UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE MATEMÁTICA

DESCOMPOSICIÓN QR DEL PRODUCTO DE MATRICES

Por

Flor Norma Quiñónez Cuyubamba

Informe de Suficiencia
Para optar el Título
Profesional de Licenciada
En Matemática

Prof. William Carlos Echegaray Castillo
Asesor

LIMA, PERÚ
2003

Resumen del Informe de Suficiencia

Bachiller: Flor Norma Quiñónez Cuyubamba

Tema : Descomposición QR del Producto de Matrices

El presente informe tiene como objetivo mostrar el producto de matrices descomposición QRP^T, es decir si se tiene $M_m = A_1A_2A_3...A_m$ de orden n en dicho producto se presentara una dificultad al efectuar en forma directa, es que cuando m aumenta, el número de operaciones aumenta.
Es por ello que dicho producto será realizado mediante una descomposición de la forma $A = QRP^T$, donde Q es ortogonal, R triangular superior y P es una permutación.

Indice

“Este trabajo va dedicado a mis padres, por su constante apoyo, sus consejos y sobre todo por brindarme mucho amor, en especial a mi madre Aurea Cuyubamba Ninahuanca, ser que ha guiado mis pasos.”

También va dedicado a la memoria de mi tía Graciela Quiñonez Rosales, por ser un ejemplo de fortaleza.

Agradecimientos

Deseo agradecer a todos los profesores por impartirme sus conocimientos, en especial a quien me asesoró en el desarrollo de este trabajo, me estoy refiriendo al Mg. William Carlos Echegaray Castillo.

También es preciso mencionar mi agradecimiento a la familia Pajuelo Rojas, por su incondicional apoyo, y a mi amiga Silvia por su paciencia. Del mismo modo a mi hermanita Cinthya quien supo alentarme en todo momento.

Finalmente, agradecer a un gran amigo por su invalorable contribución para el presente trabajo, gracias Piscis.

Capítulo 1

Introducción

La presente monografía se encuentra orientado a la descomposición QR del producto de matrices $M_m = A_1 A_2 A_3 ... A_m$ de orden n. La principal dificultad natural de dicho producto, es que cuando m aumenta, el número de operaciones aumentan.

En este trabajo enfocaremos como dicho producto puede realizarse a través de la descomposición de la forma $A = QRP^T$; donde Q es ortogonal, R triangular superior y P es una permutación.

En el capítulo 2 definiremos algunos conceptos preliminares que serán empleados durante el desarrollo del capítulo 3 donde se verá el algoritmo y su comportamiento.

Capítulo 2

Conceptos Preliminares

2.1 Imagen, Espacio Nulo y Rango de una Matriz

Existen dos subespacios importantes asociados a una matriz $A \in \mathbb{R}^{m \times n}$.

La imagen de A, denotada por $R(A)$ está definida por:

$$R(A) = \{y \in \mathbb{R}^m / y = Ax \quad \text{para algún} \quad x \in \mathbb{R}^n\},$$

y el espacio nulo de A, denotado por $N(A)$ está definido por:

$$N(A) = \{x \in \mathbb{R}^n / Ax = 0\}$$

Si $A = [a_1, \ldots, a_n]$, donde $a_i \in \mathbb{R}^m$ es un vector columna para todo $i = 1, \ldots, n$, entonces $R(A) = span\,\{a_1, \ldots, a_n\}$.

El rango de la matriz A, denotado por $rank(A)$ está definido por:

$$rank(A) = \dim(R(A)).$$

2.2 Algunas Matrices Especiales

2.2.1 Matrices Triangular y Diagonal

1. Una matriz $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ es una matriz diagonal si $a_{ij} = 0$, $\forall i \neq j$. Denotaremos $A = diag(a_{11}, \ldots, a_{ss})$, donde $s = \min\{m, n\}$.

2. Una matriz cuadrada $A = (a_{ij})$ es una matriz triangular superior, si $a_{ij} = 0$, $\forall i > j$.

3. La transpuesta de una matriz triangular superior es una matriz triangular inferior, esto es $A = (a_{ij})$ es triangular inferior si $a_{ij} = 0$, $\forall i < j$.

Algunas Propiedades de las Matrices Triangulares

Las siguientes propiedades de las matrices triangulares son útiles:

1. El producto de dos matrices triangular superior (inferior), es una matriz triangular superior (inferior).

2. La inversa de una matriz triangular superior (inferior) no singular, es una matriz triangular superior (inferior).

3. Los valores propios de una matriz triangular, son exactamente los elementos de su diagonal.

4. El determinante de una matriz triangular, es exactamente el producto de los elementos de su diagonal.

2.2.2 Matriz Ortogonal

Una matriz cuadrada U es ortogonal si $UU^T = U^TU = I$.

Propiedades

1. La inversa de una matriz ortogonal U, es justamente su transpuesta.

2. El producto de matrices ortogonales es una matriz ortogonal,

$$URV(URV)^T = I.$$

donde U, R, V son matrices ortogonales.

2.2.3 Matriz de Permutación

Sea $P \in \mathbb{R}^{n\times n}$, decimos que P es una *matriz de permutación*, si existe exactamente un elemento no nulo en cada fila y columna que es el 1 y los restantes son ceros. Luego si $(\alpha_1, \ldots, \alpha_n)$ es una permutación de $(1, 2, \ldots, n)$ entonces:

$$P = \begin{pmatrix} e_{\alpha_1}^T \\ e_{\alpha_2}^T \\ \vdots \\ e_{\alpha_n}^T \end{pmatrix},$$

donde e_i es la i-ésima columna de la matriz identidad $I \in \mathbb{R}^{n\times n}$, es una matriz de permutación.

Similarmente $P = \begin{pmatrix} e_{\alpha_1}, & e_{\alpha_2}, & \ldots & , e_{\alpha_n} \end{pmatrix}$, donde e_i es la i-ésima columna de la matriz identidad $I \in \mathbb{R}^{n\times n}$, es una matriz de permutación.

