Tema 10

Bases ortogonales

En el Tema 9 derivamos conceptos tales como la distancia y el ángulo entre vectores (y con éste el de ortogonalidad) o la longitud de vectores a partir de la noción de producto interno. Si dotamos a un espacio vectorial de tal producto, también podemos aplicar el término "ortogonal" a ciertos subespacios vectoriales. En lo que sigue, vamos a aplicar dicho término a otros conceptos.

10.1. Conjuntos ortogonales y bases ortogonales

Un conjunto de vectores $\{u_1, \dots, u_p\}$ de V se denomina **conjunto ortogonal** si cada par de vectores diferentes del conjunto es ortogonal, es decir, si $\langle u_i, u_j \rangle = 0$ siempre que $i \neq j$.

Ejemplo

Vamos a demostrar que $\{u_1, u_2, u_3\} \subset \mathbb{R}^3$ con $u_1 = (3,1,1)^t$, $u_2 = (-1,2,1)^t$ y $u_3 = \dfrac{1}{2}(-1,-4,7)^t$ es un conjunto ortogonal $\boxed{\text{respecto al producto escalar usual.}}$
Debemos considerar los tres posibles pares de vectores diferentes:

$$\begin{aligned} \langle u_1, u_2 \rangle &= 3 \cdot (-1) + 1 \cdot 2 + 1 \cdot 1 = 0\,, \\ \langle u_1, u_3 \rangle &= 3 \cdot (-1)\frac{1}{2} + 1 \cdot (-2) + 1 \cdot \frac{7}{2} = 0\,, \\ \langle u_2, u_3 \rangle &= -1 \cdot (-1)\frac{1}{2} + 2 \cdot (-2) + 1 \cdot \frac{7}{2} = 0\,. \end{aligned}$$

Por tanto, $\{u_1, u_2, u_3\}$ es un conjunto ortogonal.

Teorema

Si $S = \{u_1, \dots, u_p\}$ es un conjunto ortogonal de vectores no nulos de V, entonces S es linealmente independiente.

Obsérvese que la implicación inversa no es cierta en general.

Ejemplo

Ya sabemos que $\left\{(3,1,1)^t, (-1,2,1)^t, \frac{1}{2}(-1,-4,7)^t\right\}$ es un conjunto ortogonal. Comprobemos que, efectivamente, sus vectores son linealmente independientes; si escribimos la combinación lineal:

$$\alpha(3,1,1)^t + \beta(-1,2,1) + \gamma\frac{1}{2}(-1,-4,7)^t$$

y la igualamos a cero, es trivial ver que ha de ser $\alpha = \beta = \gamma = 0$. En cambio, el conjunto $\{(1,1,1)^t, (1,0,1)^t\}$ es linealmente independiente y, obviamente, no es ortogonal.

Un conjunto $\{u_1, \dots, u_p\}$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios.

Ejemplo

Podemos obtener un conjunto ortonormal a partir del conjunto ortogonal $\left\{(3,1,1)^t, (-1,2,1)^t, \frac{1}{2}(-1,-4,7)^t\right\}$ simplemente dividiendo cada vector entre su norma. Así, tendremos que

$$\left\{\frac{1}{\sqrt{11}}(3,1,1)^t, \frac{1}{\sqrt{6}}(-1,2,1)^t, \frac{1}{\sqrt{66}}(-1,-4,7)^t,\right\}$$

es un conjunto ortonormal, como se comprueba de manera trivial.

Obviamente, si S es un conjunto ortogonal de vectores no nulos de V, como es linealmente independiente, será una base del subespacio W generado por S. Puesto que S es tanto un conjunto ortogonal como una base de W la denominaremos **base ortogonal** de W. Si además los vectores de S son unitarios, diremos que S es una **base ortonormal** de W. Si hay $n = \dim(V)$ vectores en S, entonces $W = V$ y S es una base ortogonal (ortonormal) de V.

