PRACTICA 9

Método de ortogonalización de Gram–Schmidt. Aplicaciones.

Método de Gram-Schmidt. El *método de ortogonalización de Gram–Schmidt* permite construir conjuntos ortogonales en cualquier espacio con producto interno, de dimensión finita o infinita, a partir de una sucesión de vectores no nulos como sigue: Sea V un espacio con producto interno y sea $v_1, v_2, \ldots$ una sucesión (finita o infinita) de vectores no nulos cualesquiera de V. Entonces la sucesión de vectores $w_1, w_2, \ldots$ de V generados a partir de las relaciones

$$\begin{aligned} w_1 &= v_1 \\ w_j &= v_j - \sum_{i=1}^{j-1} \frac{\langle w_i | v_j \rangle}{\|w_i\|^2} w_i \qquad j = 2, 3, \ldots \end{aligned}$$

tiene las siguientes propiedades para cada $j = 1, 2, \ldots$

i) El elemento w_j es ortogonal a todo vector del subespacio generado por los vectores precedentes $w_1, w_2, \ldots, w_{j-1}$,

$$\langle w_i | w_j \rangle = 0, \qquad \text{para } i = 1, 2, \ldots, j-1.$$

ii) El subespacio generado por $w_1, \ldots, w_j$ es el mismo que el subespacio generado por $v_1, \ldots, v_j$,

$$\overline{\{w_1, \ldots, w_j\}} = \overline{\{v_1, \ldots, v_j\}}.$$

iii) La sucesión de vectores es única salvo factores escalares. Esto es, si $w'_1, w'_2, \ldots$ es otra sucesión que satisfacen i y ii, entonces para cada j existe un escalar c_j tal que $w'_j = c_j w_j$.

☞ Según ii, cada vector w_j es combinación lineal de de $v_1, v_2, \ldots, v_j$.

☞ Si tenemos n vectores linealmente independientes $v_1, \ldots, v_n$, entonces el proceso de ortogonalización conduce a una *base ortogonal* del subspacio S generado por $v_1, \ldots v_n$. En particular, si V es de dimensión *finita* y $\{v_1, \ldots, v_n\}$ es una base de V, entonces el proceso de ortogonalización determina una base ortogonal de V. Una base *ortonormal* resulta de normalizar cada w_j, esto es, de tomar cada vector w_j multiplicado por el recíproco de su norma:

$$w'_j = \frac{w_j}{\|w_j\|}$$

☞ En los cálculos a mano suele resultar útil suprimir los denominadores en cada nuevo w_j generado multiplicando éste por un escalar apropiado, lo cual, según la propiedad iii no afecta la ortogonalidad.

☞ Sea V un espacio de dimensión finita y sea S un subespacio siendo $\{v_1, \ldots, v_k\}$ una base del mismo. Extendiendo ésta a una base $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ de V y aplicando el método de Gram–Schmidt obtenemos una base ortonormal $\{w_1, \ldots, w_n\}$ de V con $\{w_1, \ldots, w_k\}$ base ortonormal de S y $\{w_{k+1}, \ldots, w_n\}$ base ortonormal de $S^\perp$, el complemento ortogonal de S.

Una forma alternativa de aplicar el método de manera de obtener un conjunto *ortonormal* en forma directa (es decir, sin normalizar *a posteriori*) procede como sigue

$$\begin{aligned} w_1 &= \frac{v_1}{\|v_1\|} \\ w_j &= \frac{v_j - \sum_{i=1}^{j-1} \langle w_i | v_j \rangle w_i}{\left\| v_j - \sum_{i=1}^{j-1} \langle w_i | v_j \rangle w_i \right\|} \qquad j = 2, 3, \ldots \end{aligned}$$

Una implementación algorítmica de este procedimiento puede plantearse como sigue.

