Descomposición QR de matrices con columnas linealmente independientes
Notas para los cursos 21 y 22 (J.L. Mancilla Aguilar)

En lo que sigue veremos que toda matriz $A \in \mathbb{K}^{n\times m}$ ($\mathbb{K} = \mathbb{R}$ ó $\mathbb{C}$) con rango$(A) = m$, es decir, cuyas columnas forman un conjunto linealmente independiente, se puede factorizar como producto de dos matrices, una de $n \times m$ cuyas columnas forman un conjunto ortonormal y otra de $m \times m$ que es triangular superior e inversible. Tal factorización, que se denomina descomposición QR, es muy utilizada en la resolución numérica de ecuaciones lineales y en el cálculo de autovalores.

Definición. Dada $A \in \mathbb{K}^{n\times m}$ con rango$(A) = m$, una descomposición QR de A es una factorización

$$A = QR \quad \text{con} \quad Q \in \mathbb{K}^{n\times m} \quad \text{y} \quad R \in \mathbb{K}^{m\times m}$$

tales que $Q^H Q = I$ y R es triangular superior e inversible.

Teorema 1. Sea $A \in \mathbb{K}^{n\times m}$ con rango$(A) = m$. Supongamos que $A = QR$ es una descomposición QR de A entonces

1. Las columnas de Q forman una base ortonormal de col(A);
2. $P = QQ^H$ es la matriz de proyección sobre col(A).

Demostración. Como el punto 2. es consecuencia directa del punto 1. por lo que ya hemos visto sobre matrices de proyección, sólo probaremos 1.

Sean $q_1, q_2, \ldots, q_m$ las columnas de Q. Como

$$Q^H Q = \begin{bmatrix} q_1^H \\ q_2^H \\ \vdots \\ q_m^H \end{bmatrix} [q_1 \ q_2 \ \cdots q_m] = \begin{bmatrix} q_1^H q_1 & q_1^H q_2 & \cdots & q_1^H q_m \\ q_2^H q_1 & q_2^H q_2 & \cdots & q_2^H q_m \\ \vdots & \vdots & \ddots & \vdots \\ q_m^H q_1 & q_m^H q_2 & \cdots & q_m^H q_m \end{bmatrix},$$

y por lo tanto $q_i^H q_j$ es el elemento ij del producto $Q^H Q$, tenemos que

$$Q^H Q = I \iff q_i^H q_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} \iff \{q_1, q_2, \ldots, q_m\} \text{ es un conjunto ortonormal.}$$

Entonces, claramente, $\{q_1, q_2, \ldots, q_m\}$ es una b.o.n. de col(Q). Si probamos que col$(Q) =$ col(A) tendremos probado el punto 1.

Para ello usaremos el siguiente resultado: si A, B y C son matrices tales que $A = BC$ entonces col$(A) \subset$ col(B).

La demostración de este resultado es la siguiente: si $y \in$ col(A) entonces existe x tal que $y = Ax = BCx$. Llamando $z = Cx$, tenemos que $y = Bz$, con lo cual $y \in$ col(B). Por lo tanto elemento de col(A) es a su vez elemento de col(B) con lo cual col$(A) \subset$ col(B).

Entonces, dado que $A = QR$ tenemos que col$(A) \subset$ col(Q). Como R es inversible por la definición de descomposición QR, tenemos que $Q = AR^{-1}$, con lo cual col$(Q) \subset$ col(A). Por lo tanto

$$\text{col}(A) \subset \text{col}(Q) \quad \text{y} \quad \text{col}(Q) \subset \text{col}(A) \quad \Longrightarrow \quad \text{col}(A) = \text{col}(Q).$$

En lo que sigue veremos que toda matriz A de $n \times m$ con rango m admite una descomposición QR.

Teorema 2. Sea $A \in \mathbb{K}^{n\times m}$ con $\text{rango}(A) = m$. Entonces existe una descomposición QR de A.

Demostración. Denominemos $v_1, v_2, \dots, v_m$ a las columnas de A. Por hipótesis $B = \{v_1, v_2, \dots, v_m\}$ es un conjunto linealmente independiente. Aplicando el procedimiento de Gram-Schmidt (G-S) al conjunto B obtenemos una base ortogonal $\{u_1, u_2, \dots, u_m\}$ de $\text{col}(A)$ que satisface las igualdades:

$$\begin{array}{rcl} u_1 & = & v_1 \\ u_2 & = & v_2 - \alpha_{12}u_1 \\ u_3 & = & v_3 - \alpha_{13}u_1 - \alpha_{23}u_2 \\ \vdots & \vdots & \quad\vdots \\ u_j & = & v_j - \alpha_{1j}u_1 - \alpha_{2j}u_2 - \cdots - \alpha_{(j-1)\,j}u_{j-1} \\ \vdots & \vdots & \quad\vdots \\ u_m & = & v_m - \alpha_{1m}u_1 - \alpha_{2m}u_2 - \cdots - \alpha_{(m-1)\,m}u_{m-1} \end{array} \qquad \text{con } \alpha_{ij} = \frac{u_i^H v_j}{\|u_i\|^2} \quad 1 \le i < j.$$

