Bases Ortonormales
Método de ortogonalización de Gram-Schmidt

Algebra II FIUBA 2020

Recordemos algunas definiciones

Trabajaremos en un espacio vectorial V sobre $\mathbb{R}$ o $\mathbb{C}$ con **producto interno** $\langle \cdot, \cdot \rangle$ y **norma** definida a partir de ese producto interno :

$$\|v\| = \sqrt{\langle v, v \rangle}, \; v \in V$$

.

Una **base ortonormal** de un subespacio es una base cuyos elementos son ortogonales entre si y tienen norma 1 .

$$B = \{v_1, v_1, \cdots, v_m\}$$

$$\langle v_i, v_j \rangle = 0, \quad \forall i \neq j \qquad \|v_i\| = 1, \quad \forall i = 1, \cdots, m$$

Recordemos el
Teorema de Pitágoras : Si $v \perp w$, es decir $\langle v, w\rangle = 0$, entonces

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2$$

Entonces si v es combinación lineal de vectores ortogonales $v_1, v_2, \cdots, v_m$, la norma de $v = \sum_{i=1}^m a_i v_i$ resultará

$$\|v\|^2 = \|\sum_{i=1}^m a_i v_i\|^2 = \sum_{i=1}^m |a_i|^2 \|v_i\|^2$$

Más aún, si los vectores tienen norma 1
$\|v\|^2 = \|\sum_{i=1}^m a_i v_i\|^2 = \sum_{i=1}^m |a_i|^2$ y la norma de v puede calcularse sumando los cuadrados de sus coeficientes en la base ortonormal.
También conocemos una fórmula para calcular la proyección ortogonal $P_S(v)$ cuando conocemos una base ortogonal de S.

Resulta entonces muy conveniente tener una base ortonormal del espacio, o del subespacio donde vive un vector v cuya norma queremos calcular.

Ejemplo
Sea $V = \mathbb{R}^3$, $S = gen\{(1,1,1),(1,2,0)\}$, claramente S está descripto a partir de una base que no es ortogonal (ni ortonormal),

Podríamos hallar otra base de S que sea ortonormal ?

Empecemos por hallar una base ortogonal de S, después dividiremos los vectores por su norma y tendremos una base ortonormal.

Sean $v_1 = (1,1,1)$ y $v_2 = (1,2,0)$.

Construyamos una base ortogonal de S, $\{\tilde{v_1}, \tilde{v_2}\}$, y luego si los vectores no resultan de norma 1 podemos dividirlos por su norma:

Elijamos $\tilde{v_1} = (1,1,1)$ y hallemos un vector $\tilde{v_2}$ de manera que $\tilde{v_2} \perp \tilde{v_1}$ y $gen\{\tilde{v_1}, \tilde{v_2}\} = S$

Observemos que $\tilde{v_2} \in S$ de manera que $\tilde{v_2} = c_1\tilde{v_1} + c_2 v_2$.
Como queremos que $\tilde{v_1}$ y $\tilde{v_2}$ generen el mismo subespacio S, c_2 no puede ser nulo y como además $\tilde{v_2} \perp \tilde{v_1}$, resulta

$$\langle \tilde{v_2}, \tilde{v_1} \rangle = \langle c_1\tilde{v_1} + c_2 v_2, \tilde{v_1} \rangle$$

$$= c_1\langle \tilde{v_1}, \tilde{v_1} \rangle + c_2\langle v_2, \tilde{v_1} \rangle = 0$$

Tenemos entonces que

$$c_1\|\tilde{v_1}\|^2 + c_2\langle v_2, \tilde{v_1}\rangle = 0$$

Elijamos por ejemplo $c_2 = 1$ y despejemos $c_1 = -\frac{\langle v_2, \tilde{v_1}\rangle}{\|\tilde{v_1}\|^2} = -\frac{3}{3} = -1$ Entonces resulta

$$\tilde{v_1} = (1,1,1), \quad \tilde{v_2} = v_2 - \tilde{v_1} = (0,1,-1)$$

Para que esta nueva base resulte ortonormal dividimos los vectores por sus normas

$$\tilde{B} = \{\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{2}}(0,1,-1)\}$$

Podríamos utilizar este método para construir, a partir de una base finita , una base ortonormal con todas las ventajas que ya comentamos que tienen las BON.

Ese procedimiento se llama **Método de Ortogonalización de Gram Schmidt**.
A continuación lo presentamos.