1. $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} e_2^T \\ e_3^T \\ e_1^T \end{pmatrix},$

$$P_1 A = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{pmatrix} = \begin{pmatrix} \text{2da fila de } A \\ \text{3era fila de } A \\ \text{1era fila de } A \end{pmatrix}$$

2. $P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} e_3 & e_1 & e_2 \end{pmatrix},$

$$AP_1 = \begin{pmatrix} a_{13} & a_{11} & a_{12} \\ a_{23} & a_{21} & a_{22} \\ a_{33} & a_{31} & a_{32} \end{pmatrix}$$
$$= (\text{3era columna de } A, \text{ 1era columna de } A, \text{ 2da columna de } A)$$

Propiedades de la matriz de Permutación

1. Las matrices de permutación son ortogonales.

2. Si P es una matriz de permutación entonces $P^{-1} = P^T$.

3. El producto de matrices de permutación, es una matriz de permutación.

4. La transpuesta de una matriz de permutación, es una matriz de permutación.

Prueba. 1. Sea $P = \begin{pmatrix} e_{\alpha_1}^T \\ e_{\alpha_2}^T \\ \vdots \\ e_{\alpha_n}^T \end{pmatrix}$ una matriz de permutación, donde $(\alpha_1, \ldots, \alpha_n)$ es una permutación de $(1, 2, \ldots, n)$ y e_i es la i-ésima columna de la matriz identidad $I \in \mathbb{R}^{n\times n}$. Luego

$$PP^T = \begin{pmatrix} e_{\alpha_1}^T \\ e_{\alpha_2}^T \\ \vdots \\ e_{\alpha_n}^T \end{pmatrix} \begin{pmatrix} e_{\alpha_1} & e_{\alpha_2} & \ldots & e_{\alpha_n} \end{pmatrix}$$

$$= \begin{pmatrix} e_{\alpha_1}^T e_{\alpha_1} & e_{\alpha_1}^T e_{\alpha_2} & \ldots & e_{\alpha_1}^T e_{\alpha_n} \\ e_{\alpha_2}^T e_{\alpha_1} & e_{\alpha_2}^T e_{\alpha_2} & \ldots & e_{\alpha_2}^T e_{\alpha_n} \\ \vdots & \vdots & \ddots & \vdots \\ e_{\alpha_n}^T e_{\alpha_1} & e_{\alpha_n}^T e_{\alpha_2} & \ldots & e_{\alpha_n}^T e_{\alpha_n} \end{pmatrix} = I,$$

por lo tanto P es ortogonal. ■

2.3 Transformaciones de Householder y Aplicaciones a la Factorización QR

Definición 2.1 *Una matriz de la forma $H = I - 2\frac{uu^T}{u^Tu}$, donde $I \in \mathbb{R}^{n\times n}$ matriz identidad y $u \in \mathbb{R}^n$, $u \neq 0$, es llamado matriz de Householder.*

1. La matriz de Householder es simétrica y ortogonal.

2. $H^2 = I$.

Lema 2.1 *Sea* $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$, *donde* $x \neq 0$. *Luego si*

$$u = x + sign(x_1)\|x\|_2 e_1,$$

donde $e_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^n$, *entonces la matriz de Householder*

$$H = I - 2\frac{uu^T}{u^Tu},$$

cumple $Hx = -sign(x_1)\|x\|_2 e_1 = (-sign(x_1)\|x\|_2, 0, \ldots, 0)^T$.

Prueba. Sea

$$\begin{aligned}
\beta &= \frac{2}{\mu^T\mu} = \frac{2}{(x^T + sign(x_1) \parallel x \parallel_2 e_1^T)(x + sign(x_1) \parallel x \parallel_2 e_1)} \\
&= \frac{2}{\parallel x \parallel_2^2 + 2sign(x_1) \parallel x \parallel_2 x_1 + sign^2(x_1) \parallel x \parallel_2^2} \\
&= \frac{1}{\parallel x \parallel_2 (\parallel x \parallel_2 + \mid x_1 \mid)}
\end{aligned}$$

Ahora veamos

$$\begin{aligned}
Hx &= (I - \beta uu^T)x = x - \beta u(x + sign(x_1)\|x\|_2 e_1)^T x \\
&= x - \beta u(\parallel x \parallel_2^2 + sign(x_1) \parallel x \parallel_2 x_1) \\
&= x - u = x - (x + sign(x_1)\|x\|_2 e_1) = -sign(x_1)\|x\|_2 e_1,
\end{aligned}$$

y por lo tanto el lema 2.1 queda demostrado. ■

Observación 2.1 *Para* $x_1 \in \mathbb{R}$, *definimos*

$$sign(x_1) = \begin{cases} 1 & , si\ x_1 \geq 0 \\ -1 & , si\ x_1 < 0 \end{cases}$$

2.3.1 Factorización QR con Matrices de Householder

Teorema 2.1 (de la Factorización QR de Householder) *Dada una matriz* $A \in \mathbb{R}^{n\times n}$ *no singular, existe una matriz ortogonal* $Q \in \mathbb{R}^{n\times n}$, *y una matriz triangular superior* $R \in \mathbb{R}^{n\times n}$ *tal que* $A = QR$.

El proceso para hallar la factorización QR usando matrices de Householder consta de $n-1$ pasos.

Denotemos por $A^{(k)} = (a_{ij}^{(k)}) \in \mathbb{R}^{n\times n}$ para $k = 0, 1, 2, \ldots, n-1$.

Inicialmente hacemos $A^{(0)} = A$.

Paso 1: Halle una matriz de Householder $H_1 \in \mathbb{R}^{n\times n}$ tal que:

$$A^{(1)} = H_1 A^{(0)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \ldots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{pmatrix},$$

para esto es suficiente construir $H_1 = I_n - 2\frac{u_n u_n^T}{u_n^T u_n}$, tal que:

$$H_1 \begin{pmatrix} a_{11}^{(0)} \\ a_{21}^{(0)} \\ \vdots \\ a_{n1}^{(0)} \end{pmatrix} = \begin{pmatrix} a_{11}^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Paso 2: Halle una matriz $H_2 \in \mathbb{R}^{n\times n}$ tal que

$$A^{(2)} = H_2 A^{(1)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \ldots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \ldots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & \ldots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \ldots & a_{nn}^{(2)} \end{pmatrix}.$$

H_2 es construido como sigue. Primero construimos una matriz de Householder

$$\tilde{H}_2 = I_{n-1} - 2\frac{u_{n-1}u_{n-1}^T}{u_{n-1}^T u_{n-1}} \in \mathbb{R}^{(n-1)\times(n-1)},$$

tal que

$$\tilde{H}_2 \begin{pmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{pmatrix} = \begin{pmatrix} * \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

y luego definimos

$$H_2 = \begin{pmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & \tilde{H}_2 \end{pmatrix}.$$

