -(\text{sign}\ a_{11})\,||A^1||_2 \tag{8.4}$$

para evitar una sustracción mal condicionada en el cálculo de $v_1 = A^1 - \alpha_1 e_1$

Para *calcular* la matriz $H_1 A$ téngase en cuenta que

$$H_1 A = A - 2u_1 u_1^T A = A - \beta v_1 (v_1^T A), \quad \text{donde } \beta = \frac{2}{v_1^T v_1} \tag{8.5}$$

Este factor β puede ser calculado en la forma siguiente

$$\beta^{-1} = \frac{v_1^T v_1}{2} = \frac{1}{2}((A^1)^T A^1 - 2\alpha a_{11} + \alpha_1^2) = -\alpha_1 (a_{11} - \alpha_1). \tag{8.6}$$

$\boxed{\text{Segunda etapa}}$ Se aplica el mismo procedimiento a la submatriz de (8.2) de tipo $(m-1) \times (n-1)$; lo cual proporciona un vector $\bar{u}_2 \in \mathbb{R}^{m-1}$ y una matriz de Householder $\bar{H}_2 = I - 2\,\bar{u}_2\,\bar{u}_2^T$. Escribiendo $u_2 = (0, \bar{u}_2)^T$ y multiplicando (8.2) por $H_2 = I - 2\,u_2\,u_2^T$, se obtiene

$$H_2 H_1 A = H_2 \begin{pmatrix} \alpha_1 & * & \dots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \alpha_1 & * & \dots & * \\ 0 & & & \\ \vdots & & \bar{H}_2 C & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \alpha_1 & * & * & \dots & * \\ 0 & \alpha_2 & * & \dots & * \\ 0 & 0 & * & \vdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \dots & * \end{pmatrix}$$

Este proceso se debe continuar n etapas ($n-1$, si $m = n$) para obtener la matriz triangular

$$H_n \dots H_2 H_1 A = R = \begin{pmatrix} R' \\ 0 \end{pmatrix}$$

y la matriz ortogonal $Q^T = H_n \dots H_2 H_1$.

Nótese que a la hora de construir un programa para este algoritmo, no es necesario calcular explicitamente las matrices H_i, ni la matriz Q, es suficiente retener, además de R, los valores de β_i (o α_i) y los vectores v_i, los cuales contienen toda la información necesaria.

Puede verse que el coste computacional de la factorización $Q\,R$ es aproximadamente el doble que la factorización $L\,U$.

Ejercicio 8.4.5 *Obtener la factorización $Q\ R$ de la matriz* $A = \begin{bmatrix} 1 & -4 & 5 \\ 1 & -5 & 1 \\ -1 & 6 & 1 \\ 1 & -5 & 0 \end{bmatrix}$

Solución: Por Hoseholder.

$$1)\ \alpha = -sign(a_{11})\,||A^1|| = -2, \quad \beta = 1/6, v = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, R = \begin{bmatrix} -2 & 10 & -4 \\ 0 & -1/3 & -2 \\ 0 & 4/3 & 4 \\ 0 & -1/3 & 0 \end{bmatrix}$$

$$2)\ \alpha = -sign(R_{22})\,||R^{2*}|| = \sqrt{2},\ \beta = \frac{3}{6+\sqrt{2}},\ v = \begin{bmatrix} 0 \\ -1/3 - \sqrt{2} \\ 4/3 \\ -1/3 \end{bmatrix}, R = \begin{bmatrix} -2 & 10 & -4 \\ 0 & \sqrt{2} & 3\sqrt{2} \\ 0 & 0 & -\sqrt{2}/2 \\ 0 & 0 & 1.190 \end{bmatrix}$$

$$3)\, \alpha = -sign(R_{33})\,||R^{3*}|| = \sqrt{2}, \quad \beta = 0.324, v = [\,]\,, R = \begin{bmatrix} -2 & 10 & -4 \\ 0 & \sqrt{2} & 3\sqrt{2} \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}$$

$$Q = H_1\, H_2\, H_3 = \begin{bmatrix} -1/2 & \sqrt{2}/2 & 0 & -1/2 \\ -1/2 & 0 & -/2 & 1/2 \\ 0.5 & \sqrt{2}/2 & 0 & 1/2 \\ -1/2 & 0 & \sqrt{2}/2 & 1/2 \end{bmatrix}. \blacksquare$$

