

Tema 4: Espacios vectoriales euclídeos

Introducción

En este tema sólo vamos a considerar espacios vectoriales sobre $\mathbb{R}$. Lo que queremos es obtener formas de medir vectores y distancias entre ellos y esto nos va a servir para poder aproximar funciones.

Producto escalar

Definición: Si E es un $\mathbb{R}$ - e.v. se dice que una función

$$\langle, \rangle: E \times E \rightarrow \mathbb{R}$$

$$(u, v) \rightarrow \langle u, v \rangle$$

es un producto escalar real (o producto interior real) si:

1. $\forall u, v, w \in E, \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
2. $\forall u, v \in E, \forall \alpha \in \mathbb{R}, \langle \alpha u, v \rangle = \alpha \langle u, v \rangle$
3. $\forall u, v \in E, \langle u, v \rangle = \langle v, u \rangle$
4. $\forall u \in E, \langle u, u \rangle \geq 0$ y $\langle u, u \rangle = 0 \Leftrightarrow u = 0$

Si $\langle, \rangle: E \times E \rightarrow \mathbb{R}$ es un producto escalar, al par $(E, \langle, \rangle)$ le denominaremos espacio vectorial euclídeo.

Ejemplos:

- En $\mathbb{R}^2$ la función

$$\langle, \rangle: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

definida para cada par de vectores $(x, y), (x', y') \in \mathbb{R}^2$ mediante la expresión

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1 y_1 + x_2 y_2$$

es un producto escalar. Además este producto escalar es el que llamamos producto escalar usual de $\mathbb{R}^2$.

- En general en el $\mathbb{R}$ - e.v. $\mathbb{R}^n$ la función

$$\langle, \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

definida para cada par de vectores

$$(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$$

mediante la expresión

$$\langle (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \rangle = x_1 y_1 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i$$

es un producto escalar. Además es el usual de $\mathbb{R}^n$.

- En el $\mathbb{R}$ - e.v. de los polinomios con coeficientes reales $\mathbb{R}[x]$ consideraremos los productos escalares siguientes:

$$\langle, \rangle_1: \mathbb{R}[x] \times \mathbb{R}[x] \rightarrow \mathbb{R}$$

$$(p(x), q(x)) \rightarrow \int_{-1}^1 p(x) \cdot q(x) dx$$

$$\langle, \rangle_0: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

$$(p(x), q(x)) \rightarrow \sum_{i=0}^{\min\{n, m\}} a_i b_i$$

donde $p(x) = a_n x^n + \dots + a_1 x + a_0$ y $q(x) = b_m x^m + \dots + b_1 x + b_0$.

Observación: Nótese que si $(E, \langle, \rangle)$ es un espacio vectorial euclídeo, a partir de las propiedades del producto escalar se verifica que:

$$\forall u, v, w \in E \quad \langle u, w + v \rangle = \langle w + v, u \rangle = \langle w, u \rangle + \langle v, u \rangle = \langle u, w \rangle + \langle u, v \rangle$$

y

$$\forall u, v \in E, \forall \alpha \in \mathbb{R}, \langle u, \alpha v \rangle = \langle \alpha v, u \rangle = \alpha \langle v, u \rangle = \alpha \langle u, v \rangle$$

Además si $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ y $u_1, \dots, u_n, v \in E$

$$\langle \alpha_1 u_1, \dots, \alpha_n u_n, v \rangle = \alpha_1 \langle u_1, v \rangle + \dots + \alpha_n \langle u_n, v \rangle = \sum_{i=1}^n \alpha_i \langle u_i, v \rangle$$

Se sigue entonces que un producto escalar es lineal en la primera y en la segunda variable, es decir, es una función bilineal.

Longitud o norma euclídea de un vector

Definición: Si $(E, \langle, \rangle)$ es un e.v. y $u \in E$, una norma (o módulo) es un número real, $\|u\|$, tal que

1. $\forall u \in E; \|u\| \geq 0$ y $\|u\| = 0 \Leftrightarrow u = 0$
2. $\forall u \in E, \forall k \in \mathbb{R}; \|ku\| = |k| \|u\|$
3. $\forall u, v \in E; \|u + v\| \leq \|u\| + \|v\|$ (desigualdad triangular)