Sea V un espacio con producto interno $\langle , \rangle$ y $B = \{v_1, v_2, \cdots, v_m\}$ un conjunto linealmente de V.

Vamos a construir de manera recursiva, un conjunto ortogonal $\tilde{B} = \{w_1, w_2, \cdots, w_m\}$ que satisfacerá

$$gen\{v_1, v_2, \cdots, v_k\} = gen\{w_1, w_2, \cdots, w_k\}, \quad k = 1, \cdots m$$

Finalmente dividiremos los vectores por sus norma para obtener una base ortonormal.

Elijamos $w_1 = v_1$.
Ciertamente se satisface $gen\{v_1\} = gen\{w_1\}$

Para elegir w_2 tengamos en cuenta que debe ser

$$gen\{v_1, v_2\} = gen\{w_1, w_2\} \tag{1}$$

por lo tanto w_2 será de la forma $w_2 = c_1 v_1 + c_2 v_2$.
Como además debe ser ortogonal a w_1, obtenemos

$$\langle w_2, w_1 \rangle = \langle c_1 v_1 + c_2 v_2, v_1 \rangle = 0$$

Notemos que no puede ser $c_2 = 0$ ya que en ese caso no valdría (1), entonces si elegimos $c_2 = 1$ podemos despejar $c_1 = -\frac{\langle v_2, v_1 \rangle}{\| v_1 \| \|^2}$ y obtener

$$w_2 = v_2 - \frac{\langle v_2, v_1 \rangle}{\| v_1 \| \|^2} v_1 = v_2 - \frac{\langle v_2, v_1 \rangle}{\| v_1 \| \|^2} w_1$$

Este vector w_2 satisface todo lo pedido.

Con este mismo procedimiento podemos construir $w_3, \cdots, w_k$. Supogamos que lo hemos hecho .

Es decir que tenemos vectores ortogonales $w_1, \cdots, w_k$ tales que

$$gen\{v_1, v_2, \cdots, v_k\} = gen\{w_1, w_2, \cdots, w_k\}$$

Cómo construiríamos w_{k+1}?
w_{k+1} debe satisfacer

$$\langle w_{k+1}, w_j \rangle = 0, \quad j = 1, \cdots, k$$

$$gen\{v_1, v_2, \cdots, v_{k+1}\} = gen\{w_1, w_2, \cdots, w_{k+1}\}$$

A partir de la construcción que ya hicimos para w_2 proponemos

$$w_{k+1} = v_{k+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\| w_j \| \|^2} w_j \qquad (2)$$

Los productos internos con los otros elementos de la base resultan para $i \leq k$

$$\langle w_{k+1}, w_i \rangle = \langle v_{k+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\| w_j \| \|^2} w_j, w_i \rangle =$$

$$\langle v_{k+1}, w_i \rangle - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\| w_j \| \|^2} \langle w_j, w_i \rangle$$

Como $\langle w_i, w_j \rangle = 0$ cuando $i \neq j$ resulta

$$\langle w_{k+1}, w_i \rangle = \langle v_{k+1}, w_i \rangle - \frac{\langle v_{k+1}, w_j \rangle}{\|w_j|\|^2} \langle w_j, w_i \rangle = 0$$

Ya hemos visto qu se satisface la condición de ortogonalidad. Nos faltaría ver que

$$gen\{v_1, v_2, \cdots, v_k, v_{k+1}\} = gen\{w_1, w_2, \cdots, w_k, w_{k+1}\}$$

Esta condición se satisface ya que,
$gen\{v_1, v_2, \cdots, v_k\} = gen\{w_1, w_2, \cdots, w_k\}$ y por la construcción que realizamos, $w_{k+1} = v_{k+1} - \sum c_j w_j$.

$$\{w_1, w_2, \cdots, w_m\}$$

es el conjunto ortogonal que buscábamos.

Para obtener un conjunto ortonormal que genere el mismo espacio, basta simplemente con dividir cada vector w_j por su norma y obtener $z_j = \frac{w_j}{\|w_j\|}$.

El conjunto $\{z_1, z_2, \cdots, z_m\}$ es un conjunto ortonormal de vectores que satisface

$$gen\{z_1, z_2, \cdots, z_m\} = gen\{v_1, v_2, \cdots, v_m\}$$

.

