Guía de estudio:
Espacios euclídeos y Ortogonalización

Prof. Maria-Angeles Zurro

29 de febrero de 2024

Índice

1. Espacios euclídeos

El contenido de esta sección sigue la línea de presentación dada en [1].

Un *espacio euclídeo* es un espacio vectorial real E dotado de un producto escalar, es decir, de una regla que asigne a cada par de vectores $\vec{u}, \vec{v}$ un número real designado por $< \vec{u}, \vec{v} >$, de forma que se cumplen las siguientes propiedades:

1. Propiedad simétrica: $< \vec{u}, \vec{v} >=< \vec{v}, \vec{u} >$.

2. Propiedad distributiva: $< \vec{u} + \vec{w}, \vec{v} >=< \vec{u}, \vec{v} > + < \vec{w}, \vec{v} >$.

3. $< \lambda\vec{u}, \vec{v} >= \lambda < \vec{u}, \vec{v} >$ para cada $\lambda \in \mathbb{R}$.

4. $< \vec{u}, \vec{v} >\geq 0$, y sólo es 0 si $\vec{u} = \vec{v}$.

Se define el módulo o longitud (o norma) de un vector $\vec{u} \in E$ como

$$||\vec{u}|| = \sqrt{\langle \vec{u}, \vec{u} \rangle}. \tag{1}$$

Ejemplo 1.1. *1. Si $E = \mathbb{R}^2$, y $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, entonces la fórmula*

$$< \vec{u}, \vec{v} >:= x_1x_2 + y_1y_2 \tag{2}$$

define un producto escalar en E.

2. *Si $E=\mathbb{R}^2$, y $\vec{u}=(x_1,y_1)$, $\vec{v}=(x_2,y_2)$, entonces la fórmula*

$$< \vec{u},\vec{v} >:= (x_1,y_1)\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}x_2\\ y_2\end{pmatrix} \tag{3}$$

   define un producto escalar en E. En este caso, el módulo de un vector tiene la fórmula

$$||\vec{u}|| = \sqrt{x_1^2+2x_1y_1+2y_1^2} = \sqrt{(x_1+y_1)^2+y_1^2}.$$

3. *Si E es el espacio de funciones continuas reales en el intervalo $[a,b]$, denotado por $C([a,b])$, se define para cada par de funciones f,g en E el producto escalar*

$$< f,g >:= \int_a^b f(t)g(t)\,dt\ , \tag{4}$$

   obteniéndose así un espacio euclídeo de funciones.

En un espacio euclídeo se cumple la siguiente desigualdad.

Proposición 1 (Desigualdad de Schwarz)**.** *En un espacio euclídeo E, se da*

$$|< \vec{u},\vec{v} >| \le ||\vec{u}||\cdot||\vec{v}||\,,$$

para cualesquiera $\vec{u},\vec{v}$ en E.

Demostración. Consideremos dos vectores cualesquiera $\vec{u},\vec{v}$ en E, y λ un escalar real arbitrario. Entonces,

$$< \lambda\vec{u}-\vec{v},\lambda\vec{u}-\vec{v} >\ge 0, \tag{5}$$

por las propiedades del producto escalar. Desarrollando la desigualdad anterior se obtiene

$$\lambda^2||\vec{u}||^2 - 2 < \vec{u},\vec{v} > + ||\vec{v}||^2 \ge 0.$$

La parte izquierda de esta ecuación es un polinomio cuadrático en λ que no puede tener raíces reales y distintas. En consecuencia su discriminante ha de ser siempre ≤ 0. Por tanto $< \vec{u},\vec{v} >^2 -||\vec{u}||^2||\vec{v}||^2 \le 0$. Tomando raíces cuadradas obtenemos $| < \vec{u},\vec{v} > | \le ||\vec{u}||||\vec{v}||$, que es la desigualdad pedida. □

Consecuencia de la desigualda de Cauchy-Schwarz es que en los espacios vectoriales euclídeos se puede establecer un concepto de ángulos entre sus elementos. Por tanto podremos hablar de elementos ortogonales, concepto análogo al de vectores perpendiculares en el espacio euclídeo $\mathbb{R}^2$ con la métrica usual.