Paso k ($k = 2, 3, \ldots, n-1$): Primero creamos una matriz de Householder

$$\tilde{H}_k = I_{n-k+1} - 2\frac{u_{n-k+1}u_{n-k+1}^T}{u_{n-k+1}^T u_{n-k+1}} \in \mathbb{R}^{(n-k+1)\times(n-k+1)},$$

tal que

$$\tilde{H}_k \begin{pmatrix} a_{kk}^{(k-1)} \\ a_{k+1,k}^{(k-1)} \\ \vdots \\ a_{nk}^{(k-1)} \end{pmatrix} = \begin{pmatrix} * \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

y luego, definimos

$$H_k = \begin{pmatrix} I_{k-1} & 0_{(k-1)\times(n-k+1)} \\ 0_{(n-k+1)\times(k-1)} & \tilde{H}_k \end{pmatrix}.$$

Calculamos $A^{(k)} = H_k A^{(k-1)}$.

Aquí la matriz $A^{(k)}$ tendrá ceros debajo del elemento (k, k), en la k-ésima columna, y los ceros ya creados en las columnas previas no serán modificados.

Al final del paso $n-1$, tendremos que $A^{(n-1)} = R$, donde R es una matriz triangular superior.

Ahora como

$$A^{(k)} = H_k A^{(k-1)}, \text{ para } k = 1, 2, \ldots, n-1,$$

entonces

$$R = A^{(n-1)} = H_{n-1}A^{(n-2)} = H_{n-1}H_{n-2}A^{(n-3)} = H_{n-1}H_{n-2}\ldots H_1 A.$$

Llamamos $Q^T = H_{n-1}H_{n-2}\ldots H_1$, entonces $R = Q^T A$, y $Q = H_1^T H_2^T \ldots H_{n-1}^T$, pero puede comprobarse que $H_1, H_2, \ldots, H_{n-1}$ son simétricas y ortogonales, entonces $Q = H_1 H_2 \ldots H_{n-1}$, y Q es ortogonal, luego $A = QR$.

Ejemplo 2.1 *Sea*

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Paso *$k = 1$: Construimos la matriz de Householder H_1 tal que*

$$H_1 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} * \\ 0 \\ 0 \end{pmatrix},$$

$$u_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + sign(0)(0^2 + 1^2 + 1^2)^{1/2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \sqrt{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix},$$

$$\begin{aligned} H_1 &= I_3 - 2\frac{u_3 u_3^T}{u_3^T u_3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 2\frac{\begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 1 & 1 \end{pmatrix}}{\sqrt{2}^2 + 1^2 + 1^2} \\ &= \begin{pmatrix} 0 & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}. \end{aligned}$$

Calculamos

$$A^{(1)} = H_1 A^{(0)} = H_1 A = \begin{pmatrix} -1.4142 & -2.1213 & -2.8284 \\ 0 & -0.2071 & 0.2929 \\ 0 & -1.2071 & -1.7071 \end{pmatrix}.$$

Paso *$k = 2$: Construimos la matriz de Householder $\tilde{H}_2$ tal que*

$$\tilde{H}_2 \begin{pmatrix} -0.2071 \\ -1.2071 \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix},$$

$$
\begin{aligned}
u_2 &= \begin{pmatrix} -0.2071 \\ -1.2071 \end{pmatrix} + sign(-0.2071)\sqrt{(-0.2071)^2 + (-1.2071)^2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} -1.4318 \\ -1.2071 \end{pmatrix},
\end{aligned}
$$

$$
\tilde{H}_2 = I_2 - 2\frac{u_2 u_2^T}{u_2^T u_2} = \begin{pmatrix} -0.1691 & -0.9856 \\ -0.9856 & 0.1691 \end{pmatrix}.
$$

Construimos

$$
H_2 = \begin{pmatrix} I_{1\times 1} & 0_{1\times 2} \\ 0_{2\times 1} & \tilde{H}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.1691 & -0.9856 \\ 0 & -0.9856 & 0.1691 \end{pmatrix}.
$$

Calculamos

$$
A^{(2)} = H_2 A^{(1)} = \begin{pmatrix} -1.4142 & -2.1213 & -2.8284 \\ 0 & 1.2247 & 1.6330 \\ 0 & 0 & -0.5774 \end{pmatrix} = R.
$$

Ahora

$$
Q = H_1 H_2 = \begin{pmatrix} 0 & 0.8165 & 0.5774 \\ -0.7071 & 0.4082 & -0.5774 \\ -0.7071 & -0.4082 & 0.5774 \end{pmatrix}.
$$

2.4 Rotaciones de Givens y Aplicaciones a la Factorización QR

Una matriz de Givens es de la forma:

$$
G(i,k,\theta) = \begin{array}{cc}
\begin{array}{cccccccc} & & & & i & k & & \\ & & & & \downarrow & \downarrow & & \end{array} & \text{columnas} \\
\begin{pmatrix}
1 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & & & & \cdots & \vdots \\
0 & 0 & 0 & \cdots & c & s & \cdots & 0 \\
\vdots & \vdots & \vdots & & & & \cdots & \vdots \\
0 & 0 & 0 & \cdots & -s & c & \cdots & 0 \\
\vdots & \vdots & \vdots & & & & \cdots & \vdots \\
0 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 1
\end{pmatrix}_{n\times n} & \begin{array}{l} \\ \\ \\ \leftarrow i \\ \qquad \text{filas} \\ \leftarrow k \\ \\ \\ \end{array}
\end{array}
$$

donde $c = \cos\theta$, $s =$sen θ

El efecto de una matriz Givens aplicado a un vector es rotarlo un ángulo de θ grados.

