Ortogonalización de Gram–Schmidt

Objetivos. Estudiar el proceso de ortogonalización de Gram–Schmidt que permite construir de una lista arbitraria de vectores $a_1, \ldots, a_m$ una lista ortogonal $b_1, \ldots, b_m$ que genere al mismo subespacio.

Requisitos. Listas ortogonales de vectores, listas ortonormales de vectores, proyección ortogonal de un vector sobre el subespacio generado por vectores ortogonales, matriz de Gram.

En esta sección suponemos que V es un espacio vectorial complejo o real con un producto interno. El el caso complejo suponemos que el producto interno es lineal con respecto al segundo argumento.

1. Proyección ortogonal de un vector sobre el subespacio generado por una lista ortogonal (repaso). Sean V un espacio vectorial real o complejo con producto interno, $b_1, \ldots, b_j$ algunos vectores ortogonales no nulos y $v \in V$. Definimos los vectores $u, w \in V$ de la siguiente manera:

$$u = \sum_{k=1}^{m} \frac{\langle b_k, v \rangle}{\langle b_k, b_k \rangle} \, b_k, \qquad w = v - u. \tag{1}$$

Entonces $w \perp \ell(b_1, \ldots, b_j)$.

2. Proceso de ortogonalización de Gram–Schmidt. Sea V un espacio vectorial real o complejo con producto interno y sean $a_1, \ldots, a_m \in V$. Queremos construir vectores ortogonales $b_1, \ldots, b_m \in V$ de tal manera que para todo $j \in \{1, \ldots, m\}$

$$\ell(b_1, \ldots, b_j) = \ell(a_1, \ldots, a_j).$$

Idea del proceso de ortogonalización de Gram–Schmidt: en el j-ésimo paso definir el vector b_j como a_j menos la proyección ortogonal del vector a_j al subespacio generado por los vectores $b_1, \ldots, b_{j-1}$.

En el j-ésimo paso suponemos que los vectores $b_1, \ldots, b_{j-1}$ ya están construidos y son ortogonales entre si. Buscamos b_j de la forma

$$b_j = a_j - \sum_{k=1}^{j-1} \lambda_{j,k} b_k. \tag{2}$$

Para memorizar los índices del coeficiente $\lambda_{j,k}$ puede notar que este coeficiente sirve para "corregir" el vector a_j usando el vector b_k.

Para calcular el coeficiente $\lambda_{j,q}$ multipliquemos la igualdad (2) por b_q en el sentido del producto interno:

$$\langle b_q, b_j\rangle = \left\langle b_q, a_j - \sum_{k=1}^{j-1} \lambda_{j,k} b_k \right\rangle = \langle b_q, a_j\rangle - \sum_{k=1}^{j-1} \delta_{q,k}\lambda_{j,k}\|b_k\|^2 = \langle b_q, a_j\rangle - \lambda_{j,q}\|b_q\|^2.$$

Queremos que $\langle b_q, b_j\rangle$ sea igual a 0. Si $b_q \ne \mathbf{0}$, entonces $\lambda_{j,q}$ debe ser igual a

$$\lambda_{j,q} = \frac{\langle b_q, a_j\rangle}{\|b_q\|^2} = \frac{\langle b_q, a_j\rangle}{\langle b_q, b_q\rangle}.$$

Si $b_q = \mathbf{0}$, entonces el sumando $\lambda_{j,q} b_i$ no depende de $\lambda_{j,q}$, y $\lambda_{j,q}$ se puede elegir de manera arbitraria. En este caso por simplicidad ponemos $\lambda_{j,q} = 0$.

Así obtenemos las fórmulas principales:

$$\boxed{b_j := a_j - \sum_{k=1}^{j-1} \lambda_{j,k} b_k,} \qquad \text{donde} \qquad \boxed{\lambda_{j,k} := \begin{cases} \dfrac{\langle b_k, a_j\rangle}{\|b_k\|^2}, & b_k \ne \mathbf{0}; \\ 0, & b_k = \mathbf{0}. \end{cases}} \tag{3}$$

3. Observación. Es importante que el vector b_j se construye como una combinación lineal de los vectores $b_1, \ldots, b_{j-1}, a_j$, con el uso de los *vectores nuevos* $b_1, \ldots, b_{j-1}$. Los vectores $b_1, \ldots, b_{j-1}$ ya son ortogonales entre si, por eso las fórmulas para los coeficientes $\lambda_{j,k}$ son tan simples. Sería muy incómodo construir b_j como una combinación lineal de los vectores originales $a_1, \ldots, a_{j-1}, a_j$.

4. Ejemplo. Aplicar la ortogonalización de Gram–Schmidt a la lista de vectores a_1, a_2, a_3:

$$a_1 = \begin{bmatrix} 4 \\ -2 \\ -1 \\ 2 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} -6 \\ 3 \\ 4 \\ -8 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} 5 \\ -5 \\ -3 \\ -4 \end{bmatrix}.$$

Usando la matriz de Gram compruebe que la lista de vectores b_1, b_2, b_3 que se obtiene al final es ortogonal.