Ortogonalidad. Sea E un espacio euclídeo con producto escalar $\langle,\rangle$. Dados dos vectores $\vec{u},\vec{v}$ en E, diremos que son ortogonales (respecto de $\langle,\rangle$) si $\langle\vec{u},\vec{v}\rangle=0$. Además si W es un subespacio de E, se define su complemento ortogonal $W^\perp$ como el conjunto

$$W^\perp = \{\vec{w}\in E \mid \langle\vec{w},\vec{u}\rangle = 0 \text{ para todo } \vec{u}\in E\ \}\,. \tag{6}$$

En estos espacios existe una noción de *ángulo* entre vectores. En concreto, si $\vec{u},\vec{v}$ están E, definimos el *coseno del ángulo* que forman $\vec{u}$ y $\vec{v}$ como

$$\cos\angle(\vec{u},\vec{v}) := \frac{\langle\vec{u},\vec{v}\rangle}{||\vec{u}||\cdot||\vec{v}||}. \tag{7}$$

En consecuencia, el *ángulo* que forman $\vec{u}$ y $\vec{v}$ es

$$\angle(\vec{u},\vec{v}) := \arccos\left(\frac{\langle\vec{u},\vec{v}\rangle}{||\vec{u}||\cdot||\vec{v}||}\right)\in[0,\pi]\,. \tag{8}$$

Ejemplo 1.2. *Si $E = \mathbb{R}^2$, y consideramos el producto escalar definido por la fórmula*

$$< \vec{u}, \vec{v} >:= (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \tag{9}$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E, entonces

$$\angle((1,0),(0,1)) = \arc\cos\left(\frac{1}{1 \cdot \sqrt{2}}\right) = \frac{\pi}{4}. \tag{10}$$

En los espacios euclítos se verifica el Teorema de Pitágoras generalizado.

Teorema 2 (Teorema de Pitágoras generalizado)**.** *Sea E un espacio euclídeo, y $\vec{u}, \vec{v}$ dos vectores ortogonales de E, entonces:*

$$||\vec{u} + \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2 \ . \tag{11}$$

Definición 3. *Dos subespacios W_1 y W_2 de un espacio euclídeo E se dicen* ortogonales*, y se escribe $W_1 \perp W_2$, si todos los vectores de W_1 son ortogonales a todos los vectores de W_2.*

En general, en un espacio euclídeo todo subespacio W de E tiene asociado otro formado por los vectores que son ortogonales a los de W, se llama *el complemento ortogonal de* W, y se denota por $W^{\perp}$. Es decir,

$$W^{\perp} := \{\vec{y} \in E | \ \vec{y} \perp \vec{x}, \ \text{para cada vector } \vec{x} \in W\} \tag{12}$$

Además, se dan las siguientes propiedades:

A. $W_1 \perp W_2$ si y sólo si todos los vectores de una base de W_1 son ortogonales a todos los vectores de una base de W_2.

B. Si $W_1 \perp W_2$, entonces $W_1 \cap W_2 = \{\vec{0}\}$.

C. Se tiene que $E = W + W^{\perp}$, y la suma es directa.

Ejemplo 1.3. *El complemento ortogonal del subespacio $W = L(\vec{u}, \vec{v})$, con $\vec{u} = (1,0,1)$, $\vec{v} = (2,-1,1)$ de $\mathbb{R}^3$ cuando se considera el producto escalar habitual es la recta $W^{\perp}$ de vector director $\vec{r} = (1,1,-1)$. Se tiene que $\mathbb{R}^3 = W + W^{\perp}$. Entonces, cada vector de $\mathbb{R}^3$ se pone de forma única como suma de dos vectores, uno quenestá en W (a veces llamada la componente horizontal del vector) y otro en $W^{\perp}$ (a veces llamada la componente vertical del vector). Por ejemplo:*

$$(7,2,2) = (4,-1,3) + (3,3,-1) = 2 \cdot (1,0,1) + (2,-1,1) + 3 \cdot (1,1,-1).$$