Proposición 2.2 *1. Como $c^2 + s^2 = 1$, los $G(i,k,\theta)$ son ortogonales.*

2. Multiplicación por un vector.

Cuando un vector n-dimensional:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

es multiplicado por una matriz de Givens, solo las componentes i y k son *afectadas, las otras componentes permanecen iguales, es decir:*

$$Y = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ cx_i + sx_k \\ x_{i+1} \\ \vdots \\ x_{k-1} \\ -sx_i + cx_k \\ x_{k+1} \\ \vdots \\ x_n \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow k \\ \\ \\ \\ \end{array}$$

luego cada componente de Y puede ser expresado como sigue:

$$y_j = \begin{cases} cx_i + sx_k, & \text{si } j = i \\ -sx_i + cx_k, & \text{si } j = k \\ x_j, & \text{si } j \neq i, k \end{cases}$$

3. *Multiplicando una matriz $A \in \mathbb{R}^{m\times n}(m \geq n)$ por la izquierda con una matriz de Givens $G(i,k,\theta)$ tal que $1 \leq i \leq k \leq m$ el único efecto es un cambio de la i-ésima y k-ésima filas. Las nuevas filas son una combinación lineal de las anteriores y están dadas por:*

$$\begin{aligned} (G_{ik}A)_{ij} &= ca_{ij} + sa_{kj} \\ (G_{ik}A)_{kj} &= -sa_{ij} + ca_{kj}, \qquad j = 1, 2, ..., n, \end{aligned}$$

en particular, si $a_{ij} = a_{kj} = 0$, entonces

$$(G_{ik}A)_{ij} = 0, \ (G_{ik}A)_{kj} = 0.$$

Ejemplo 2.2 *Sea* $G(1,2,\theta) = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$ *y el vector* $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, *hallar los valores de c y s para transformar X en* $\begin{pmatrix} * \\ 0 \end{pmatrix}$.

Solución

$$G(1,2,\theta)X = \begin{pmatrix} cx_1 + sx_2 \\ -sx_1 + cx_2 \end{pmatrix}$$

igualando la segunda fila a cero resulta $c = \dfrac{sx_1}{x_2}$ y usando el hecho que $c^2+s^2=1$ se llega a que $s^2(\dfrac{x_1^2}{x_2^2}+1)=1$, de donde $s = \dfrac{x_2}{\sqrt{x_1^2+x_2^2}}$, $c = \dfrac{x_1}{\sqrt{x_1^2+x_2^2}}$.

Sin embargo, el cálculo anterior puede causar overflow o underflow y se recomienda lo siguiente:

i) Si $|x_2| \geqslant |x_1|$ entonces $t = \dfrac{x_1}{x_2}$, $s = \dfrac{1}{\sqrt{1+t^2}}$, $c = st$.

ii) Si $|x_2| < |x_1|$ entonces $t = \dfrac{x_2}{x_1}$, $c = \dfrac{1}{\sqrt{1+t^2}}$, $s = ct$.

2.4.1 Haciendo ceros en una posición especificada de un vector

Sea el vector

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix}$$

supongamos que queremos hacer cero solamente la posición x_k del vector. Entonces se debe usar la matriz $G(i,k,\theta)$ con $i < k$, de manera que

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x_i \\ x_k \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

Ejemplo 2.3 *Sea el vector*

$$X = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$$

hallar la matriz de Givens que hace cero la tercera componente y hallar el nuevo vector.

Solución:

Paso1:

$k = 3, i = 2$

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$s = \dfrac{3}{\sqrt{10}}, c = \dfrac{-1}{\sqrt{10}}$

Paso 2:

$$G(2,3,\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{-1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ 0 & \frac{-3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \end{pmatrix}$$

entonces

$$G(2,3,\theta)X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{-1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ 0 & \frac{-3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ \sqrt{10} \\ 0 \end{pmatrix}.$$

2.4.2 Haciendo ceros en una posición especificada de una matriz

Si se desea crear un cero en la posición (k, i), $k > i$ de una matriz de Givens $G(i, k, \theta)$, la cual al multiplicar por A sólo afectará las filas i, k de A.

Ejemplo 2.4 *Sea la matriz*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{pmatrix}$$

hacer cero la posición $(2,1)$ *usando rotación de Givens.*

Solución:

Paso 1:

$G(1,2,\theta)$

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$$c = \frac{1}{\sqrt{5}}, \quad s = \frac{2}{\sqrt{5}}$$

Paso 2:

$$G(1,2,\theta) = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$G(1,2,\theta)A = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -\frac{3}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \\ 0 & -\frac{1}{\sqrt{5}} \\ 4 & 5 \end{pmatrix}$$

2.4.3 Factorización QR con rotaciones de Givens

Similar a lo que se hizo con las matrices de Householder, la factorización QR de una matriz puede ser obtenida usando rotaciones de Givens, lo cual es más conveniente para calcular los valores propios de una matriz y para resolver sistemas de ecuaciones lineales con matrices que tienen muchos ceros. Aunque el tiempo de cómputo es casi el doble comparado con matrices Householder se pueden adaptar más facilmente a computación paralela.

Consideremos la matriz $A \in \mathbb{R}^{m \times n}$. Lo que se busca es convertir a A en una matriz triangular superior R paso a paso.

En el primer paso se hace cero debajo de la posición (1,1) de la primera columna obteniéndose una matriz $A^{(1)} = Q_1 A$ donde

$$Q_1 = G(1,m,\theta)G(1,m-1,\theta)\ldots G(1,2,\theta)$$

es ortogonal, ya que el producto de matrices ortogonales es ortogonal.

$$A^{(1)} \to Q_1 A = \begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ & & & & \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \end{pmatrix}.$$

En el segundo paso se hace debajo de la posición (2,2) de la segunda columna obteniéndose una matriz $A^{(2)} = Q_2 A^{(1)}$ donde

$$Q_2 = G(2,m,\theta)G(2,m-1,\theta)\ldots G(2,3,\theta)$$

$$A^{(2)} \to Q_2 A^{(1)} = \begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & * & \cdots & * \end{pmatrix}$$

es también ortogonal y así sucesivamente.

Sea s = min(n,m-1), entonces:

$$R = A^{(s)} = Q_s A^{(s-1)} = Q_s Q_{s-1} A^{(s-2)} = \ldots = Q_s Q_{s-1} \ldots Q_2 Q_1 A = Q^T A$$

o equivalentemente $A = QR$ con $Q^T = Q_s Q_{s-1} A^{(s-2)} = \ldots = Q_s Q_{s-1} \ldots Q_2 Q_1$

$$A^{(s)} \to Q_s A^{(s-1)} = \begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & * \\ 0 & 0 & 0 & \cdots & * \end{pmatrix} = R$$

Ejemplo 2.5 *Sea la matriz*

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

encontrar la factorización QR usando la rotacion de Givens.