Ejemplo

Dada $B = \{(1,0,1),(-1,1,-1),(2,2,0)\}$ construyamos a partir de ella una base ortonomal de $\mathbb{R}^3$ utilizando el método de ortogonalización de Gram-Schmidt.
Consideraremos el producto interno canónico.
Utilizando la notación anterior

$$v_1 = (1,0,1),\ v_2 = (-3,1,-1),\ v_3 = (2,2,1)$$

Entonces

$$w_1 = v_1 = (1, 0, 1)$$

Para hallar w_2 planteamos $w_2 = c_2 v_2 + c_1 w_1$, elegimos $c_2 = 1$ y pedimos $w_2 \perp w_1$ de donde resulta

$$w_2 = v_2 - \frac{\langle v2, w_1 \rangle}{\|w_1\|^2} w_1 = (-3, 1, -1) - \frac{-4}{2}(1, 0, 1)$$

Entonces $w_2 = v_2 + 2w_1 = (-3, 1, -1) + 2(1, 0, 1) = (-1, 1, 1)$

$$w_2 = (-1, 1, 1)$$

Para calcular w_3 planteamos

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2 = (2,2,1) - \frac{3}{2}(1,0,1) - \frac{1}{3}(-1,1,1)$$

de donde resulta

$$w_3 = (\tfrac{5}{6}, \tfrac{5}{3}, -\tfrac{5}{6})$$

La base obtenida

$$\{(1,0,1),(-1,1,1),(\frac{5}{6},\frac{5}{3},-\frac{5}{6})\}$$

es ortogonal pero no es ortonormal.

Si dividimos cada vector por su norma obtenemos la base ortonormal

$$\{\frac{(1,0,1)}{\sqrt{2}},\frac{(-1,1,1)}{\sqrt{3}},\frac{(\frac{5}{6},\frac{5}{3},-\frac{5}{6})}{\sqrt{\frac{150}{36}}}\}$$

Observación

El método de ortogonalización se puede aplicar a conjuntos de m vectores linealmente independientes en un espacio de dimensión n $(n \geq m)$.

No es necesario que el conjunto inicial sea una base del espacio, como sucedió en el primer ejemplo que desarrollamos.

Escribamos una vez más la relación entre v_j y w_i a partir de la construcción que propone el método de ortogonalización de Gram Schmidt.

$$w_1 = v_1$$

$$w_2 = v_2 - a_{1,2}w_1$$

$$w_3 = v_3 - a_{1,3}w_1 - a_{2,3}w_2$$

$$\vdots$$

$$w_k = v_k - a_{1,k}w_1 - a_{2,k}w_2 \cdots - a_{j-1,k}w_{j-1}$$

$$\vdots$$

$$w_m = v_m - a_{1,m}w_1 - a_{2,m}w_2 - \cdots - a_{m-1,m}w_{m-1}$$

Si en lugar de w_j despejamos v_j en cada ecuación obtnemos

$$v_1 = w_1$$

$$v_2 = a_{1,2}w_1 + w_2$$

$$v_3 = a_{1,3}w_1 + a_{2,3}w_2 + w_3$$

$$\vdots$$

$$v_k = a_{1,k}w_1 - a_{2,k}w_2 \cdots - a_{j-1,k}w_{j-1} + w_k$$

$$\vdots$$

$$v_m = a_{1,m}w_1 + a_{2,m}w_2 + \cdots + a_{m-1,m}w_{m-1} + w_m$$

Esta relación puede escribirse matricialmente

Sea $A \in \mathbb{R}^{n \times m}$ la matriz cuyas columnas son los vectores v_i

$$A = [v_1 \, v_2 \, \cdots v_m] = [w_1 \, w_2 \, \cdots w_m] \begin{bmatrix} 1 & a_{1,2} & a_{1,3} & \cdots & a_{1,m} \\ 0 & 1 & a_{2,3} & \cdots & a_{2,m} \\ 0 & 0 & 1 & \cdots & a_{3,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Si llamamos $S \in \mathbb{R}^{m\times m}$ a la matriz triangular superior

$$S = \begin{bmatrix} 1 & a_{1,2} & a_{1,3} & \cdots & a_{1,m} \\ 0 & 1 & a_{2,3} & \cdots & a_{2,m} \\ 0 & 0 & 1 & \cdots & a_{3,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

y $P \in \mathbb{R}^{n\times m}$ a la matriz cuyas columnas son los vectores w_j (que son ortogonales), hemos obtenido una descomposición de A como $A = PS$.