1.1. Espacios afines euclídeos. Distancia

En muchas aplicaciones el espacio vectorial $E = \mathbb{R}^n$ se considera como espacio de puntos $A = \mathbb{R}^n$ y no como espacio de vectores. Esto es debido a que no se establece a priori ninguna referencia cartesiana para calcular. En estas situaciones se dice que consideramos a $\mathbb{R}^n$ como espacio afín A asociado al espacio vectorial euclídeo $E = \mathbb{R}^n$. Por esta situación podemos definir el vector que une dos puntos $P = (p_1, \ldots, p_n)$, $Q = (q_1, \ldots, q_n)$ como

$$\vec{PQ} := (q_1 - p_1, \ldots, q_n - p_n). \tag{13}$$

También podemos definir una distandia entre puntos con la fórmula siguiente:

$$d(P,Q); = ||\vec{PQ}||. \tag{14}$$

Ilustraremos estos conceptos sobre un ejemplo.

Ejemplo 1.4. *Sea $E = \mathbb{R}^2$, con el producto escalar definido por la fórmula*

$$< \vec{u}, \vec{v} >:= (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \tag{15}$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E. Los puntos $P = (1, 1)$ y $Q = (1, 0)$ en $A = \mathbb{R}^2$ definen el vector $\vec{PQ} = (0, -1)$ cuyo módulo es $\sqrt{2}$. En consecuencia, para este producto escalar, se tiene que $d(P, Q) = \sqrt{2}$.

2. Ortogonalidad. Proceso de ortonormalización

Fijaremos en esta sección un espacio vectorial euclídeo E con producto escalar $<, >$, que supondremos de dimensión finita n. En E una base se dice *ortogonal* si sus elemetos son ortogonales dos a dos; si además, los elementos de una base ortogonal son de longitud 1, la base se dice *ortonormal*.
Si $B = \{\vec{e_1}, \ldots, \vec{e_n}\}$ es una base ortonormal de E, y $\vec{u}$ es un vector, entonces se da la siguiente igualdad

$$\vec{u} =< \vec{u}, \vec{e_1} > \vec{e_1} + \cdots + < \vec{u}, \vec{e_n} > \vec{e_n}, \tag{16}$$

es decir, las coordenadas $(\lambda_1, \ldots, \lambda_n)$ de $\vec{u}$ en B se calculan según la fórmula:

$$\lambda_1 =< \vec{u}, \vec{e_1} >, \ldots, \lambda_n =< \vec{u}, \vec{e_n} > . \tag{17}$$

Además, en este espacio siempre se puede construir una base ortogonal a partir de una dada; es el método (o algoritmo) de ortogonalización de Gram-Schmidt. Se fundamenta en el siguiente teorema.

Teorema 4. *(Teorem 8.2.1 de [1]) Sea $\vec{u}_1, \ldots, \vec{u}_n, \ldots$ una sucesión finita o infinita numerable de vectores linealmente independientes en un espacio euclídeo E, y sea $L_k = L(\vec{u}_1, \ldots, \vec{u}_k)$ el espacio vectorial generado por los k primeros vectores. Entonces, existe un conjunto de vectores $\vec{v}_1, \ldots, \vec{v}_n, \ldots$ tal que, para cada $k = 1, 2, \ldots,$ se verifica*

1. *El espacio vectorial $L'_k := L(\vec{v}_1, \ldots, \vec{v}_k)$ coincide con L_k.*

2. *El vector $\vec{v}_{k+1}$ es ortogonal a L_k.*

Observa que del teorema anterior se deduce que la recta de vector director $\vec{v}_{k+1}$ es ortogonal al espacio L_k (de dimensión k). Si denotamos por ℓ_{k+1} a esta recta, entonces escribimos abreviadamente $\ell_{k+1} \perp L_k$.
A continuación exponemos un algoritmo para calcular los vectores $\vec{v}_1, \ldots, \vec{v}_n, \ldots$ del teorema anterior. Es el llamado Algoritmo de ortogonalización de Gram-Schimdt.