Solución:

***Paso 1**: Encontrar c y s tal que*

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$$\begin{aligned} a_{11} &= 0, & a_{21} = 1 \\ c &= 0, & s = 1 \end{aligned}$$

$$J(1,2,\theta) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A \equiv J(1,2,\theta)A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

Encontraremos c y s tal que

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{31} \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$$\begin{aligned} a_{11} &= 1, \qquad & a_{31} = 1 \\ c &= \frac{1}{\sqrt{2}}, \qquad & s = \frac{1}{\sqrt{2}} \end{aligned}$$

$$J(1,3,\theta) = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$A^{(1)} \equiv J(1,3,\theta)A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & -1 & -1 \\ 0 & -\frac{1}{\sqrt{2}} & -\sqrt{2} \end{pmatrix}$$

Paso 2: *Encontrar c y s tal que*

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{22} \\ a_{32} \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$$\begin{aligned} a_{22} &= -1, \qquad & a_{32} = -\frac{1}{\sqrt{2}} \\ c &= -\frac{\sqrt{2}}{\sqrt{3}}, \qquad & s = -\frac{1}{\sqrt{3}} \end{aligned}$$

$$J(2,3,\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{\sqrt{2}}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}$$

$$A^{(2)} \equiv J(2,3,\theta)A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{\sqrt{2}}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & -1 & -1 \\ 0 & -\frac{1}{\sqrt{2}} & -\sqrt{2} \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{2\sqrt{2}}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

por lo tanto

$$R = \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{2\sqrt{2}}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

2.5 Transformación Rápida de Givens

La habilidad de introducir ceros en forma selectiva hace que las rotaciones de Givens sea una herramienta importante en ciertos problemas estructurados. Esto ha conducido al desarrollo de procedimientos rápidos de Givens. La idea de este procedimiento equivale a una representación mejor de Q, cuando Q es el producto de Rotaciones Givens.

En particular, Q es representado por un par de la matriz (M, D) donde $M^T M = D = diag(d_i)$ y cada d_i es positivo. Las matrices Q, M y D están relacionadas a través de la formula

$$Q = MD^{-1/2} = M\ diag(1/\sqrt{d_i})$$

Notar que $Q^TQ = I$, pues $(MD^{-1/2})^T(MD^{-1/2}) = I$

y así la matriz $MD^{-1/2}$ es ortogonal.

Los detalles son mejor explicados en una matriz de 2x2. Sean $X = (x_1, x_2)^T$ y $D = diag(d_1, d_2)$ dados y se asumirá que d_1 y d_2 son positivos.

Se define

$$M_1 = \begin{pmatrix} \beta_1 & 1 \\ 1 & \alpha_1 \end{pmatrix}$$

y se tiene que

$$M_1^T X = \begin{pmatrix} \beta_1 x_1 + x_2 \\ x_1 + \alpha x_2 \end{pmatrix}$$

y

$$M_1^T D M_1 = \begin{pmatrix} d_2 + \beta_1^2 d_1 & d_1\beta_1 + d_2\alpha_1 \\ d_1\beta_1 + d_2\alpha_1 & d_1 + \alpha_1^2 d_2 \end{pmatrix} = D_1$$

si $x_2 \neq 0$ y $\alpha_1 = -\dfrac{x_1}{x_2}$ y $\beta_1 = -\dfrac{\alpha_1 d_2}{d_1}$ entonces

$$M_1^T X = \begin{pmatrix} x_2(1+\gamma_1) \\ 0 \end{pmatrix}$$

$$M_1^T D M_1 = \begin{pmatrix} d_2(1+\gamma_1) & 0 \\ 0 & d_1(1+\gamma_1) \end{pmatrix}$$

donde $\gamma_1 = -\alpha_1\beta_1 = \left(\frac{d_2}{d_1}\right)\left(\frac{x_1}{x_2}\right)^2$

Análogamente, asumimos $x_1 \neq 0$ y definimos M_2 por

$$M_2 = \begin{pmatrix} 1 & \alpha_2 \\ \beta_2 & 1 \end{pmatrix}$$

donde $\alpha_2 = -\frac{x_2}{x_1}$ y por $\beta_2 = -\left(\frac{d_1}{d_2}\right)\alpha_2$, entonces

$$M_2^T X = \begin{pmatrix} x_1(1+\gamma_2) \\ 0 \end{pmatrix}$$

y

$$M_2^T D M_2 = \begin{pmatrix} d_1(1+\gamma_2) & 0 \\ 0 & d_2(1+\gamma_2) \end{pmatrix} = D_2$$

donde $\gamma_2 = -\alpha_2\beta_2 = \left(\frac{d_1}{d_2}\right)\left(\frac{x_2}{x_1}\right)^2$

Notar que , $J = D^{1/2} M_i D_i^{-1/2}$ es ortogonal, $i = 1, 2$.

En efecto: para $i = 1$

$$\begin{aligned} J^T J &= (D_1^{-1/2})^T M_1{}^T (D^{1/2})^T (D^{1/2} M_1 D_1^{-1/2}) \\ &= (D_1^{-1/2})^T M_1{}^T D M_1 D_1^{-1/2} \\ &= (D_1^{-1/2})^T D_1 D_1^{-1/2} \\ &= I \end{aligned}$$

Análogamente para $i = 2$.

Además se tiene que el segundo componente de $J^T(D^{-1/2}x)$ es cero.

Veamos

$$\begin{aligned}
J^T(D^{-1/2}x) &= (D_1^{-1/2})^T {M_1}^T (D^{-1/2})^T D^{-1/2}x \\
&= (D_1^{-1/2})^T {M_1}^T x \\
&= (D_1^{-1/2})^T \begin{pmatrix} x_2(1+\gamma_1) \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\sqrt{d_1}} x_2(1+\gamma_1) \\ 0 \end{pmatrix}
\end{aligned}$$

Note que γ_i satisface , $\gamma_1\gamma_2 = \left(\frac{d_2}{d_1}\right)\left(\frac{x_1}{x_2}\right)^2\left(\frac{d_1}{d_2}\right)\left(\frac{x_2}{x_1}\right)^2 = 1$, siempre podemos seleccionar el M_i de manera que el factor de crecimiento $(1+\gamma_i)$ sea acotado por 2. Las matrices de la forma:

$$M_1 = \begin{pmatrix} \beta_1 & 1 \\ 1 & \alpha_1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 1 & \alpha_2 \\ \beta_2 & 1 \end{pmatrix}$$

entonces $-1 \leq \alpha_i\beta_i \leq 0$.

veamos para i=1 :

$$\alpha_1\beta_1 = -\left(\frac{d_2}{d_1}\right)\left(\frac{x_1}{x_2}\right)^2 \text{ entonces } \alpha_1\beta_1 \leq 0 \text{ , pues } \frac{d_2}{d_1} > 0 \text{ ; } \frac{x_1}{x_2} \geq 0.$$

Por otra parte se tiene que el factor $(1+\gamma_1) < 2$ entonces $\gamma_1 \leq 1$ pero $\gamma_1 = -\alpha_1\beta_1$ por lo tanto $-1 \leq \alpha_1\beta_1 \leq 0$.