Algoritmo de ortonormalización de Gram-Schmidt

Dados $v_1, v_2, \ldots, v_n$
Para $j = 1, 2, \ldots, n$
    Tomar $w_j = v_j$
    Para $i = 1, 2 \ldots, j-1$
        Calcular $r_{ij} = \langle w_i | v_j \rangle$
        Tomar $w_j = w_j - r_{ij} w_i$
    Calcular $r_{jj} = \|w_j\|$
    Tomar $w_j = \dfrac{w_j}{r_{jj}}$

Ejercicio 1. Encontrar una base ortonormal para el subespacio S de $\mathbb{R}^4$ (con el producto interno canónico) generado por el conjunto de vectores $\{v_1, v_2, v_3\}$, donde

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

Determinar, también, una base para el complemento ortogonal $S^\perp$.

Ejercicio 2. Explicar que sucede si el procedimiento de ortogonalización es aplicado a

a) un conjunto de vectores ortonormales,

b) un conjunto de vectores linealmente dependientes.

Ejercicio 3. Sea V el espacio vectorial de las funciones reales continuas sobre el intervalo $[-1, 1]$ con el producto interno

$$\langle f | g \rangle = \int_{-1}^{1} f(t) g(t) \mathrm{d}t$$

Sea S el subespacio generado por los tres polinomios $q_0 = 1, q_1 = t, q_2 = t^2$.

a) Usando el procedimiento de ortogonalización determine una base ortogonal $\{p_0, p_1, p_2\}$ de S.

b) Multiplicando por un escalar apropiado, determine a partir de los polinomios del punto anterior, los polinomios que son ortogonales y cumplen que $p_k(1) = 1$ para $k = 0, 1, 2$. Los polinomios así obtenidos son los tres primeros *polinomios de Legendre*, los cuales juegan un papel relevante en la resolución de ciertas ecuaciones diferenciales.

Ejercicio 4. Sea S el subespacio de $\mathbb{R}^4$ (con el producto interno canónico) generado por $v_1 = (1, 1, 1, 1)$ y $v_2 = (1, 2, 3, 2)$. Determinar la proyección ortogonal de $v = (1, 3, 5, 7)$ sobre S.

Mejor aproximación. La *proyección ortogonal* aparece como la solución de un importante problema de aproximación que pasamos a detallar. Supongamos que S es un subespacio de un espacio con producto interno V y sea v un vector arbitrario en V. El problema consiste en hallar un vector s en S cuya distancia a v, dada por la norma $\|v - s\|$ sea lo más pequeña posible. El problema puede precisarse estableciendo la siguiente definición

Definición. Sea S un subespacio de un espacio con producto interno V y sea v un vector de V. Una *mejor aproximación* a v por vectores de S es un vector s de S tal que

$$\|v - s\| \leq \|v - w\|$$

para todo vector w en S.

Ejercicio 5. Demostrar que si S es un subespacio de dimensión *finita*, entonces la proyección ortogonal de v sobre S es la mejor aproximación a v por vectores de S.

Ejercicio 6. Sea V el espacio de las funciones reales continuas sobre el intervalo $[-1, 1]$ con el producto interno $\langle f|g\rangle = \int_{-1}^{1} f(t)g(t)\mathrm{d}t$ y sea S el subespacio de V de los polinomios de grado ≤ 2. Encontrar el polinomio de p de grado ≤ 2 que constituye la mejor aproximación a $f(t) = t^5$, en el sentido de que

$$\int_{-1}^{1} [f(t) - p(t)]^2 \mathrm{d}t$$

es mínima.

Estabilidad numérica del método de Gram-Schmidt. El procedimiento de Gram-Schmidt, si bien es una herramienta teórica poderosa, no constituye un buen algoritmo *numérico*, puesto los cálculos involucrados en la forma presentada pueden producir pérdida de precisión cuando se emplea aritmética de punto flotante. Un algoritmo matemáticamente equivalente pero numéricamente más estable es dado a continuación:

Algoritmo de ortonormalización de Gram-Schmidt modificado

Dados $v_1, v_2, \ldots, v_n$
Para $j = 1, 2, \ldots, n$
    Tomar $w_j = v_j$
    Para $i = 1, 2 \ldots, j - 1$
        Calcular $r_{ij} = \langle w_i | w_j\rangle$
        Tomar $w_j = w_j - r_{ij} w_i$
    Calcular $r_{jj} = \|w_j\|$
    Tomar $w_j = \dfrac{w_j}{r_{jj}}$

Ejercicio 7. Utilizar aritmética de punto flotante de tres dígitos para obtener una base ortonormal del subespacio de $\mathbb{R}^3$ generado por los vectores

$$v_1 = (1, 10^{-3}, 10^{-3}), \quad v_2 = (1, 10^{-3}, 0), \quad v_3 = (1, 0, 10^{-3})$$

utilizando

a) el método de ortonormalización no modificado,

b) el método de ortonormalización modificado.