2.1. Algoritmo de ortogonalización de Gram-Schimdt

El algoritmo de Gram-Schimdt consta de los siguientes pasos:

- **Input:** Un conjunto de vectores $B = \{\vec{u}_1, \ldots, \vec{u}_n\}$ linealmente independiente.

- **Output:** Un conjunto de vectores $B_o = \{\vec{v}_1, \ldots, \vec{v}_n\}$ linealmente independiente y ortogonales dos a dos.

1. Definimos $\vec{v_1} := \vec{u_1}$.
2. Para $i = 2, \dots, n$, calculamos

$$\vec{v_i} = \vec{u_i} - \left(\frac{< \vec{u_i}, \vec{v_1}, >}{||\vec{v_1}||^2} \vec{v_1} + \cdots + \frac{< \vec{u_i}, \vec{v_{i-1}}, >}{||\vec{v_{i-1}}||^2} \vec{v_{i-1}} \right) \tag{18}$$

3. Definimos $B_o := \{\vec{v}_1, \dots, \vec{v}_n\}$.

Ejemplo 2.1. *Sea $E = \mathbb{R}^2$, con el producto escalar definido por la fórmula*

$$< \vec{u}, \vec{v} >:= (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \tag{19}$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E. Supongamos que nos dan la base $B = \{\vec{u}_1 = (1, -1), \vec{u}_2 = (0, 1)\}$. Observa que

$$< \vec{u_1}, \vec{u_2} >= (1, -1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (1, -1) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -1 \neq 0.$$

Además $||\vec{u_1}||^2 =< \vec{u_1}, \vec{u_1} >= 1$. Aplicando el algoritmo de Gram-Scmidt a la base B obtenemos

$$\vec{v_1} = \vec{u_1} = (1, -1) \ , \quad \vec{v_2} = \vec{u_2} - \frac{< \vec{u_2}, \vec{v_1}, >}{||\vec{v_1}||^2} \vec{v_1} = (0, 1) - \frac{-1}{1}(1, -1) = (1, 0).$$

Luego la base ortogonal asociada a B es

$$B_o = \{\vec{v_1} = (1, -1), \vec{v_2} = (1, 0)\}$$

En consecuencial la base ortonormal asociada a B según el algoritmo de Gram.Schmidt es

$$\tilde{B} = \left\{ \vec{e_1} = \frac{1}{||\vec{v_1}||} \vec{v_1} = (1, -1), \vec{e_2} = \frac{1}{||\vec{v_2}||} \vec{v_2} = (1, 0) \right\},$$

ya que $||\vec{v_2}|| = 1$ para este productor escalar.

2.2. Proyecciones ortogonales

El concepto clásico de proyeccción de un vector de $\mathbb{R}^3$ sobre un plano que pasa por el origen admite una generación a espacios euclídeos de dimensión finita. También se puede generalizar en algunos casos a espacios de funciones de dimensión infinita. En el aprńdice A se muestra un ejemplo.
Describiremos a conticuación el concepto de proyección ortogonal sobre un subespacio W de un espacio euclídeo E.

Consideramos un espacio eiclídeo E de dimensión n, y un subespacio W de dimensión $d < n$ sobre el que queremos proyectar. Se tiene que

$$E = W + W^{\perp} \ , \text{ con } \ W \cap W^{\perp} = \{\vec{0}\}. \tag{20}$$

Consideremos una base ortonormal $B_1 = \{\vec{e}_1, \dots, \vec{e}_d\}$ de W, respectivamente $B_2 = \{\vec{e}_{d+1}, \dots, \vec{e}_n\}$ base ortonormal de $W^{\perp}$. Entonces $B = B_1 \cup B_2$ es una base de E, y en esta base se tiene que las coordenadas de $\vec{u} \in E$ son

$$(< \vec{u}, \vec{e_1} >, \dots, < \vec{u}, \vec{e_d} >, < \vec{u}, \vec{e}_{d+1} >, \dots, < \vec{u}, \vec{e_n} >) \,. \tag{21}$$

por (16). Entonces la *proyección sobre W* queda definida por:

$$pr_W : E \to W \ , \quad pr_W(\vec{u}) :=< \vec{u}, \vec{e_1} > \vec{e_1} + \cdots + < \vec{u}, \vec{e_d} > \vec{e_d}. \tag{22}$$

Ilustraremos la construcción anterior sobre un ejemplo.

