Conceptos previos

Definición: un conjunto de vectores $\{q_1, q_2, \cdots q_m\}$ donde $q_i \in I\!R^n$ es un conjunto ortonormal si satisface

$$< q_i, q_j >= \begin{cases} 0 & \text{si} \quad i \neq j \\ 1 & \text{si} \quad i = j \end{cases}$$

Definición: Sea $Q \in I\!R^{n\times n}$. Se dice que Q es ortogonal si sus columnas forman un conjunto ortonormal. Observe que necesariamente esta matriz satisface que $Q^{-1} = Q^T$.

Propiedades:

1. $Q^T Q = QQ^T = I$
2. $\|Qx\|_2 = \|x\|_2, \forall x \in I\!R^n$
3. El producto de matrices ortogonales es una matriz ortogonal.

Factorización QR

La idea es, dada $A \in I\!R^{n\times n}$, hallar una matriz ortogonal Q y una matriz R triangular superior, tal que $A = QR$. Al igual que el proceso de eliminación Gaussiana, la matriz A se multiplicará por matrices ortogonales hasta obtener la matriz R, esto es

$$\underbrace{Q_{n-1}Q_{n-2}\cdots Q_2Q_1}_{Q^{-1}} A = R \tag{1}$$

$$Q^{-1}A = R \tag{2}$$

$$QQ^{-1}A = QR \tag{3}$$

$$A = QR \tag{4}$$

Note que si $Q^{-1} = Q_{n-1}Q_{n-2}\cdots Q_2Q_1$ entonces

$$Q = (Q_{n-1}Q_{n-2}\cdots Q_2Q_1)^{-1} = Q_1^{-1}Q_2^{-1}\cdots Q_{n-2}^{-1}Q_{n-1}^{-1} = Q_1^T Q_2^T \cdots Q_{n-2}^T Q_{n-1}^T$$

Esquema de la factorización QR

Primer paso: Es necesario hallar Q_1 tal que

$$Q_1 A = \begin{bmatrix} \alpha_1 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \dots & * \end{bmatrix} = A^{(1)}$$

Segundo paso: Es necesario hallar Q_2 tal que

$$Q_2 A^{(1)} = Q_2 Q_1 A = \begin{bmatrix} \alpha_1 & * & * & \dots & * \\ 0 & \alpha_2 & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix} = A^{(2)}$$

siguiendo con esta metodología es posible hallar $n-1$ matrices ortogonales tales que

$$Q_{n-1} \cdots Q_2 Q_1 A = \begin{bmatrix} \alpha_1 & * & * & \dots & * \\ 0 & \alpha_2 & * & \dots & * \\ 0 & 0 & \alpha_3 & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * \end{bmatrix} = A^{(n-1)} = R$$

Transformada de Householder

Una forma de obtener las matrices ortogonales necesarias para obtener la factorización QR de una matriz dada $A \in I\!R^{n \times n}$ es a través del uso de las matrices de Householder.
Definición: Dado $u \in I\!R^n$ una matriz de Householder, denotada por H, se define como

$$H = I_n - \frac{2uu^t}{u^t u},$$

donde I_n es la matriz identidad de orden n. Las matrices de Householder también son conocidas con el nombre de **Reflectores de Householder.** Para ilustrar este hecho considere la Figura 1. En esta Figura, H representa una matriz de Householder generada por el vector $u \in I\!R^3$ y P es el subespacio conformado por todos los vectores ortogonales al vector u es decir: $P = \{y \in I\!R^3 : u^t y = 0\}$ Al aplicar la matriz H sobre algún vector x se

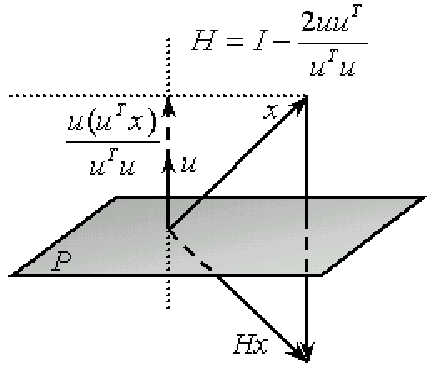


Figura 1: Reflector de Householder

obtiene un vector $y = Hx$ que es el **reflejo** de x a través de P.