Si dividimos cada columna de P por su norma obtenemos una matriz cuyas columnas son ortogonales y unitarias

$$Q = [q_1\, q_2\, \cdots q_n] = [\frac{w_1}{\|w_1\|}, \frac{w_2}{\|w_2\|}, \cdots, \frac{w_m}{\|w_m\|}]$$

Para mantener la descomposición anterior debemos alterar S multiplicando la fila j por $\|w_j\|$

$$R = \begin{bmatrix} \|w_1\| & a_{1,2}\|w_1\| & a_{1,3\|w_1\|} & \cdots & a_{1,m}\|w_1\| \\ 0 & \|w_2\| & a_{2,3}\|w_2\| & \cdots & a_{2,m}\|w_2\| \\ 0 & 0 & \|w_3\| & \cdots & a_{3,m}\|w_3\| \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & 1 & \end{bmatrix}$$

La descomposición de A resulta entonces

$$A = QR$$

donde Q es una matriz de columnas ortogonales y unitarias $Q^T Q = Id$ y R es una matriz triangular superior con elementos positivos en su diagonal.

Observación
Así como sucediera con el método de otogonalización, la descomposición QR, se puede aplicar a conjuntos de m vectores linealmente independientes en un espacio de dimensión n ($n \geq m$), es decir a matrices $A \in \mathbb{R}^{n \times m}$ de rango m. No es necesario que el conjunto inicial sea una base del espacio.

Resultado

Si $A \in \mathbb{R}^{n\times m}$ con $rg(A) = m$ (es decir que sus columnas son linealmente independientes) exiten matrices Q que satisface $Q^T Q = Id$ y R triangular superior con elementos positivos en su diagonal, que satisfacen

$$A = QR$$

Observación

Si $A \in \mathbb{C}^{n\times m}$ la matriz Q de la desomposición es unitaria , decir $\overline{Q^T} Q = Id$.

Por qué sería útil o conveniente poder expresar una matriz de esta forma?

Supongamos que queremos encontrar una solución de $Ax = b$ por cuadrados mínimos

$$A^T Ax = A^T b \Rightarrow R^T Q^T QRx = A^T b \Rightarrow$$

$$\Rightarrow R^T Rx = R^T Q^T b \Rightarrow Rx = Q^T b$$

esta última ecuación es fácil de resolver ya que R es una matriz triangular.

Ejemplo

Retomemos el ejemplo anterior y hallemos la descomposición QR de la matriz cuyas columnas son los vectores v_1, v_2, v_3.
Habiendo hallado, por medio de la ortogonalización de Gram Schmidt

$$w_1 = v_1$$

$$w_2 = v_2 + 2w_1$$

$$w_3 = v_3 - \frac{3}{2}w_1 - \frac{1}{3}w_2$$

y despejando los vectores v_1, v_2 y v_3, resulta

$$v_1 = v_1$$

$$v_2 = w_2 - 2w_1$$

$$v_3 = w_3 + \frac{3}{2}w_1 + \frac{1}{3}w_2$$

Entonces

$$A = [v_1\, v_2\, \cdots v_n] = PS = \begin{bmatrix} 1 & -1 & \frac{5}{6} \\ 0 & 1 & \frac{5}{3} \\ 1 & 1 & -\frac{5}{6} \end{bmatrix} \begin{bmatrix} 1 & -2 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix}$$

Para obtener la descomposición QR de A, debemos dividir por la norma de w_i la i.ésima columna de P y, consistentemente, multiplicar la i-ésima fila d S por ese mismo valor.

Como $\|w_1\| = \sqrt{2}, \quad \|w_2\| = \sqrt{3}, \quad \|w_3\| = \sqrt{\frac{150}{36}} = \frac{5\sqrt{6}}{6}$
obtenemos

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{6}}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix}$$

$$R = \begin{bmatrix} \sqrt{2} & -2\sqrt{2} & \frac{3}{2}\sqrt{2} \\ 0 & \sqrt{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & 5\frac{\sqrt{6}}{6} \end{bmatrix}$$

y la descomposición resulta

$$\begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{6}}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & -2\sqrt{2} & \frac{3}{2}\sqrt{2} \\ 0 & \sqrt{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & 5\frac{\sqrt{6}}{6} \end{bmatrix}$$