Solución. 1. Ponemos $b_1 = a_1$. Calculamos la norma de b_1:

$$\|b_1\|^2 = 16 + 4 + 1 + 4 = 25, \qquad \|b_1\| = 5.$$

2. Construimos el vector b_2.

$$\lambda_{2,1} = \frac{\langle b_1, a_2\rangle}{\|b_1\|^2} = \frac{-24 - 6 - 4 - 16}{25} = -2.$$

De aquí

$$b_2 = a_2 - \lambda_{2,1} b_1 = a_2 + 2b_1 = \begin{bmatrix} -6 \\ 3 \\ 4 \\ -8 \end{bmatrix} + \begin{bmatrix} 8 \\ -4 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 2 \\ -4 \end{bmatrix}.$$

Calculamos la norma de b_2:

$$\|b_2\|^2 = 4 + 1 + 4 + 16 = 25, \qquad \|b_2\| = 5.$$

3. Construimos el vector b_3.

$$\lambda_{3,1} = \frac{\langle b_1, a_3 \rangle}{\|b_1\|^2} = \frac{20 + 10 + 3 - 8}{25} = 1, \qquad \lambda_{3,2} = \frac{\langle b_2, a_3 \rangle}{\|b_2\|^2} = \frac{10 + 5 - 6 + 16}{25} = 1.$$

De aquí

$$b_3 = a_3 - \lambda_{3,1} b_1 - \lambda_{3,2} b_2 = a_3 - b_1 - b_2 = \begin{bmatrix} 5 \\ -5 \\ -3 \\ -4 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \\ 1 \\ -2 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -4 \\ -2 \end{bmatrix}.$$

Calculamos la norma de b_3:

$$\|b_3\|^2 = 1 + 4 + 16 + 4 = 25, \qquad \|b_3\| = 5.$$

Para comprobar que los vectores b_1, b_2, b_3 son ortogonales calculamos su matriz de Gram:

$$G(b_1, b_2, b_3) = \begin{bmatrix} 4 & -2 & -1 & 2 \\ 2 & -1 & 2 & -4 \\ -1 & -2 & -4 & -2 \end{bmatrix} \begin{bmatrix} 4 & 2 & -1 \\ -2 & -1 & -2 \\ -1 & 2 & -4 \\ 2 & -4 & -2 \end{bmatrix} = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 25 \end{bmatrix}.$$

Podemos normalizar los vectores b_1, b_2, b_3 (dividirlos entre sus normas) y obtener una lista ortonormal:

$$c_1 = \begin{bmatrix} 4/5 \\ -2/5 \\ -1/5 \\ 2/5 \end{bmatrix}, \qquad c_2 = \begin{bmatrix} 2/5 \\ -1/5 \\ 2/5 \\ -4/5 \end{bmatrix}, \qquad c_3 = \begin{bmatrix} -1/5 \\ -2/5 \\ -4/5 \\ -2/5 \end{bmatrix}. \qquad \square$$

5. Ejemplo. Ortogonalizar la siguiente lista de vectores en $\mathbb{R}^4$:

$$a_1 = \begin{bmatrix} 5 \\ 1 \\ 1 \\ -3 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} 9 \\ 3 \\ 3 \\ -7 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} 7 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \qquad a_4 = \begin{bmatrix} -5 \\ 5 \\ -1 \\ 5 \end{bmatrix}.$$

6. Criterio de la contención de subespacios en términos de sus generadores, repaso. Sean $a_1, \ldots, a_j \in V$ algunos vectores y sea S un subespacios de V. Entonces:

$$a_1, \ldots, a_j \in S \qquad \Longleftrightarrow \qquad \ell(a_1, \ldots, a_j) \subseteq S.$$

7. Teorema (conservación de los subespacios en el proceso de ortogonalización de Gram–Schmidt). Sea V un espacio vectorial real o complejo con producto interno y sean $a_1, \ldots, a_m \in V$. Denotemos por $b_1, \ldots, b_m$ a los vectores obtenidos de $a_1, \ldots, a_m$ al aplicar el método de ortogonalización de Gram-Schmidt:

$$b_j = a_j - \sum_{k=1}^{j-1} \lambda_{j,k} b_k, \tag{4}$$

$$\lambda_{j,k} = \begin{cases} \dfrac{\langle b_k, a_j \rangle}{\|b_k\|^2}, & b_k \neq \mathbf{0}; \\ 0, & b_k = \mathbf{0}. \end{cases} \tag{5}$$

Entonces para todo $j \in \{1, \ldots, m\}$ los vectores $a_1, \ldots, a_j$ generan al mismo subespacio que los vectores $b_1, \ldots, b_j$:

$$\ell(a_1, \ldots, a_j) = \ell(b_1, \ldots, b_j).$$

Demostración. 1. De la fórmula (4) podemos expresar a_j como una combinación lineal de $b_1, \ldots, b_j$:

$$a_j = \sum_{k=1}^{j-1} \lambda_{j,k} b_k + b_j.$$

Esto implica que $a_1, \ldots, a_j \in \ell(b_1, \ldots, b_j)$.