Ejemplo 2.2. *Consideramos* $\mathbb{R}^3$ *con el producto escalar habitual. Sea* W *el plano engendrado por los vectores* $\vec{v_1} = (1, 0, -1), \vec{v_2} = (0, 1, -1)$*. Una base ortonormal de* W *es*

$$B_1 = \left\{ \vec{e_1} = \frac{1}{\sqrt{2}}(1, 0, -1),\ \vec{e_2} = \frac{1}{\sqrt{6}}(-1, 2, -1) \right\}$$

El complemento ortogonal de W *es la recta* $W^\perp$ *engendrada por* $\vec{e_3} = \frac{1}{\sqrt{3}}(1, 1, 1)$*. En consecuencia, en la base* $B = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ *se tiene*

$$pr_W : \mathbb{R}^3 \to W \ , \quad pr_W(\vec{u}) :=< \vec{u}, \vec{e_1} > \vec{e_1}+ < \vec{u}, \vec{e_2} > \vec{e_2}.$$

Por ejemplo, si $\vec{u} = (1, 2, 3) =< \vec{u}, \vec{e_1} > \vec{e_1}+ < \vec{u}, \vec{e_2} > \vec{e_2}+ < \vec{u}, \vec{e_3} > \vec{e_3} = \frac{-2}{\sqrt{2}}\vec{e_1} + 0 \cdot \vec{e_2} + \frac{6}{\sqrt{3}}\vec{e_3}$*, se tiene que*

$$pr_W(\vec{u}) = \frac{-2}{\sqrt{2}}\vec{e_1} + 0 \cdot \vec{e_2} = \frac{-2}{\sqrt{2}}\vec{e_1} = (-1, 0, 1) \ .$$

3. Aplicaciones autoadjuntas

Adjunta de una aplicación, Consideraremos en esta sección un espacio euclídeo E con un productor escalar $<, >$, y denotaremos por $\mathcal{L}(E)$ al conjunto de todas las aplicaciones lineales $f : E \to E$.

Definición 5. *Sea* $f \in \mathcal{L}(E)$*. Diremos que una aplicación* $f^* \in \mathcal{L}(E)$ *es adjunta de* f*, si satisface la siguiente condición:*

$$< f(\vec{u}), \vec{v} >=< \vec{u}, f^*(\vec{v}) > \ , \ \textit{para cualesquiera } \vec{u}, \vec{v} \in E.$$

Diremos que f *es autoadjunta si* $f = f^*$*.*

Para cualquier aplicación lineal f siempre existe su adjunta. En concreto se da el siguiente resultado.

Teorema 6. *Sea* $f \in \mathcal{L}(E)$ *una aplicación lineal en un espacio vectoria euclídeo de dimensión finita. Entonces, existe una única aplicación adjunta de* f *cuya matriz en una base ortonormal es la matriz traspuesta de la matriz de* f*.*

Tenemos las siguientes propiedades de la aplicación adjunta. Consideremos $A, B \in \mathcal{L}(E)$.

1. $I^* = I$.
2. $(A^*)^* = A$.
3. $(A + B)^* = A^* + B^*$.
4. $(A \circ B)^* = B^* \circ A^*$.
5. Si A posee inversa, entonces $(A^{-1})^* = (A^*)^{-1}$.

Las transformaciones autoadjuntas presentan algunas características propias que detallaremos a continuación.

Proposición 7. *Sea* $f \in \mathcal{L}(E)$ *autoadjunta. Entonces en una base ortonormal de* E *se tiene que la matriz de* A *de* f *es simétrica. Además, se dan las siguientes propiedades:*

1. *Para $f, g \in \mathcal{L}(E)$ autoadjuntas, entonces $f+g$ es autoadjunta.*

2. *Si f es autoadjunta y posee inversa, entonces F^{-1} es autoadjunta.*

3. *Para $f, g \in \mathcal{L}(E)$ autoadjuntas y tales que conmutan, es decir $f \circ g = g \circ f$, entonces $f \circ g$ es autaadjunta.*

Las transformaciones autoadjuntas tienen mucha importancia en distintos campos científicos debido al resultado que enunciamos a continuación.