Entonces, despejando cada v_i obtenemos la serie de igualdades

$$\begin{array}{rcl} v_1 & = & u_1 \\ v_2 & = & \alpha_{12}u_1 + u_2 \\ v_3 & = & \alpha_{13}u_1 + \alpha_{23}u_2 + u_3 \\ \vdots & \vdots & \quad\vdots \\ v_j & = & \alpha_{1j}u_1 + \alpha_{2j}u_2 + \cdots + \alpha_{(j-1)\,j}u_{j-1} + u_j \\ \vdots & \vdots & \quad\vdots \\ v_m & = & \alpha_{1m}u_1 + \alpha_{2m}u_2 + \cdots + \alpha_{(m-1)\,m}u_{m-1} + u_m \end{array}.$$

que pueden escribirse en forma matricial

$$[v_1\ v_2\ \cdots\ v_m] = [u_1\ u_2\ \cdots u_m] \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1m} \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2m} \\ 0 & 0 & 1 & \cdots & \alpha_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Llamando $Q_0 = [u_1\ u_2\ \cdots u_m]$ y R_0 a la matriz triangular superior que aparece arriba, tenemos que $A = Q_0R_0$, que es casi la factorización que estamos buscando, ya que las columnas de Q_0 forman un conjunto ortogonal y R_0 es tringular superior e inversible. Lo que hacemos ahora es normalizar cada columna de Q_0, es decir, definimos

$$Q = [q_1\ q_2\ \cdots q_m] \quad \text{con} \quad q_i = \frac{u_i}{\|u_i\|}$$

y modificamos R_0 de modo tal que su producto con Q siga dando A, para ello multiplicamos cada fila de R_0 por el número por el cual dividimos la correspondiente columna de Q_0 obteniendo la matriz triangular superior e inversible:

$$R = \begin{bmatrix} \|u_1\| & \alpha_{12}\|u_1\| & \alpha_{13}\|u_1\| & \cdots & \alpha_{1m}\|u_1\| \\ 0 & \|u_2\| & \alpha_{23}\|u_2\| & \cdots & \alpha_{2m}\|u_2\| \\ 0 & 0 & \|u_3\| & \cdots & \alpha_{3m}\|u_3\| \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \|u_m\| \end{bmatrix} \tag{1}$$

Entonces $A = QR$ es la factorización QR buscada.

Ejemplo. Hallar una descomposición QR de

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

De acuerdo con la demostración del teorema anterior, para hallar tal descomposición deberíamos aplicar el precedimiento de G-S a las columnas de A y con los vectores obtenidos, previa normalización, construir la matriz Q. La matriz R podría obtenerse directamente mediante (1), calculando los α_{ij} mediante la fórmula $\alpha_{ij} = \frac{u_i^H v_j}{\|u_i\|^2}$. Sin embargo ello no es necesario, pues, una vez obtenida Q, como $A = QR$ y $Q^H Q = I$, tenemos que $Q^H A = Q^H(QR) = (Q^H Q)R = R$.

Procedemos entonces a aplicar G-S a las columnas de A: llamando v_i a la columna i de A tenemos que

$$u_1 = v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

$$u_2 = v_2 - \frac{u_1^H v_2}{\|u_1\|^2} u_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

$$u_3 = v_3 - \frac{u_1^H v_3}{\|u_1\|^2} u_1 - \frac{u_2^H v_3}{\|u_2\|^2} u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}.$$

Normalizando los u_i obtenidos formamos Q:

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

y calculamos R mediante

$$R = Q^H A = \begin{bmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{6}}{2} & \frac{\sqrt{6}}{6} \\ 0 & 0 & \frac{2\sqrt{3}}{3} \end{bmatrix}.$$

Nota. Cuando se calcula la descomposición QR de una matriz en forma numérica, es decir, empleando una computadora digital para hacer los cálculos, no se emplea el procedimiento de

Gram-Schmidt para calcular las columnas de Q debido a que los errores de redondeo pueden ser muy grandes. La descomposición se hace empleando otros métodos que involucran la utilización de las denominadas matrices de Householder (hay una introducción en wikipedia).

Varios programas que efectúan cálculos con matrices, como Matlab, Mathematica, Maple, Scilab (de uso libre, se baja de la red en la dirección http://www.scilab.org/) contienen instrucciones que calculan la descomposición QR de una matriz.

Descomposición QR y cuadrados mínimos.

La descomposición QR de una matriz A cuyas columnas forman un conjunto l.i. es muy útil en la resolución de ecuaciones lineales por cuadrados mínimos, porque permite hacerlo en forma eficiente y con gran precisión. Observamos que si A posee columnas l.i. y $A = QR$ es una descomposición QR de A entonces

$$A^H Ax = A^H b \quad \Longleftrightarrow \quad R^H Q^H QRx = R^H Q^H b \quad \Longleftrightarrow \quad R^H Rx = R^H Q^H b \quad \Longleftrightarrow \quad Rx = Q^H b,$$

la última equivalencia debido a que R^H es inversible por serlo R. Luego, las soluciones por cuadrados mínimos de la ecuación $Ax = b$ se pueden obtener resolviendo la ecuación

$$Rx = Q^H b,$$

lo cual tiene dos ventajas, una es que R es triangular y la otra es que, en general, el error que se comete al resolver de esta manera mediante una computadora digital es menor que el que se comete empleando la ecuación normal $A^H Ax = A^H b$.