2. Demostremos por inducción sobre j la siguiente afirmación $P(j)$:

$$b_1, \ldots, b_j \in \ell(a_1, \ldots, a_j).$$

El caso $j = 1$ es trivial: $b_1 = a_1 \in \ell(a_1)$. Supongamos que la afirmación $P(j-1)$ es válida, esto es, $b_1, \ldots, b_{j-1} \in \ell(a_1, \ldots, a_{j-1})$. Entonces cada sumando escrito en el lado derecho de la fórmula (4) pertenece al subespacio $\ell(a_1, \ldots, a_j)$ y, por lo tanto, $b_j \in \ell(a_1, \ldots, a_j)$. Acabamos de demostrar que $P(j-1)$ implica $P(j)$.

3. Del resultado de la parte 1 de la demostración se sigue que $\ell(a_1, \ldots, a_j) \subseteq \ell(b_1, \ldots, b_j)$, y del resultado de la parte 2 se sigue que $\ell(b_1, \ldots, b_j) \subseteq \ell(a_1, \ldots, a_j)$. $\square$

8. Corolario (ortogonalización de Gram–Schmidt y dependencias lineales). Sean $a_1, \ldots, a_m$ una lista de vectores en V y $b_1, \ldots, b_m$ la lista obtenida de $a_1, \ldots, a_m$ al aplicar la ortogonalización de Gram-Schmidt. Entonces para todo $j \in \{1, \ldots, m\}$ las siguientes condiciones son equivalentes:

a) $a_j \in \ell(a_1, \ldots, a_{j-1})$.

b) $b_j = \mathbf{0}$.

Demostración. (a)⇒(b). Supongamos que $a_j \in \ell(a_1, \ldots, a_{j-1})$. Entonces

$$b_j \in \ell(b_1, \ldots, b_j) = \ell(a_1, \ldots, a_j) = \ell(a_1, \ldots, a_{j-1}) = \ell(b_1, \ldots, b_{j-1}),$$

lo que significa que b_j es una combinación lineal de $b_1, \ldots, b_{j-1}$. Como $b_j \perp \{b_1, \ldots, b_{j-1}\}$, los coeficientes de esta combinación lineal son nulos y $b_j = \mathbf{0}$.

(b)⇒(a). Supongamos que $b_j = \mathbf{0}$. Entonces

$$\begin{aligned} a_j \in \ell(a_1, \ldots, a_j) &= \ell(b_1, \ldots, b_j) = \ell(b_1, \ldots, b_{j-1}, \mathbf{0}) \\ &= \ell(b_1, \ldots, b_{j-1}) = \ell(a_1, \ldots, a_{j-1}). \end{aligned}$$ □

9. Corolario. En las notaciones del teorema, las siguientes dos condiciones son equivalentes:

(a) $a_1, \ldots, a_m$ son linealmente independientes.

(b) todos los vectores $b_1, \ldots, b_m$ son no nulos.

10. Observación. En muchos libros el proceso de ortogonalización de Gram–Schmidt se estudia solamente en el caso si los vectores originales $a_1, \ldots, a_m$ son linealmente independientes. En este caso los vectores $b_1, \ldots, b_m$ son no nulos, y la fórmula para los coeficientes $\lambda_{j,k}$ se simplifica (se quita el caso $b_k = \mathbf{0}$). El autor de estos apuntes (Egor Maximenko) agradece al profesor Vladimir Borísovich Dybin por la explicación del caso general.

11. Ejercicio. Ortogonalizar la siguiente lista de vectores en $\mathbb{R}^4$:

$$a_1 = \begin{bmatrix} 2 \\ -4 \\ 5 \\ 2 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} -6 \\ 5 \\ -1 \\ -6 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} -10 \\ 13 \\ -4 \\ -3 \end{bmatrix}.$$

12. Ejercicio. Usando el proceso de Gram–Schmidt ortogonalice la siguiente lista de vectores en $\mathbb{R}^4$:

$$a_1 = \begin{bmatrix} 4 \\ -2 \\ -1 \\ 2 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} -6 \\ 3 \\ 4 \\ -8 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} 5 \\ -5 \\ -3 \\ -4 \end{bmatrix}.$$

13. Ejercicio. Aplique la ortogonalización de Gram–Schmidt a la siguiente lista de vectores en $\mathbb{R}^4$:

$$a_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} 5 \\ 3 \\ -3 \\ -3 \end{bmatrix}, \qquad a_3 = \begin{bmatrix} 1 \\ -1 \\ -7 \\ 1 \end{bmatrix}, \qquad a_4 = \begin{bmatrix} -2 \\ 0 \\ 12 \\ 6 \end{bmatrix}.$$

14. Tarea adicional. Consideramos el espacio de los polinomios $\mathcal{P}(\mathbb{R})$ con el producto interno

$$\langle f, g \rangle := \int_{-1}^{1} f(x)\, g(x)\, dx.$$

Aplique el proceso de Gram–Schmidt a los monomios $e_0(x) = 1$, $e_1(x) = x$, $e_2(x) = x^2$, $e_3(x) = x^3$.