8.4.2 Resolución de sistemas lineales

La factorización A = Q R se puede utilizar, igual que la factorización L U, para la resolución de un sistema lineal A X = B, con la ventaja de que ahora la inversa de Q es inmediata y todo se reduce a la resolución del sistema lineal

$$A\,X = B \Longleftrightarrow R\,X = Q^T\,B$$

Ejercicio 8.4.6 *Resolver el sistema A X = B mediante factorización Q R, siendo*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \qquad b = \begin{bmatrix} 3 \\ 3 \\ -1 \end{bmatrix}$$

Solución:

$$[R, Q^T\,b] = \begin{bmatrix} -\sqrt{2} & -3/\sqrt{2} & -\sqrt{2} & -3\sqrt{2} \\ 0 & -\dfrac{3\sqrt{2}+3}{\sqrt{2}+2} & -\dfrac{11\sqrt{2}+22}{3\sqrt{2}+3} & \dfrac{2\sqrt{2}}{3} \\ 0 & 0 & 1/3 & -1/3 \end{bmatrix} \Longrightarrow \begin{bmatrix} x \\ y \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}. \blacksquare$$

8.4.3 Cálculo de valores y vectores propios

Método de factorización Q R

En primer lugar, consideraremos el caso en que la matriz A es simétrica.

En esta situación, el método de factorización QR consiste en transformar sucesivamente la matriz A en otra congruente ortogonal que sea cada vez más próxima a una diagonal aprovechando la factorización $Q\,R$ de las sucesivas matrices. Para ello, en un primer paso, se realiza la mencionada factorización

$$A = Q_1\,R_1$$

En segundo paso, se construye la matriz

$$A_1 = R_1\,Q_1 \quad (= Q_1^T\,A\,Q_1),$$

que se vuelve a factorizar. Y así sucesivamente.

Para obtener dicha factorización se puede utilizar el método de Householder expuesto más arriba, o bien, transforma la matriz A en otra congruente ortogonal que tenga la forma de Hessenberg, mediante multiplicación a izquierda y derecha por matrices de Householder (ver (8.4.3) y, a continuación, realizar $n-1$ rotaciones de Givens para transformarla en una triangular superior.

El algoritmo se detiene cuando el elemento $(A_k)_{nn-1}$ es suficientemente pequeño; entonces el elemento $(A_k)_{nn}$ es un valor propio de A.

Construcción de la matriz de Householder para llevar a forma de Hessenberg

Hay que buscar una matriz de Householder $H = I - 2\,\omega\,\omega^T$, $||\omega|| = 1$ tal que la primera columna de la matriz $A_1 = H\,A\,H$ sea de la forma

$$A_1^1 = (a_{11}, \alpha, 0, \dots, 0)^T,$$

Para ello, es suficiente elegir (cf. [5])

$$\alpha = -sign(a_{21})\,(\textstyle\sum_{j=2}^n a_{j1}^2)^{1/2}$$

$$r = (\tfrac{1}{2}\,\alpha\,(\alpha - a_{21})\,)^{1/2}$$

$$\omega_1 = 0, \quad \omega_2 = \frac{a_{21}-\alpha}{2\,r}, \quad \omega_k = \frac{a_{k1}}{2\,r}, \quad k = 3, 4, ..., n$$

En sucesivos pasos, se construirán matrices de Householder que vayan transformando en ceros todos los elementos por debajo de la primera paralela a la diagonal de la segunda, tercera, etc. columnas, de forma que la matriz final tenga la forma de Hessenberg.