Las matrices mencionadas son transformaciones rápidas de Givens de orden 2x2.

Notar que la premultiplicación por una transformación rápida de Givens implica la mitad del número de multiplicaciones que la premultiplicación por una transformación ordinaria de Givens.

La primera forma de la transformación es :

$$F(i,k,\alpha,\beta) = \begin{array}{cl} \begin{array}{cccccccc} & & & & i & k & & \\ & & & & \downarrow & \downarrow & & \end{array} & \text{columnas} \\ \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \alpha & \cdots & 0 \\ \vdots & \vdots & \vdots & & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i \\ \quad \text{filas} \\ \leftarrow k \\ \\ \\ \end{array} \end{array}$$

mientras que la segunda forma de la transformación es :

$$F(i,k,\alpha,\beta) = \begin{array}{cl} \begin{array}{cccccccc} & & & & i & k & & \\ & & & & \downarrow & \downarrow & & \end{array} & \text{columnas} \\ \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \alpha & \cdots & 0 \\ \vdots & \vdots & \vdots & & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & & & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i \\ \quad \text{filas} \\ \leftarrow k \\ \\ \\ \end{array} \end{array}$$

Así , si $M \epsilon \mathbb{R}^{nxn}$ y $D = diag(d_i)$ satisface $M^T M = D$ y si F es una transformación rápida de Givens de orden $n \times n$, con $F^T D F = D_{nuevo}$ diagonal, entonces $M^T_{nuevo} M_{nuevo} = D_{\ nuevo}$ donde $M_{nuevo} = MF$.

2.6 Factorización QR con Columna Pivoteada

Si $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ y $rank(A) < n$, entonces la factorización QR no necesariamente produce una base ortonormal para el $R(A)$.

Ejemplo 2.6 *Si A tiene la siguiente forma:*

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$$

y su factorización QR es:

$$A = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Veamos que su $rank(A) = 2$.

Se tiene que

$$A = \begin{bmatrix} q_1 & q_1 & q_1 + q_2 + q_3 \end{bmatrix}$$

ahora demostraremos que $\{q_1, q_1 + q_2 + q_3\}$ son linealmente independientes.

Supongamos que $\{q_1, q_1 + q_2 + q_3\}$ son linealmente dependientes.Entonces existen escalares $(\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0,0)\}$ tal que:

$$\alpha q_1 + \beta(q_1 + q_2 + q_3) = 0$$

Si $\alpha \neq 0$ y $\beta = 0 \Rightarrow \alpha q_1 = 0 \Rightarrow q_1 = 0\,(\Rightarrow\Leftarrow)$.

Si $\alpha = 0$ y $\beta \neq 0 \Rightarrow \beta(q_1 + q_2 + q_3) = 0 \Rightarrow q_1 + q_2 + q_3 = 0$.

Se sabe que Q es ortogonal entonces $q_i^T q_i > 0$, y $q_i^T q_j = 0$, $i \neq j$.

Luego multiplicando por q_1^T se tiene:

$q_1^T q_1 + q_1^T q_2 + q_1^T q_3 = 0 \Rightarrow q_1^T q_1 = 0\,(\Rightarrow\Leftarrow)$.

Si $\alpha \neq 0$ y $\beta \neq 0$ y multiplicando por q_2^T se tiene:

$\alpha q_2^T q_1 + \beta q_2^T q_1 + \beta q_2^T q_2 + \beta q_2^T q_3 = 0 \Rightarrow \beta q_2^T q_2 = 0 \Rightarrow q_2^T q_2 = 0\,(\Rightarrow\Leftarrow)$.

Por lo tanto $\{q_1, q_1 + q_2 + q_3\}$ es linealmente independiente.

Luego $rank(A) = 2 = \dim(Span\{a_1, a_2, a_3\})$ y

$$\begin{aligned} R(A) &= Span\{q_1, q_1 + q_2 + q_3\} \\ R(A) &\neq Span\{q_1, q_2\} \\ R(A) &\neq Span\{q_2, q_3\} \\ R(A) &\neq Span\{q_1, q_3\} \end{aligned}$$

Afortunadamente el procedimiento de factorización QR (por ejemplo el método de Householder), puede ser modificado de manera simple, para producir una base ortonormal, para el $R(A)$.

La idea es generar una matriz de permutación P tal que

$$AP = QR \tag{2.1}$$

donde

$$R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix}$$

donde $r = rank(A)$, Q es ortogonal, $R_{11} \in \mathbb{R}^{r\times r}$ es triangular superior no singular. Las r primeras columnas de Q entonces formarán una base ortonormal para el $R(A)$.

Antes de probar el teorema 2.4 que garantiza la existencia de la factorización dada en (2.1) enunciaremos el siguiente teorema que usaremos en su demostración.

Teorema 2.3 *Sea $A \in \mathbb{R}^{m\times n}$ con $rank(A) = n$. Entonces existe una matriz ortogonal $Q \in \mathbb{R}^{m\times m}$ tal que*

$$Q^T A = \begin{pmatrix} R \\ 0 \end{pmatrix},$$

donde $R \in \mathbb{R}^{n\times n}$ es triangular superior con elementos diagonales positivos. La matriz R es única, como son las primeras n columnas de Q.

Prueba. Ver demostración en [4]. ■

Teorema 2.4 *(Columna Pivoteada QR). Sea $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ con $rank(A) = r \leq n$. Entonces existen una matriz de permutación $P \in \mathbb{R}^{n\times n}$ y una matriz ortogonal $Q \in \mathbb{R}^{m\times m}$ tal que*

$$Q^T AP = \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix}$$

donde R_{11} es triangular superior con elementos en la diagonal no nulos.