Definición: Si $(E, \langle, \rangle)$ es un e.v. y $u \in E$, llamaremos norma euclídea (o módulo) de u al número real

$$\|u\| = \sqrt{\langle u, u \rangle}$$

Propiedades de la norma euclídea

1. $\forall u \in E, \|u\| \geq 0$ y $\|u\| = 0 \Leftrightarrow \langle u, u \rangle = 0 \Leftrightarrow u = 0$
2. $\forall u \in E, \forall k \in \mathbb{R}$

$$\|ku\| = \sqrt{\langle ku, ku \rangle} = \sqrt{k^2 \langle u, u \rangle} = |k| \cdot \|u\|$$
3. $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$ (ley del paralelogramo)
4. $\langle u, v \rangle = \frac{1}{2}(\|u + v\|^2 - \|u\|^2 - \|v\|^2)$
5. $\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$

Teorema: (Desigualdad de Cauchy-Schwarz) Si $(E, \langle, \rangle)$ es un espacio vectorial euclídeo, $\forall u, v \in E$ se verifica que

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$$

Además

$$|\langle u, v \rangle| = \|u\| \cdot \|v\| \Leftrightarrow \{u, v\} \text{ es ligado}$$

Proposición: Si $(E, \langle, \rangle)$ es un espacio vectorial euclídeo, $\forall u, v \in E$ se verifica que

$$\|u + v\| \leq \|u\| + \|v\|$$

Definición: Si $(E, \langle, \rangle)$ es un e.v. euclídeo y $u, v \in E$, $u \neq 0, v \neq 0$ se denomina ángulo entre los vectores de u y v al número real $\hat{u}v \in [0, \pi]$ tal que

$$\cos(\hat{u}v) = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$$

Definición: Si $(E, \langle, \rangle)$ es un e.v. euclídeo y $u, v \in E$, se dice que u y v son ortogonales (o perpendiculares) si

$$\langle u, v \rangle = 0$$

Ejemplos: En el e.v.e. $(\mathbb{R}^3, \langle, \rangle)$, donde $\langle, \rangle$ es el producto escalar usual, determinar el ángulo formado por los vectores $(1, 0, 1)$ y $(0, 0, 1)$.

Definición: Si $(E, \langle, \rangle)$ es un e.v. euclídeo y $u, v \in E$, se denomina distancia euclídea de u a v y se denota por $d(u, v)$ al número real

$$d(u, v) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}$$

A partir de las propiedades vistas de la norma euclídea, es fácil ver que se satisfacen las propiedades que debe verificar cualquier función distancia, a saber:

- $\forall x, y \in E, d(x, y) \geq 0$
- $\forall x, y \in E, d(x, y) = 0 \Leftrightarrow x = y$
- $\forall x, y, z \in E, d(x, z) \leq d(x, y) + d(y, z)$

Ejemplos: En el e.v.e. $(\mathbb{R}^3, \langle, \rangle)$, donde $\langle, \rangle$ es el producto escalar usual, determinar la distancia entre los vectores $(1, 0, 1)$ y $(0, 1, 1)$.

Método de ortogonalización de Gram-Schmidt

Definición: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo y $\{u_1, \dots, u_r\}$ un sistema de vectores de E . Se dice que el sistema $\{u_1, \dots, u_r\}$ es ortogonal si se verifica:

1. $\forall i \in \{1, \dots, r\}, u_i \neq 0$
2. $\forall i, j \in \{1, \dots, r\}$ si $i \neq j \Rightarrow \langle u_i, u_j \rangle = 0$

Si además de ser un sistema ortogonal, el sistema $\{u_1, \dots, u_r\}$ satisface

1. $\forall i \in \{1, \dots, r\}, \|u_i\| = 1$

se dice que el sistema $\{u_1, \dots, u_r\}$ es ortonormal.

Proposición: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo y $\{u_1, \dots, u_r\}$ un sistema de vectores de $E - \{0\}$. Se verifica que

$$\{u_1, \dots, u_r\} \text{ ortogonal} \rightarrow \{u_1, \dots, u_r\} \text{ libre}$$

Definición: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo y $B = \{u_1, \dots, u_n\}$ una base de E . Se dice que B es una base ortogonal si el sistema $\{u_1, \dots, u_n\}$ es ortogonal y, obviamente, se dice que B es una base ortonormal si el sistema $\{u_1, \dots, u_n\}$ es ortonormal.

Obviamente si $(E, \langle, \rangle)$ es un espacio vectorial euclídeo tal que $\dim(E) = n$ y $\{u_1, \dots, u_n\}$ es un sistema ortogonal, necesariamente $\{u_1, \dots, u_n\}$ es una base ortogonal, siendo libre como consecuencia de la proposición anterior.