Ejercicio 8.4.7 *Mediante factorización $Q\,R$, hallar valores y vectores propios de*

$$A = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

Solución:

Primera iteración: $Q = \begin{bmatrix} -0.87 & 0.142 & 0.471 \\ 0.246 & -0704 & 0.667 \\ 0.426 & 0.696 & 0.577 \end{bmatrix}$, $\quad R = \begin{bmatrix} 1.91 & -0.246 & 0.426 \\ 0 & 1.81 & -0.696 \\ 0 & 0 & 1.15 \end{bmatrix}$

Segunda iteración: $Q = \begin{bmatrix} -0.966 & 0.425 & 0.87 \\ 0.075 & -0.904 & 0.42 \\ 0.249 & 0.696 & 0.577 \end{bmatrix}$, $\quad R = \begin{bmatrix} 1.98 & -0.075 & -0.249 \\ 0 & 1.94 & -0.425 \\ 0 & 0 & 1.04 \end{bmatrix}$

Tercera iteración: $Q = \begin{bmatrix} -0.991 & 0.010 & 0.132 \\ 0.0199 & -0.974 & 0.224 \\ 0.13 & 0.225 & 0.966 \end{bmatrix}$, $\quad R = \begin{bmatrix} 1.99 & -0.0199 & -0.13 \\ 0 & 1.98 & -0.225 \\ 0 & 0 & 1.01 \end{bmatrix}$

Los valores propios exactos son: 2, 2, 1. ■

8.5 Proyección ortogonal. Mejor aproximación.

Definición 8.5.1 *Se denomina* **proyección ortogonal** *del vector v sobre un subespacio S del espacio euclídeo V al vector*

$$x = \sum_{i=1}^r \vec{proy}_{v_i} v, \qquad \textit{siendo} \qquad \vec{proy}_{v_i} v = \frac{v \cdot v_i}{||v_i||^2}\, v_i$$

y $\{v_i\}$ una base ortogonal de S.

Definición 8.5.2 *Se denomina* **mejor aproximación** *en el subespacio S del vector v del espacio euclídeo V al vector x de S que hace mínima $||v-x||$*

Teorema 8.5.3 *Sea $S \leq V$, espacio vectorial euclídeo. Sean $v \in V$ y $x \in S$, entonces, las siguientes afirmaciones son equivalentes:*
i) $v - x \perp S$
ii) x es una mejor aproximación a v sobre S.

Dem.: i) $\Rightarrow$ ii). Sea un vector cualquiera $w \in S$, se tiene:

$$||v-w||^2 = ||(v-x)+(x-w)||^2 = ||v-x||^2 + ||x-w||^2 \geq ||v-x||^2.$$

i) $\Leftarrow$ ii). Sea $h \in S$ y $\lambda > 0$. Se tiene:

$$\begin{aligned} 0 &\leq ||v-x+\lambda\, h||^2 - ||v-x||^2 \\ &= ||v-x||^2 + 2\lambda < v-x, h > +\lambda^2\, ||h||^2 - ||v-x||^2 \\ &= \lambda\,(2 < v-x, h > +\lambda\, ||h||^2). \end{aligned}$$

Haciendo $\lambda \to 0$, se obtiene que $< v-x, h > \geq 0$. Como la misma desigualdad debe ser satisfecha por $-h$, también se tiene: $< v-x, h > \leq 0$; luego, $< v-x, h >= 0$; pero como h es arbitraria en S, se cumple que $v - x \perp S$. ■

Definición 8.5.4 *El sistema de ecuaciones lineales que resulta al expresar la condición i) respecto de alguna base de S se denomina* **ecuaciones normales**

Si $\{u_j\}$ es la base de S, las ecuaciones normales se pueden escribir en la forma:

$$x \cdot u_j = v \cdot u_j, \quad j = 1, 2, ..., n$$

Ejemplo 8.5.5 *Aproximar $f(x) = \operatorname{sen} x$ en [-1,1] por $c_1 x + c_2 x^3 + c_3 x^5$ considerando el producto escalar*

$$< f, h >= \int_{-1}^{1} f(x)\, h(x)\, dx$$

Soluc.: Podemos tomar como S el subespacio engendrado por los polinomios x, x^3, x^5 que componen una base de S. Por lo tanto, las ecuaciones normales serán en este caso:

$$c_1 < x, x^{2j-1} > +c_2 < x^3, x^{2j-1} > +c_3 < x^5, x^{2j-1} >=< f, x^{2j-1} >, \quad j = 1, 2, 3$$

que componen un sistema lineal que se puede escribir en la forma matricial siguiente:

$$A\, c = b,$$

siendo:

$$1/2\, A = \begin{pmatrix} 1/3 & 1/5 & 1/7 \\ 1/5 & 1/7 & 1/9 \\ 1/7 & 1/9 & 1/11 \end{pmatrix}, \qquad b = \begin{pmatrix} \alpha - \beta \\ -3\alpha + 5\beta \\ 65\,\alpha - 101\,\beta \end{pmatrix}, \quad \alpha = \operatorname{sen} 1, \quad \beta = \cos 1,$$

después de haber calculado las integrales debidas.