Veamos una consecuencia importante de esta idea:

Proposición

Sea $(w_1, \dots, w_p)$ una base ortogonal del subespacio W de V. Entonces, cada w de W tiene una representación única como combinación lineal de $w_1, \dots, w_p$. De hecho, si

$$w = c_1 w_1 + \cdots + c_p w_p,$$

entonces

$$c_j = \frac{\langle w, w_j\rangle}{\langle w_j, w_j\rangle}, \quad \text{para todo } j = 1, \dots, p.$$

Ejemplo

Los vectores $u_1 = (3, 1, 1)^t$, $u_2 = (-1, 2, 1)^t$ y $u_3 = \frac{1}{2}(-1, -4, 7)^t$ forman un conjunto ortogonal S; así, son linealmente independientes y, en consecuencia, forman una base ortogonal de $\mathbb{R}^3$. Vamos a expresar el vector $v = (6, 1, -8)^t$ como combinación lineal de estos vectores. Se tiene que

$$\begin{aligned} \langle v, u_1 \rangle &= 11\,, \quad \langle v, u_2 \rangle = -12\,, \quad \langle v, u_3 \rangle = -33\,, \\ \langle u_1, u_1 \rangle &= 11\,, \quad \langle u_2, u_2 \rangle = \ \ 6\,, \quad \langle u_3, u_3 \rangle = \frac{33}{2}\,. \end{aligned}$$

Entonces

$$v = \frac{11}{11} u_1 - \frac{12}{6} u_2 - \frac{33}{33/2} u_3 = u_1 - 2u_2 - 2u_3\,.$$

En otras palabras, las coordenadas de v con respecto a la base S son:

$$[v]_S = (1, -2, -2)^t\,.$$

El siguiente resultado relaciona el concepto de conjuntos ortonormales con las matrices ortogonales.

Teorema

Una matriz A $(m \times n)$ es ORTOGONAL (es decir, $A^t A = I$) si y sólo si tiene columnas ORTONORMALES.

Además, como veremos en el Tema 12, dada una matriz ortogonal cuadrada Q, ésta representa, respecto a alguna base ortogonal, un cierto tipo de transformaciones lineales con la propiedad de transformar bases ortonormales en bases ortonormales.

A continuación estudiamos un método que nos permite obtener una base ortogonal para cualquier subespacio W de un espacio vectorial dado V, a partir de cualquier base de W. Obviamente, si queremos obtener una base ortonormal, bastará con normalizar los vectores obtenidos con dicho método.

10.1.1. Procedimiento de Gram-Schmidt

El procedimiento de Gram-Schmidt es un algoritmo simple para obtener una base ortogonal (u ortonormal, normalizando los nuevos vectores) a partir de otra que no lo es. Para ilustrar cómo funciona, consideremos el siguiente ejemplo:

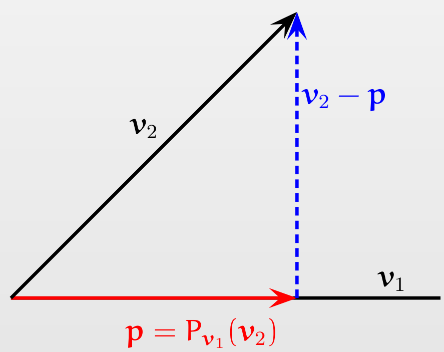
Ejemplo

Sea S el subespacio de $\mathbb{R}^2$ con el producto escalar usual y definido por $S = \mathtt{Gen}(v_1, v_2)$, con $v_1 = (3,6,0)^t$ y $v_2 = (1,2,2)^t$, linealmente independientes y no ortogonales (su producto interno no es nulo).

Si queremos construir una base ortogonal de S, podemos considerar

$$p = P_{v_1}(v_2) = \frac{15}{45}\,(3,6,0)^t\,,$$

de manera que $v_2 - p$ sea ortogonal a v_1 (ver la figura).



Obviamente, $w_2 = v_2 - p = (0,0,2)^t$ pertenece a S, puesto que es una combinación lineal de v_1 (con coeficiente $-\|P_{v_1}(v_2)\|$) y de v_2 (con coeficiente 1). Además es ortogonal a v_1. De esta manera, el par (v_1, w_2) constituye una base ortogonal de S.

Finalmente obtenemos una base ortonormal B_S de S dividiendo cada vector por su norma:

$$B_S = \left(\frac{1}{\sqrt{5}}\,(1,2,0)^t\,, (0,0,1)^t\right)\,.$$

Generalicemos esta idea.

Método de Gram-Schmidt.