Teorema 8. *Sea $f \in \mathcal{L}(E)$ una aplicación autoadjunta en un espacio euclídeo de dimensión finita. Entonces, existe una base ortonormal (de autovectores) en la que la matriz A de f es diagonal. Además todos sus autovalores son reales.*

El siguiente ejemplo ilustra el teorema anterior.

Ejemplo 3.1. *Consideremos en $\mathbb{R}^3$ el producto escalar habitual. Sea*

$$f(x, y, z) = \left(\frac{3}{2}x - \frac{1}{2}y, -\frac{1}{2}x + \frac{3}{2}y, -z\right).$$

Un cálculo fácil demuestra que

$$< f(\vec{u}), \vec{v} > = < \vec{u}, f(\vec{v}) > \text{ , para cualesquiera } \vec{u}, \vec{v} \in \mathbb{R}^3.$$

Por otro lado, los autovalores de f son $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$. Los autoespacios correspondientes son:

$$E(1) = (1, 1, 0)\mathbb{R}, \quad E(2) = (1, -1, 0)\mathbb{R}, \quad E(-1) = (0, 0, 1)\mathbb{R}.$$

La base ortonormal es por tanto $\vec{v}_1 = \frac{1}{\sqrt{2}}(1, 1, 0)$, $\vec{v}_2 = \frac{1}{\sqrt{2}}(1, -1, 0)$, $\vec{v}_3 = (0, 0, 1)$. En esta base la matriz de f es:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Nota 3.1. *En el ejemplo anterior se observa que cada autoespacio W es invariante por f, es decir $f(W) \subset W$. En general, las transformaciones autoadjuntas satisfacen que si un subespacio genérico, W es invariante, entonces su ortonal $W^{\perp}$ también es invariante.*

4. Aplicaciones ortogonales.

Consideraremos en esta sección un espacio euclídeo E con un productor escalar $<, >$, y denotaremos por $\mathcal{L}(E)$ al conjunto de todas las aplicaciones lineales $f : E \to E$.

Definición 9. *Sea $f \in \mathcal{L}(E)$. Diremos que f es una aplicación ortogonal si satisface la siguiente condición:*

$$< f(\vec{u}), f(\vec{v}) > = < \vec{u}, \vec{v} > \text{ , para cualesquiera } \vec{u}, \vec{v} \in E.$$

Nota 4.1. *Se suele decir que una aplicación es ortogonal si conserva el producto escalar. Observa que en este caso conserva la longitud de los vectores, pues $||\vec{u}||^2 = < \vec{u}, \vec{u} > = < f(\vec{u}), f(\vec{u}) > = ||f(\vec{u})||^2$.*

Este tipo de transformaciones verifican las siguientes propiedades.

Proposición 10. *Sea f una aplicación lineal ortogonal en un espacio euclídeo E. Entonces,*

1. *f transforma bases ortonormales en bases ortonormales.*
2. *Los autovalores reales de f son 1 ò -1 (puede tener también autovalores complejos de módulo 1).*
3. *Si W es un subespacio invariante por f, entonces $W^{\perp}$ es también invariante por f.*

Denotaremos por $\mathrm{O}(n;\mathbb{R})$ a las transformaciones ortogonales en $\mathbb{R}^n$. En este conjunto encontramos las matrices A de tamaño $n\times n$ tales que $AA^t = I_n$. Es decir,

$$\mathrm{O}(n;\mathbb{R}) = \left\{ A \in \mathbb{M}_{n,n}(\mathbb{R}) \mid AA^t = I_n \right\}.$$

Ejemplo 4.1. *Consideremos en $\mathbb{R}^3$ una base ortonorma $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, las sigientes transformaciones son ortogonales:*