Resolviendo el sistema se obtiene: $c_1 \approx 348.9, c_2 \approx -1491.1, c_3 \approx 1275.6$.

Se debe tener en cuenta que el problema puede estar mal condicionado. Ese mal condicionamiento suele ser frecuente cuando se utilizan polinomios respecto de la base natural. En general, es conveniente considerar otras bases, por ejemplo, ortonormales.

Si la base $\{u_j\}$ considerada en S es una base ortonormal se cumple: $< u_i, u_j >= \delta_{i\,j}$ y, si expresamos el elemento g en dicha base mediante: $g = a_1\, u_1 + ... + a_n\, u_n$, las ecuaciones normales serán en este caso:

$$a_1 \; < u_1, u_j > + ... + a_n \; < u_n, u_j > = < f, u_j >, \quad j = 1, 2, ..., n,$$

es decir,

$$a_j = < f, u_j >, \quad j = 1, 2, ..., n,$$

lo que quiere decir que g es la suma de las proyecciones ortogonales de f sobre los vectores de la base, o sea, la proyección ortogonal sobre S, como implícitamente se encuentra contenido en el teorema 8.5.3.

En el ejemplo precedente, se puede considerar la base de polinomios ortogonales de Legendre (ortonormada respecto del producto escalar mencionado) siguiente:

$$\begin{aligned}
P_1(x) &= x/\sqrt{2/3}\\
P_2(x) &= (5\,x^3 - 3\,x)\sqrt{2/3}\\
P_3(x) &= (63\,x^5 - 70\,x^3 + 15\,x)/\sqrt{2/11}
\end{aligned}$$

y buscar el polinomio $g(x) = a_1\, P_1(x) + a_2\, P_2(x) + a_3\, P_3(x)$. En este caso obtendríamos, con mucho menor esfuerzo computacional, la siguiente solución: $a_1 \approx 0.73731, a_2 \approx -6.7391 * 10^{-2}, a_3 \approx 2.2904 * 10^{-2}$.

Ejercicio 8.5.6 *(***Aproximación por polinomios de Chebyshev.***) Aproximar la función* $f(x) = \operatorname{sen} x$ *en* $\mathbb{R}_3[x]$

Ejercicio 8.5.7 *(***Aproximación por series de Fourier (limitadas).***) Obtener la mejor aproximación de la función* $f(x) = -1, si\, x < 0; = 1, si\, x > 0$ *en el subespacio* $S = < 1, \cos x, \operatorname{sen} x, \cos 2x, \operatorname{sen} 2x >$

Ayuda:

Considerar el producto escalar

$$< f, h > = \int_{-\pi}^{\pi} f(x)\, h(x)\, dx$$

respecto del cual la familia generadora dada es una base ortogonal del subespacio.

Solo falta construir la proyección ortogonal de $f(x)$ sobre S. ■

8.5.1 El problema del ajuste

Un planteamiento muy común en las ciencias experimentales conduce también a problemas de Aproximación.

Sea cierto fenómeno descrito por la variable dependiente y, la variable independiente t y ciertos parámetros $a_1, ..., a_n$, todo ellos reales y relacionados mediante la ecuación:

$$y = \phi(t; a_1, ..., a_n), \tag{8.7}$$

que se supone conocida. Por ejemplo, experimentalmente se sabe que la cantidad de sustancia radiactiva en una muestra en un instante dado está definida por la ecuación

$$y = a_1 e^{-a_2 t},$$

donde los parámetros a_1, a_2 dependen de la muestra concreta.

Interesa, pues, determinar los parámetros, para lo cual se dispone de tantas medidas experimentales como se quiera, en general, muchas más que parámetros a determinar. Los parámetros deben verificar las relaciones

$$y_j = \phi(t_j; a_1, ..., a_n), \quad j = 1, 2, ..., m \gg n \tag{8.8}$$

Debido a que el modelo matemático considerado no es absolutamente preciso y a que las medidas realizadas no son exactas, las igualdades (8.8) no se cumplen. Se plantea entonces determinar el valor de los parámetros $a_1, ..., a_n$ (es decir, ajustar) de forma que el vector de primeros miembros y el de segundos miembros de (8.8) estén a la mínima distancia en el sentido de la norma euclídea de $\mathbb{R}^m$. Notar que se podrían utilizar otras normas.