Teorema: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo y $\{u_1, \dots, u_r\}$ un sistema libre de E . En estas condiciones existe un sistema ortogonal $\{v_1, \dots, v_r\}$ tal que

$$\forall i \in \{1, \dots, r\} \quad L[\{v_1, \dots, v_i\}] = L[\{u_1, \dots, u_i\}]$$

MÉTODO DE ORTOGONALIZACIÓN DE GRAM-SCHMIDT

Dado el sistema libre $\{u_1, \dots, u_r\}$ el método de ortogonalización de Gram-Schmidt consiste en definir el sistema de vectores

$$\begin{aligned} v_1 &= u_1 \\ v_2 &= u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} \cdot v_1 \\ v_3 &= u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} \cdot v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} \cdot v_2 \end{aligned}$$

...

$$v_r = u_r - \frac{\langle u_r, v_1 \rangle}{\langle v_1, v_1 \rangle} \cdot v_1 - \dots - \frac{\langle u_r, v_{r-1} \rangle}{\langle v_{r-1}, v_{r-1} \rangle} \cdot v_{r-1}$$

que es un sistema ortogonal y además satisface la condición

$$\forall i \in \{1, \dots, r-1\} \quad L[\{v_1, \dots, v_i\}] = L[\{u_1, \dots, u_i\}]$$

Observación: Nótese que si $\{v_1, \dots, v_r\}$ es ortogonal, definiendo

$$\forall i \in \{1, \dots, r\} \quad w_i = \frac{v_i}{\|v_i\|}$$

el sistema $\{v_1, \dots, v_r\}$ es ortonormal, puesto que

$$\forall i \in \{1, \dots, r\} \quad \|w_i\| = \left\| \frac{v_i}{\|v_i\|} \right\| = \frac{1}{\|v_i\|} \cdot \|v_i\| = 1$$

Por tanto, podemos afirmar que todo espacio vectorial euclídeo de dimensión finita tiene una base ortonormal.

Ejemplos: A partir de la base

$$\{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$$

obtener por el método de Gram-Schmidt una base ortonormal de $(\mathbb{R}^3, \langle, \rangle)$ donde $\langle, \rangle$ es el producto escalar habitual.

Proyecciones ortogonales

Definición: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo y $A \subseteq E$. Diremos que $v \in E$ es un vector ortogonal a A si $\forall u \in A, \langle v, u \rangle = 0$.

Ejemplo: En $(\mathbb{R}^3, \langle, \rangle)$, donde $\langle, \rangle$ es el producto escalar usual de $\mathbb{R}^3$, cualquier vector de la recta $L[(0, 0, 1)]$ es ortogonal al conjunto $A = \{(1, 1, 0), (0, 1, 0)\}$.

Observación: Nótese que el vector $0 \in E$ es ortogonal a cualquier subconjunto $A \subseteq E$. Pues, siendo A cualquier subconjunto de E , $\forall u \in A$ se verifica que

$$\langle u, 0 \rangle = \langle u, u - u \rangle = \langle u, u \rangle - \langle u, u \rangle = 0$$

ó

$$\langle u, 0 \rangle = \langle u, 0 \cdot 0 \rangle = 0 \cdot \langle u, 0 \rangle = 0$$

Proposición: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo y $A \subseteq E$. Se verifica que el conjunto

$$A^\perp = \{v \in E | \forall u \in A, \langle v, u \rangle = 0\}$$

es un subespacio vectorial de E . A este subespacio se le denomina subespacio ortogonal de A .

Observación: En el caso particular de que $(E, \langle, \rangle)$ sea un espacio vectorial euclídeo de dimensión finita y $H \prec E$, al subespacio ortogonal de H se le denomina complemento ortogonal de H .