Sea una base $(v_1, \dots, v_p)$ del subespacio S de V. Definimos:

$$\begin{aligned}
w_1 &= v_1 \,, \\
w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \,, \\
w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 \,, \\
&\vdots \\
w_p &= v_p - \frac{\langle v_p, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_p, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 - \cdots - \frac{\langle v_p, w_{p-1} \rangle}{\langle w_{p-1}, w_{p-1} \rangle} w_{p-1} \,.
\end{aligned}$$

Entonces $(w_1, \dots, w_p)$ es una base ortogonal de S. Además

$$\mathrm{Gen}(v_1, \dots, v_k) = \mathrm{Gen}(w_1, \dots, w_k), \quad 1 \leqslant k \leqslant p \,.$$

Ejemplo

Consideremos el espacio vectorial $\mathbb{P}_1$ sobre $\mathbb{R}$ con el producto interno:

$$\langle a_0 + a_1 x, b_0 + b_1 x \rangle = a_0 b_0 + 2 a_1 b_1 \,.$$

Consideremos la base $B_1 = (p_1, p_2) = (1 + x, 1 - x)$. Claramente, los vectores de esta base no son ortogonales, ya que $\langle p_1, p_2 \rangle = 1 \cdot 1 + 2 \cdot 1 \cdot (-1) = -1 \neq 0$. Vamos a utilizar el método de Gram-Schmidt para obtener una base ortogonal de $\mathbb{P}_1$. En primer lugar, consideramos el vector $q_1 = p_1$; el segundo vector de la base q_2 lo calculamos mediante:

$$q_2 = p_2 - \frac{\langle p_2, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 = p_2 + \frac{1}{3} q_1 = \frac{2}{3}(2 - x) \,.$$

Por tanto, la base $B_2 = (q_1, q_2)$ es ortogonal, como se comprueba fácilmente.

Si estamos interesados en utilizar una base ortonormal, dividimos cada vector de B_2

por su norma:

$$q_1' = \frac{1}{\sqrt{3}}(1+x), \quad q_2' = \frac{2}{\sqrt{3}}(2-x).$$

10.1.2. Factorización QR

Una aplicación adicional del método de Gram-Schmidt es la siguiente. Si $\{A_1, \ldots, A_n\}$ son las columnas de una matriz A de dimensión $m \times n$ con columnas linealmente independientes, podemos aplicar Gram-Schmidt a $A_1, \ldots, A_n$. Normalizando, es posible obtener una descomposición de A de la forma descrita en el siguiente teorema. Esta factorización es muy útil para simplificar diversos problemas, como la resolución de sistemas de ecuaciones o el cálculo de la inversa de una matriz.

Teorema: La factorización QR

Si A es una matriz $m \times n$ con columnas linealmente independientes, entonces A puede ser factorizada de la forma $A = QR$, donde Q es una matriz $m \times n$ cuyas columnas forman una base ortonormal de $\mathcal{C}(A)$ y R es una matriz $n \times n$ triangular superior e invertible con componentes positivas en su diagonal principal.

Ejemplo

Vamos a encontrar la factorización QR de

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Las columnas de A son linealmente independientes, por lo que podemos encontrar una base ortonormal para $\mathrm{Gen}(A_1, A_2, A_3) = \mathcal{C}(A)$. En primer lugar, hacemos $w_1 =$

$A_1 = (1,1,1,1)^t$. A continuación hacemos

$$w_2 = A_2 - \frac{\langle A_2, w_1\rangle}{\langle w_1, w_1\rangle}\, w_1 = \left((0,1,1,1)^t - \frac{3}{4}(1,1,1,1)^t\right) = \frac{1}{4}\,(-3,1,1,1)^t\,.$$

Para simplificar los cálculos, escalamos w_2 usando un factor de 4, es decir:

$$w_2' = (-3,1,1,1)^t\,.$$

Finalmente, obtenemos el tercer vector de la base ortogonal de la forma:

$$\begin{aligned} w_3 &= A_3 - \frac{\langle A_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle A_3, w_2'\rangle}{\langle w_2', w_2'\rangle} w_2' \\ &= (0,0,1,1)^t - \frac{2}{4}(1,1,1,1)^t - \frac{2}{12}(-3,1,1,1)^t \\ &= \frac{1}{3}(0,-2,1,1)^t\,. \end{aligned}$$

Es decir, una base ortogonal sencilla de $\mathcal{C}(A)$ sería:

$$\left((1,1,1,1)^t, (-3,1,1,1)^t, (0,-2,1,1)^t\right)\,.$$

Ahora podemos normalizar la base; los correspondientes vectores formarán las columnas de Q en la factorización:

$$Q = \begin{pmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{pmatrix}\,.$$

Puesto que las columnas de Q son ortonormales, tenemos que $Q^t\,Q = I_3$. Ahora necesitamos encontrar la matriz triangular superior R que verifica $A = Q\,R$. Si multiplicamos ambos miembros de esta expresión por Q^t resulta

$$Q^t\,A = Q^t\,Q\,R = I_3\,R = R\,.$$

Así:

$$
\begin{aligned}
R = Q^t A &= \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{pmatrix}.
\end{aligned}
$$