- *La identidad, cuya matriz es* $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- *La simetría respecto del plano π de ecuación $x = 0$, cuya matriz es* $S_\pi = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- *La simetría con respecto al eje $\ell = OX$, cuya matriz es* $S_\ell = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.
- *La simetría respecto al origen O (ò simetría central), cuya matriz es* $S_O = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.
- *El giro de eje OX y ángulo α, cuya matriz es* $G_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$.
- *El giro de eje OX y ángulo α compuesto con una simetría respecto del plano $x = 0$ (ò simetría rotatoria), cuya matriz es* $\sigma_\alpha = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$. *Observa que se da la igualdad $\sigma_\alpha = S_\ell \circ G_\alpha$.*

Finalmente podemos enunciar el llamado **Teorema Espectral** que da los modelos matriciales de las transformaciones ortogonales y autoadjunta. Este teorema tiene como consecuancia el teorema 12 que da la estructura de las aplicaciones lineales inversibles en los espacios euclídeos de dimensión finita.

Teorema 11 (Teorema espectral)**.** *1. Dada una transformación ortogonal O en $O(n;\mathbb{R})$, existen una transformación ortogonal $C_1 \in O(n;\mathbb{R})$ y una transformación del tipo*

$$J = \begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \cdots & 0 \\ 0 & 0 & G_{\alpha_1} & \cdots & 0 \\ 0 & 0 & \cdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & G_{\alpha_k} \end{pmatrix}, \text{ donde } G_{\alpha_i} \text{ es el giro de ángulo } \alpha_i\ , \tag{23}$$

tales que se da la factorización:

$$O = C_1 J C_1^t \ . \tag{24}$$

2. *Dada una transformación autoadjunta S en E, entonces existen Una transformación ortogonal $C_2 \in O(n;\mathbb{R})$, y una matriz diagonal $D \in \mathbb{M}_{n,n}(\mathbb{R})$ tal que*

$$S = C_2 D C_2^t \ . \tag{25}$$

Teorema 12. *Dada una transformación invertible $A \in GL(n;\mathbb{R})$, entonces existen una transformación ortogonal $O \in O(n;\mathbb{R})$, y una transformación autoadjunta S tal que se da la factorización:*

$$A = OS \ . \tag{26}$$

Ejemplo 4.2. *Observa que se da la siguiente factorización de la matriz* $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{pmatrix} = G_{\pi/4} \cdot D,$$

donde $G_{\pi/4}$ es el giro de ángulo $\pi/4$ radianes y D es una matriz diagonal.

Apéndice

A. Sistemas ortogonales de funciones

Consideremos en esta sección un intervalo I de la recta real. Denotaremos por $L(I)$ las funciones integrables en I, y por $L^2(I)$ a las *funciones cuadrado integrable*, es decir

$$L^2(I) = \left\{ f : I \to \mathbb{C} \ : \ ||f||^2 \in L(I) \ \right\}. \tag{27}$$

El producto interior de dos funciones $f, g \in L^2(I)$ se define por

$$< f, g >:= \int_I f(x)\overline{g(x)} dx \tag{28}$$

El número no negativo $||f|| = \sqrt{\langle f, f \rangle}$ existe siempre para funciones de este espacio y se llama *la norma L^2 de f.*

Definición 13. *Sea $\mathcal{S} = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$ una familia de funciones de $L^2(I)$. Si*

$$< \varphi_n, \varphi_m >= 0 \quad , \text{ siempre que } m \neq n, \tag{29}$$

La familia $\mathcal{S}$ se llama un sistema ortogonal en I. Si además, cada φ_n tiene norma 1, entonces $\mathcal{S}$ se llama un sistema ortonormal en I

Nos interesará particularmente el sistema trigonométrico especial $\mathcal{S} = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$, en donde

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \ \varphi_{2n-1}(x) = \frac{1}{\sqrt{\pi}} \cos(nx), \ \varphi_{2n}(x) = \frac{1}{\sqrt{\pi}} \sin(nx), \tag{30}$$

para $n = 1, 2, \dots$. Es fácil comprobar que $\mathcal{S}$ es ortonormal en todo intervalo de longitud 2π. El sistema (30) está formado por funciones reales. Un sistema ortonormal de funciones complejas en todo intervalo de longitud 2π lo constituye las funciones

$$\psi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}} = \frac{\cos(nx) + i\sin(nx)}{\sqrt{2\pi}} \quad , \text{ para } n = 0, 1, 2, \dots \tag{31}$$