Prueba. Si $rank(A) = r$, existe una matriz de permutación P tal que

$$AP = \begin{pmatrix} A_1 & A_2 \end{pmatrix}$$

donde $A_1 \in \mathbb{R}^{m\times r}$ y sus columnas son linealmente independientes. Consideremos ahora la factorización QR de A_1:

$$Q^T A_1 = \begin{pmatrix} R_{11} \\ 0 \end{pmatrix}$$

donde $R_{11} \in \mathbb{R}^{r\times r}$ es triangular superior y sus elementos en la diagonal son no nulos. Entonces

$$Q^T AP = \begin{pmatrix} Q^T A_1 & Q^T A_2 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$$

$rank(Q^T AP) = rank(A) = r$, y $rank(R_{11}) = r$, tenemos que $R_{22} = 0$. ■

Observación 2.2 *Sea $A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$ la factorización QR de A . Entonces para cualquier matriz de permutacion P, $Q^T AP = \begin{pmatrix} RP \\ 0 \end{pmatrix}$.*

Sea $U^T(RP) = \hat{R}$ la factorización QR de RP.

Entonces $diag(U, I)^T Q^T AP = \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix}$ es una factorización QR de AP.

2.6.1 Creando una matriz de Permutación

La matriz de permutación P es formado a través del producto de r matrices de permutación desde P_1 hasta P_r. Estas matrices de permutaciones son aplicadas a la matriz A, uno por uno, antes formamos matrices de Householder creando ceros en las columnas apropiadas. Especificamente hacemos lo siguiente:

Paso 1: Encontrar la columna de A de mayor norma. Intercambiar la columna de A mayor norma con la primera columna, donde

$$\parallel x \parallel = \left(\sum x_i^2\right)^{\frac{1}{2}}, \ x = (x_1, x_2, x_3, ..., x_n)^T \epsilon \mathbb{R}^n.$$

Esto es equivalente a formar una matriz de permutación P_1 tal que la matriz AP_1 tenga en la primera columna la mayor norma. Formamos una matriz Householder H_1 de manera que:

$$A^{(1)} = H_1 A P_1$$

tenga ceros en la primera columna debajo de la posición $(1,1)$.

Paso 2: Encontrar la columna de máxima norma de la submatriz $\hat{A}^{(1)}$ obtenida de $A^{(1)}$ eliminando la primera fila y la primera columna. Permutar las columnas de $A^{(1)}$ de manera que la columna de $\hat{A}^{(1)}$ de norma máxima llegue a ser la primera columna de $\hat{A}^{(1)}$. Esto es equivalente a construir una matriz de permutación $\hat{P}_2$ tal que la primera columna de $\hat{A}^{(1)}\hat{P}_2$ tenga norma máxima y luego postmultiplicar $A^{(1)}$ por una matriz de permutación P_2 definido por:

$$P_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \hat{P}_2 & \\ 0 & & & \end{pmatrix}$$

Ahora construimos una matriz de Householder H_2 de manera que:

$$A^{(2)} = H_2 A^{(1)} P_2 = H_2 H_1 A P_1 P_2$$

tenga ceros en la segunda columna de $A^{(2)}$ debajo de la posición $(2, 2)$.

Supongamos que necesitamos r pasos. Entonces en el final del paso r-ésimo, tenemos,

$$\begin{aligned} A &= A^{(r)} = H_r \ldots H_1 A P_1 \ldots P_r \\ &= Q^T A P = R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Ejemplo 2.7 *Sea*

$$A = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 1 \\ \frac{1}{2} & 1 \end{pmatrix}_{3\times 2} = (a_1, a_2)$$

primero encontramos la columna que tenga la norma máxima, en este caso es la columna a_2.

$$\begin{aligned} P_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ AP_1 &= \begin{pmatrix} 0 & 0 \\ 1 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{pmatrix} \\ H_1 &= \begin{pmatrix} 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/2 & -1/2 \\ -1/\sqrt{2} & -1/2 & 1/2 \end{pmatrix} \end{aligned}$$

$$A^{(1)} = H_1 A P_1 = \begin{pmatrix} 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/2 & -1/2 \\ -1/\sqrt{2} & -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} -\sqrt{2} & -\sqrt{2}/2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = R$$

$$R = \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix}$$

asi, para el ejemplo,

$$\begin{aligned} Q &= H_1^T = H_1 \\ P &= P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

La matriz A tiene rango uno, pues $R_1 = \sqrt{2}$ y su orden es 1×1.

El vector columna $\begin{pmatrix} 0 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ *forma una base ortonormal para el $R(A)$.*

Capítulo 3

Descomposición QR del producto de matrices

Trabajaremos con descomposiciones de la forma $A = QRP^T$, donde Q es ortogonal R es triangular superior, y P es una permutación.

3.1 Algoritmo

En esta sección describiremos el algoritmo para actualizar la descomposiciones QRP. Esta descripción se hará en dos fases: la primera una visión general sobre la matriz, la segunda una descripción detallada.

La entrada en el algoritmo es la de descomposición QRP de la matriz

$$A = Q_A R_A P_A^T$$

y una matriz B.no factorizada.

La salida será una descomposición QRP de

$$C = AB = Q_C R_C P_C^T \tag{3.1}$$

Los pasos son los siguientes:

1. Calcular la descomposición QRP de $P_A^T B = Q_B R_B P_C^T$.

2. Calcular la matriz ortogonal U tal que $\hat{R}_A = U^T R_A Q_B$ es una matriz triangular superior.

3. $Q_C = Q_A U$.

4. $R_C = \hat{R}_A R_B$.

Verifiquemos que las cantidades mencionadas satisfacen (3.1)

$$\begin{aligned} C &= AB = (Q_A R_A)(P_A^T B) \\ &= Q_A R_A (Q_B R_B P_C^T) \\ &= (Q_A U)(U^T R_A Q_B R_B) P_C^T \\ &= (Q_A U)(U^T R_A Q_B) R_B P_C^T \\ &= Q_C \hat{R}_A R_B P_C^T \\ &= Q_C R_C P_C^T. \end{aligned}$$

En algunas aplicaciones es necesario trabajar con productos de la forma AB^{-1}. El algoritmo mencionado puede ser adaptado para calcular la descomposición QRP de AB^{-1}. La idea es calcular una descomposición PRQ de B, de tal manera que cuando la descomposición es invertida, se convierte en una descomposición QRP de B^{-1}. Esto conduce al siguiente algoritmo.