Utilizando este procedimiento, Kepler descubrió, en 1601, su tercera ley a partir de observaciones de los periodos y las distancias al Sol de los cuatro primeros planetas.

8.5.2 El problema lineal de los mínimos cuadrados

Si consideramos que la función ϕ es lineal en los parámetros, las ecuaciones (8.8) toman la forma

$$A\,x = b, \tag{8.9}$$

donde $x \in \mathbb{R}^n$ que juega el papel de los parámetros a_i, $b \in \mathbb{R}^m$ y $A \in M_{\mathbb{R}}(m, n)$. Supondremos que $m > n$ y que rg $A = n$. Se ha planteado así un problema lineal que, en general, carece de solución.

Hemos llegado al **problema lineal de los mínimos cuadrados**, que se formula de la siguiente manera:

Dado un vector $b \in \mathbb{R}^m$ y una matriz $A \in M_{\mathbb{R}}(m, n)$, $m > n$ y rg $A = n$, determinar $x^* \in \mathbb{R}^n$ tal que haga mínima la norma euclídea del vector residuo definido por

$$r(x) = b - A\,x$$

Observemos que esto es equivalente a encontrar la mejor aproximación a b sobre el espacio vectorial

$$M = \{A\,x \,|\, x \in \mathbb{R}^n\}$$

engendrado por las columnas v_i de A, que es de dimensión n.

Aplicando el teorema 8.5.3 que caracteriza la mejor aproximación, resulta:

$$< b - A\,x^*, v_i >= 0, \quad i = 1, 2, ..., n$$

(ecuaciones normales o de Gauss) o, equivalentemente:

$$A^T A\,x^* = A^T b$$

Ejercicio 8.5.8 (**Mínimos cuadrados**) *Encontrar el polinomio cuadrático que mejor se aproxima en el sentido de los mínimos cuadrados a la tabla de valores siguiente que presenta la relación medida en el laboratorio entre la temperatura y la solubilidad del* NO_3K

t	40	60	80	100	120
s	27	39	50	60	69

Solución:

Este sistema se resuelve por cualquier método de los conocidos, por ejemplo, el de Cholesky, ya que la matriz $A^T A$ es simétrica definida positiva. Sin embargo, dado que está mal condicionado, es preferible utilizar el método de factorización $Q\,R$. Veámoslo.

El polinomio a encontrar: $P(x) = a + b\,t + c\,t^2$

El sistema a resolver: $a + b\,t_j + c\,t_j^2 = s_j, \quad j = 1, 2, 3, 4, 5$

Por lo tanto:

$$A = \begin{bmatrix} 1 & 40 & 1600 \\ 1 & 60 & 3600 \\ 1 & 80 & 6400 \\ 1 & 100 & 10000 \\ 1 & 120 & 14400 \end{bmatrix}, \quad B = \begin{bmatrix} 27 \\ 39 \\ 50 \\ 60 \\ 69 \end{bmatrix}$$

Ecuaciones de Gauss:

$$A^T A = \begin{bmatrix} 5 & 400 & 36000 \\ 400 & 36000 & 3520000 \\ 36000 & 3520000 & 363840000 \end{bmatrix}, \quad A^T b = \begin{bmatrix} 245 \\ 21700 \\ 2097200 \end{bmatrix}$$

$$a = 0, \quad b = 0.725, \quad c = -0.00125$$

8.6 Ejercicios para resolver

Ejercicios de [1] (Merino) números 25, 27, 30, 31 (ejerc. resueltos), 80, 81, 83, 84, 165 y 170 (ejerc. propuestos)

Ejercicios de [3] (de la Villa) capítulo 8, números 3, 4, 6

Ejercicios de [3] (de la Villa) capítulo 9, números 7, 8, 9

Ejercicios de [3] (de la Villa) capítulo 9, números 5, 6

Ejercicios de las hojas de enunciados de [2] (Palacios).