A.1. El teorema de óptima aproximación

Sea I un intervalo de $\mathbb{R}$ y consideremos el espacio $L^2(I)$ definido en (27). Sea $\mathcal{S} = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$ un sistema ortonormal en I. Definimos el subespacio vectorial $W_n \subset L^2(I)$ generado por $\varphi_0, \varphi_1, \dots \varphi_n$. Consideraremos la siguiente sucesión creciente de subespacios

$$W_1 \subset W_2 \subset \cdots \subset W_n \subset W_{n+1} \subset \cdots .$$

Es de señalar que, por construcción de $\mathcal{S}$, esta sucesión es infinita. El siguiente teorema muestra cómo podemos utilizar el sistema $\mathcal{S}$ para aproximar una función dada f en $L^2(I)$. De hecho, el resultado afirma que si definimos la función $s_n(x) \in W_n$ como en la fórmula (32), esta combilación lineal es la mejor aproximación de f por elementos del subespacio W_n, véase (34).

Teorema 14 (Teorema de óptima aproximación)**.** *Sea $\mathcal{S} = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$ un sistema ortonormal en I, y supongamos que $f \in L^2(I)$. Definimos dos sucesiones de funciones*

$$s_n(x) := \sum_{k=0}^{n} c_k \varphi_k(x) \ , \quad \textit{en donde } c_k :=< f, \varphi_k >, k = 0, 1, 2, \dots \tag{32}$$

y también

$$t_n(x) := \sum_{k=0}^{n} b_k \varphi_k(x) \ , \quad \textit{con } b_k \textit{ números complejos arbitrarios } k = 0, 1, 2, \dots . \tag{33}$$

Entonces, para cada número natural n se tiene la desigualdad

$$||f - s_n|| \leq ||f - t_n||. \tag{34}$$

Además, en (34) se verifica la igualdad si y sólo si $b_k = c_k$ para $k = 0, 1, 2, \dots, n$.

Nota A.1. *Consideremos dos aproximaciones sucesivas de $f \in L^2(I)$, pongamos $s_\ell(x) \in W_k$ y $s_{\ell+1}(x) \in W_{\ell+1}$. Aplicando el teorema (14) para $n = \ell + 1$ obtenemos la desigualdad*

$$||f - s_{\ell+1}|| \leq ||f - t|| \ ,$$

donde t es cualquier elemento de $W_{\ell+1}$. En consecuencia, si $t = s_\ell$, la desigualdad anterior se reescribe como

$$||f - s_{\ell+1}|| \leq ||f - s_\ell||,$$

es decir, a medida que aumentamos el número de sumandos en la expresión $\sum_{k=0}^{n} c_k \varphi_k(x)$, la distancia entre f y s_n se hacer menor; luego $s_{\ell+1}$ será una mejor aproximación de f que s_ℓ.

Ejemplo A.1. *La función de $L^2([-\pi, \pi])$*

$$f(x) = \begin{cases} -1 & \textit{para } -\pi \leq x < 0 \\ 1 & \textit{para } \ 0 \leq x < \pi \end{cases} ,$$

tiene por coeficientes de Fourier[1]

$$c_0 = 0 \ , \ c_{2n-1} = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(t) \cos(nt) dt = 0 \ ,$$

$$c_{2n} = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(t) \sin(nt) dt \ = \begin{cases} 0 & \textit{si } n \textit{ es par} \\ \frac{4}{n\sqrt{\pi}} & \textit{si } n \textit{ es impar} \end{cases} .$$

[1] Con una cantidad infinita de funciones aproximantes se consigue la aproximación

$$f(x) \sim \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin((2n+1)x).$$

Referencias

[1] E. Hernández, M. J. Vázquez, and M. A. Zurro. *Álgebra Lineal y Geometría*. Pearson, 2012.