Proposición: Sean $(E, \langle, \rangle)$ un espacio vectorial euclídeo con $\dim(E) = n$ y $H \prec E$. En estas condiciones se verifica que

$$\dim(H^\perp) = n - \dim(H)$$

Observación: Si $\{w_1, \dots, w_r\}$ es una base ortonormal de un subespacio vectorial H y $\{w_1, \dots, w_r, w_{r+1}, \dots, w_n\}$ es una base ortonormal del espacio E , por la proposición anterior $\{w_{r+1}, \dots, w_n\}$ es una base de $H^\perp$, con lo que

$$\forall u \in E \quad \exists! (\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n) \in \mathbb{R}^n \text{ tales que}$$

$$u = \alpha_1 w_1 + \dots + \alpha_r w_r + \alpha_{r+1} w_{r+1} + \dots + \alpha_n w_n$$

Si denotamos por

$$h = \alpha_1 w_1 + \dots + \alpha_r w_r \in H$$

y por

$$h^\perp = \alpha_{r+1} w_{r+1} + \dots + \alpha_n w_n \in H^\perp$$

resulta que

$$\forall u \in E \quad \exists! h \in H \quad \text{y} \quad \exists! h^\perp \in H^\perp \quad \text{tales que} \quad u = h + h^\perp$$

Definición: Siendo $(E, \langle, \rangle)$ un espacio vectorial euclídeo de dimensión finita $H \prec E$ y $u \in E$ si

$$u = h + h^\perp$$

con $h \in H$ y $h^\perp \in H^\perp$, al vector h se le denomina proyección ortogonal de u sobre el subespacio H y se le denota por $p_H^\perp(u)$.

Método para hallar una proyección ortogonal

A partir de una base del subespacio H obtenemos una base ortonormal $\{w_1, \dots, w_r\}$ de H . La proyección de u sobre H será entonces el vector:

$$p_H^\perp(u) = h = \alpha_1 w_1 + \dots + \alpha_r w_r = \langle u, w_1 \rangle w_1 + \dots + \langle u, w_r \rangle w_r$$

Ejemplo: En $(\mathbb{R}^3, \langle, \rangle)$ donde $\langle, \rangle$ es el producto escalar usual de $\mathbb{R}^3$, obtener la proyección ortogonal del vector $(1, 2, 1)$ sobre los subespacios $H = L[(1, 0, 1)]$ y $H' = L[(1, 0, 1), (1, 0, 0)]$

Teorema: Sean $\{w_1, \dots, w_r\}$ un sistema ortonormal del $(E, \langle, \rangle)$ y $H = L[(w_1, \dots, w_r)]$. Entonces, para todo $u \in E$, se verifica que

$$d(u, p_H^\perp(u)) \leq d(u, w) \quad \forall w \in H$$

Es decir, $p_H^\perp(u) = \sum_{i=1}^r \langle u, w_i \rangle w_i$ es el vector de $H = L[(w_1, \dots, w_r)]$ que está a menor distancia de u (es decir, el que mejor lo aproxima).

Solución de un sistema de ecuaciones por mínimos cuadrados

Dado un sistema de ecuaciones del tipo $Ax = b$ que NO tiene solución, lo que queremos es encontrar una solución muy cerca del sistema, es decir, una aproximación a una solución variando el sistema levemente. Mediante este método, queremos encontrar una x tal que $\|b - Ax\|$ sea mínima. Tenemos dos métodos de resolución por mínimos cuadrados.

1. Expresamos vectorialmente nuestro sistema y por tanto tenemos que es $x_1 a_1 + x_2 a_2 + \dots + x_n a_n = b$. Como no tiene solución, tenemos que $b \notin L[\{a_1, a_2, \dots, a_n\}] = S$. Ahora hallamos una b' que sea la mejor aproximación de b al subespacio S , mediante el método visto anteriormente (es decir, calcular su proyección ortogonal sobre S). Una vez hallado, planteamos el sistema $Ax = b'$ y lo resolvemos, obteniendo una aproximación por mínimos cuadrados de la solución original. El error cometido es $\|b - b'\|$
2. Si tenemos que las columnas de A son linealmente independientes, entonces tenemos que $A^t \cdot A$ es invertible y por tanto tenemos que el sistema $Ax = b$ tiene una solución por mínimos cuadrados y es $x = (A^t \cdot A)^{-1} \cdot A^t \cdot b$.

Fórmula para aproximar una recta

Dada una cantidad de puntos n obtenidas mediante observación y que no forman una recta en el plano, podemos aplicar una fórmula para obtener la mejor aproximación de la recta con esos puntos. Si la recta que queremos obtener tiene la forma $y = mx + b$ entonces:

$$m = \frac{\sum_i x_i y_i - \frac{\sum_i x_i \sum_i y_i}{n}}{\sum_i x_i^2 - \frac{(\sum_i x_i)^2}{n}} \quad b = \frac{\sum_i y_i}{n} - m \frac{\sum_i x_i}{n}$$