1. Calcular la descomposición PRQ de $BP_A = P_C R_B Q_B^T$

   Se tiene que : $AB^{-1} = Q_A R_A P_A^T B^{-1}$

   Veamos: $P_A^T B^{-1} = Q_B R_B P_C^T$

$$\begin{aligned} (P_A^T B^{-1})^{-1} &= (Q_B R_B P_C^T)^{-1} \\ B(P_A^T)^{-1} &= (P_C^T)^{-1} R_B^{-1} Q_B^{-1} \\ BP_A &= P_C R_B^{-1} Q_B^T, \qquad \text{pues } P_A^{-1} = P_A^T \\ BP_A &= P_C R_B' Q_B^T \end{aligned}$$

2. Calcular la matriz ortogonal U tal que $\hat{R}_A = U^T R_A Q_B$.

3. $Q_C = Q_A U$.

4. $R_C = \hat{R}_A R_B^{-1}$.

En algunas aplicaciones en las cuales deseamos las potencias de una matriz A, debemos proceder con matrices cuadradas; esto es calcular la secuencia $A, A^2, A^4, \ldots$.

Así dada la descomposición QRP de A^k, calcularemos la descomposición QRP de $A^k A^k$.

En forma más general dadas las descomposiciones QRP de A y B es posible calcular una descomposición de su producto; es decir de $(Q_A R_A P_A^T)(Q_B R_B P_B^T)$.

El algoritmo es el siguiente:

1. Calcular la descomposición-QR $P_A^T Q_B = VD$. Note que D será diagonal porque $P_A^T Q_B$ es ortogonal.

2. Calcular la matriz ortogonal U tal que $\hat{R}_A = U^T R_A V$ es una matriz triangular.

3. $P_C = P_B$, $Q_C = Q_A U$ y $R_C = \hat{R}_A D R_B$.

Veremos algunos detalles de los tres algoritmos mencionados, los cuales son variantes uno de otro; trataremos solo el primer algoritmo.

El primer paso del algoritmo está relacionado con el plano de rotaciones. Específicamente, las rotaciones $(i, i+1)$ eliminan los elementos en $P^T B$.

Antes que la k-ésima columna sea procesada, la columna que tiene la mayor norma-2 de la submatriz es intercambiada con la columna k-ésima.

En el segundo paso, las rotaciones generadas en la primera parte son aplicadas a R_A. Cuando la rotación en $(i, i+1)$ es postmultiplicado sobre R_A, crea un elemento no nulo en la posición $(i+1, i)$. Este elemento es eliminado al ser premultiplicado por la rotación $(i, i+1)$.

Este proceso es ilustrado en la fig. 1, aquí las flechas indican las rotaciones, y el sombrero indica el elemento que será eliminado.

Las rotaciones del primer paso pueden ser aplicadas a R_A tan pronto como son generadas, las cuales son almacenadas. De modo semejante las rotaciones que forman

U pueden estar acumuladas en Q_A, con el fin que Q_A sea transformada en Q_C. Así se tiene los pasos uno, dos y tres de la implementación están completamente interelacionados.

Asimismo cuando un producto de la forma $A_1 A_2 \ldots A_m$ debe ser procesado, el procedimiento puede ser iniciado definiendo $A = Q_A = P_A = I$ y $B = A_1$. En este caso el segundo paso deberá ser obviado y las rotaciones del primer paso acumuladas en Q_A.

En este algoritmo el número de operaciones es aproximadamente $5\frac{1}{6}n^3$ de adiciones y $10\frac{1}{6}n^3$ aproximadamente de multiplicaciones. Cuando n es suficientemente grande, las rotaciones de la transformación rápida de Givens pueden ser usadas para reducir el número de multiplicaciones aproximadamente $5\frac{1}{6}n^3$.

$$
\begin{array}{cccc}
 & & \downarrow & \downarrow \\
r & r & r & r \\
0 & r & r & r \\
0 & 0 & r & r \\
0 & 0 & 0 & r
\end{array}
\Longrightarrow
\begin{array}{ccccc}
 & & & & \\
 & r & r & r & r \\
 & 0 & r & r & r \\
\rightarrow & 0 & 0 & r & r \\
\rightarrow & 0 & 0 & \hat{r} & r
\end{array}
\Longrightarrow
\begin{array}{cccc}
 & \downarrow & \downarrow & \\
r & r & r & r \\
0 & r & r & r \\
0 & 0 & r & r \\
0 & 0 & 0 & r
\end{array}
\Longrightarrow
\begin{array}{ccccc}
 & & & & \\
 & r & r & r & r \\
\rightarrow & 0 & r & r & r \\
\rightarrow & 0 & \hat{r} & r & r \\
 & 0 & 0 & 0 & r
\end{array}
\Longrightarrow
$$

$$
\begin{array}{cccc}
\downarrow & \downarrow & & \\
r & r & r & r \\
0 & r & r & r \\
0 & 0 & r & r \\
0 & 0 & 0 & r
\end{array}
\Longrightarrow
\begin{array}{ccccc}
 & & & & \\
\rightarrow & r & r & r & r \\
\rightarrow & \hat{r} & r & r & r \\
 & 0 & 0 & r & r \\
 & 0 & 0 & 0 & r
\end{array}
\Longrightarrow
\begin{array}{cccc}
 & & & \\
r & r & r & r \\
0 & r & r & r \\
0 & 0 & r & r \\
0 & 0 & 0 & r
\end{array}
$$

Fig. 1: Reducción de R_A

Capítulo 4

Conclusiones

Este algoritmo se recomienda por su simplicidad y su eficiencia. Esta eficiencia esta comprobada por la cantidad de operaciones.

Sin embargo no se puede garantizar el funcionamiento de este algoritmo.

Cabe resaltar que dicho algoritmo puede se usado para el producto de descomposición de valores singulares(PSVD).

Bibliografía

[1] Barbolla, Rosa y Paloma Sanz. Algebra Lineal y Teoría de Matrices, Prentice Hall, 1998.

[2] Datta, B.N. Numerical Linear Algebra and Applications. Editorial Brooks/Cole Publishing Company, 1995.

[3] Golub, Gene H.y Charles F. Van Loan, Matrix Computations, John Hopkins University Press, 1993.

[4] Stewart, Gilbert W, Matrix Algorithms. Volume I: Basic Decompositions. SIAM, Philadelphia, 1998.