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

$$<,>: E \times E \to \mathbb{R}$$

$$(u, v) \to < u, v >$$

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

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

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

Ejemplos:

- En $\mathbb{R}^2$ la función
  $$<,>: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$
  definida para cada par de vectores $(x, y), (x', y') \in \mathbb{R}^2$ mediante la expresión
  $$< (x_1, x_2), (y_1, y_2) >= x_1 y_1 + x_2 y_2$$
  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
  $$<,>: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
  definida para cada par de vectores
  $$(x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) \in \mathbb{R}^n$$
  mediante la expresión
  $$< (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) >= x_1 y_1 + ... + x_n y_n = \sum_{i=1}^{n} x_i y_i$$
  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:

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

$$\forall u, v, w \in E \quad < u, w + v > = < w + v, u > = < w, u > + < w, v > = < u, w > + < v, w >$$

y

$$\forall u, v \in E, \forall \alpha \in \mathbb{R}, < u, \alpha v > = < \alpha v, u > = \alpha < v, u > = \alpha < u, v >$$

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

$$< \alpha_1 u_1, ..., \alpha_n u_n, v > = \alpha_1 < u_1, v > + ... + \alpha_n < u_n, v > = \sum_{i=1}^{n} \alpha_i < u_i, v >$$

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 normal euclídea de un vector

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

$$||u|| = \sqrt{< u, u >}$$

Propiedades de la norma euclídea

1. $\forall u \in E, ||u|| \geq 0$ y $||u|| = 0 \Leftrightarrow < u, u > = 0 \Leftrightarrow u = 0$
2. $\forall u \in E, \forall k \in \mathbb{R}$
$$||ku|| = \sqrt{< ku, ku >} = \sqrt{k^2 < u, u >} = |k| \cdot ||u||$$
3. $||u + v||^2 + ||u - v||^2 = 2(||u||^2 + ||v||^2)$ (ley del paralelogramo)
4. $< u, v > = \frac{1}{2}(||u + v||^2 - ||u||^2 - ||v||^2)$
5. $< u, v > = \frac{1}{4}(||u + v||^2 - ||u - v||^2)$

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

$$| < u, v > | \leq ||u|| \cdot ||v||$$

Además

$$| < u, v > | = ||u|| \cdot ||v|| \Leftrightarrow \{u, v\} \;\; es \;\; ligado$$

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

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

Definición: Si $(E, <,>)$ 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

$$d(u,v) = ||u - v|| = \sqrt{< u - v, u - v >}$$

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. $(\mathbb{R}^3, <,>)$, donde $<,>$ 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 $(E, <,>)$ un espacio vectorial euclídeo y $\{u_1, ..., u_r\}$ un sistema de vectores de E. Se dice que el sistema $\{u_1, ..., u_r\}$ es ortogonal si se verifica:

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

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

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

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

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

$$\{u_1, ..., u_r\} \;\; ortogonal \rightarrow \{u_1, ...u_r\} \;\; libre$$

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

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

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

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

...

$$v_r = u_r - \frac{< u_r, v_1 >}{< v_1, v_1 >} \cdot v_1 - ... - \frac{< u_r, v_{r-1} >}{< v_{r-1}, v_{r-1} >} \cdot v_{r-1}$$

que es un sistema ortogonal y además satisface la condición

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

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

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

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

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

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 $(\mathbb{R}^3, <,>)$ donde $<,>$ es el producto escalar habitual.

Proyecciones ortogonales

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

Ejemplo: En $(\mathbb{R}^3, <,>)$, donde $<,>$ 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

$$< u, 0 >=< u, u - u >=< u, u > - < u, u >= 0$$

ó

$$< u, 0 >=< u, 0 \cdot 0 >= 0 \cdot < u, 0 >= 0$$

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

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

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

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

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

Si denotamos por

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

y por

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

resulta que

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

Definición: Siendo $(E, <,>)$ 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 $\{w_1, ..., w_r\}$ de H. La proyección de u sobre H será entonces el vector:

$$p_H^{\perp}(u) = h = \alpha_1 w_1 + ... + \alpha_r w_r = <u, w_1> w_1 + ... + <u, w_r> w_r$$

Ejemplo: En $(\mathbb{R}^3, <,>)$ donde $<,>$ 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 $\{w_1, ..., w_r\}$ un sistema ortonormal del $(E, <,>)$ y $H = L[(w_1, ..., w_r)]$. 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, $p_H^{\perp}(u) = \sum_{i=1}^{r} < u, w_i > w_i$ es el vector de $H = L[\{(w_1, ..., w_r)\}]$ 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 $x_1 a_1 + x_2 a_2 + ... + x_n a_n = b$. Como no tiene solución, tenemos que $b \notin L[\{a_1, a_2, ..., a_n\}] = S$. 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 cuadradaos 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 $x = (A^t \cdot A)^{-1} \cdot A^t \cdot b$.