Comparar los resultados obtenidos.

Factorización QR. El procedimiento de ortogonalización de Gram-Schmidt permite construir una factorización matricial llamada *factorización QR* como se indica a continuación. Sea $A = (\mathbf{a}_1|\mathbf{a}_2|\cdots|\mathbf{a}_n)$ una matriz $m \times n$ sobre el cuerpo K, real o complejo, cuyos n vectores columnas $\mathbf{a_k}$ son un conjunto linealmente independiente de K^m. La aplicación del procedimiento de ortonormalización a las columnas de A conduce a una base ortonormal $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ del espacio columna de A (respecto al producto interno canónico), siendo

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{d_1}, \ \mathbf{q}_j = \frac{\mathbf{a}_j - \sum_{i=1}^{j-1}\langle \mathbf{q}_i|\mathbf{a}_j\rangle \mathbf{q}_i}{d_j} \ j = 2, 3, \ldots, n$$

donde

$$d_1 = \|\mathbf{a}_1\|, \ d_j = \left\|\mathbf{a}_j - \sum_{i=1}^{j-1}\langle \mathbf{q}_i|\mathbf{a}_j\rangle \mathbf{q}_i\right\| \ j = 2, 3 \ldots, n$$

Estas relaciones pueden ser reescritas como

$$\mathbf{a}_1 = d_1\mathbf{q}_1, \ \mathbf{a}_j = \sum_{i=1}^{j-1}\langle \mathbf{q}_i|\mathbf{a}_j\rangle \mathbf{q}_i + d_j\mathbf{q}_j \ j = 2, 3, \ldots, n$$

las cuales implican la relación matricial

$$(\mathbf{a}_1|\mathbf{a}_2|\cdots|\mathbf{a}_n) = (\mathbf{q}_1|\mathbf{q}_2|\cdots|\mathbf{q}_n)\times$$
$$\begin{pmatrix} d_1 & \langle \mathbf{q}_1|\mathbf{a}_2\rangle & \langle \mathbf{q}_1|\mathbf{a}_3\rangle & \cdots & \langle \mathbf{q}_1|\mathbf{a}_n\rangle \\ 0 & d_2 & \langle \mathbf{q}_2|\mathbf{a}_3\rangle & \cdots & \langle \mathbf{q}_2|\mathbf{a}_n\rangle \\ 0 & 0 & d_3 & \cdots & \langle \mathbf{q}_3|\mathbf{a}_n\rangle \\ \vdots & \vdots & \vdots & \ddots & \ldots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}$$

De este modo, la matriz A $m \times n$ puede ser factorizada como $A = QR$, donde Q es una matriz $m \times n$ cuyas columnas son una base ortonormal del espacio columna de A y R es una matriz $n \times n$ triangular superior con elementos positivos sobre la diagonal.

La implementación algoritmica del método descrito para determinar la factorización QR de una matriz A, es entonces, como sigue:

Dadas $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$
Para $j = 1, 2, \dots, n$
    Tomar $\mathbf{q}_j = \mathbf{a}_j$
    Para $i = 1, 2, \dots, j-1$
        Calcular $r_{ij} = \langle \mathbf{q}_i | \mathbf{q}_j \rangle$
        Tomar $\mathbf{q}_j = \mathbf{q}_j - r_{ij}\mathbf{q}_i$
    Calcular $r_{jj} = \|\mathbf{q}_j\|$
    Tomar $\mathbf{q}_j = \mathbf{q}_j / r_{jj}$

De este modo, $Q = [\, \mathbf{q}_1 | \mathbf{q}_2 | \dots | \mathbf{q}_n]$ y $R = (r_{ij})$.

Ejercicio 8. Determinar la factorización QR de la matriz

$$A = \begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}$$