<u>Propiedades:</u>

1. H es ortogonal.
2. H es simétrica
3. $I = HH^{-1} = HH^T = HH = H^2$.
4. $Hu = -u$
5. $Hy = y$, para todo $y \in P$
6. $\|Hx\|_2 = \|x\|_2$

7. Producto matriz vector

$$Hx = \left(I_n - \frac{2uu^t}{u^t u}\right) x = x - \frac{2uu^t x}{u^t u} = x - \left(\frac{2u^t x}{u^t u}\right) u = x - \beta u,$$

donde $\beta = \frac{2u^t x}{u^t u}$

A continuación se enuncia un Lema que será de utilidad para emplear las matrices de Householder para obtener la factorización QR de una matriz dada. **NOTA**: La última página de este material contiene una explicación gráfica de este Lema

Lema: Sea $x, y \in I\!R^n$, tal que $\|x\| = \|y\|$. El vector $u = x - y$ genera una matriz de Householder tal que $Hx = y$.

Matrices de Householder y Factorización QR

Las matrices ortogonales Q_i con $i = 1, 2, \cdots (n-1)$ requeridas en la expresión (1) serán matrices de Householder. Es por ello que en esta sección la ecuación (1) se reescribirá como

$$H_{n-1} H_{n-2} \cdots H_2 H_1 A = R. \tag{5}$$

Como se mencionó anteriormente, es fácil ver que este proceso de factorización, requiere de $n-1$ pasos para completarse.

Primer paso:

Es necesario hallar H_1 tal que

$$H_1 A = H_1 [a_1 \; a_1 \; \cdots \; a_n] = [H_1 a_1 \; H_1 a_2 \; \cdots \; H_1 a_n] = \begin{bmatrix} \alpha_1 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \dots & * \end{bmatrix} = A^{(1)}$$

donde a_i con $i = 1, 2, \cdots n$ representan las columnas de la matriz A.

> **NOTA:** Voy a denotar por $\widetilde{A}^{(1)} \in I\!R^{(n-1)\times(n-1)}$ a aquella submatriz de $A^{(1)}$ que no contiene ni la primera fila ni la primera columna de $A^{(1)}$. Esto será de utilidad en el segundo paso.

Observe que H_1 debe ser construida con un objetivo muy especial, H_1 debe ser tal que

$$H_1 a_1 = \begin{bmatrix} \alpha_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times n}$$

o lo que es lo mismo H_1 debe ser tal que

$$H_1 a_1 = \alpha_1 e_1, \tag{6}$$

donde e_1 es el primer canónico de tamaño n.

Observe que para hallar H_1 sólo se require del vector u_1 que la define, ya que $H_1 = I - \frac{2u_1u_1^t}{u_1^t u_1}$. Ahora bien, teniendo en cuenta **el Lema visto en la sección anterior** es posible forzar a que $\alpha_1 \in I\!R$ sea tal que $\|a_1\| = \|\alpha_1 e_1\|$ y luego, basados en dicho Lema, definir $u_1 = a_1 - \alpha_1 e_1$. Para hacer la discusión más general (y menos engorroza), no utilizaré subíndices, es decir, $H_1 = H$, $a_1 = a$, $u_1 = u$ y $\alpha_1 = \alpha$. Con este cambio de notación la ecuación (6) se reescribe como

$$Ha = \alpha e_1. \tag{7}$$

En el siguiente recuadro, se explicará cómo aplicar el Lema para obtener la matriz H (vector u que la genera) que logra que la ecuación (7) se satisfaga.

Como se dijo anteriormente si α es tal que,

$$\|a\| = \|\alpha e_1\| = |\alpha|\|e_1\| = |\alpha|, \tag{8}$$

entonces $u = a - \alpha e_1$. De (8) se desprende que α bien puede ser igual a $\|a\|$ o igual a $-\|a\|$, con lo cual el vector u puede definirse de dos formas

$$u = a - \|a\|e_1 \quad \text{si se usa} \quad \alpha = \|a\| \tag{9}$$
$$u = a + \|a\|e_1 \quad \text{si se usa} \quad \alpha = -\|a\|. \tag{10}$$

Note que si la primera componente de a es negativa conviene usar el u definido por (9) para asíevitar una resta que necesariamente ocurriría si se usa la ecuación (10) para obtener u. Mientras que si la primera componente de a es positiva conviene usar el u definido por (10), pues si se usase (9) estaríamos realizando una resta: Recuerde que siempre es una buena práctica evitar las restas.
De lo anterior se concluye que, dependiendo del signo de la primera componente de a se puede escoger una expresión para u que evite las restas. Por lo tanto

$$u = a + signo(a_1)\|a\|e_1, \tag{11}$$

donde a_1 es la primera componente del vector a.

Segundo paso:

Es necesario hallar H_2 tal que

$$H_2 \underbrace{A^{(1)}}_{paso1} = H_2[\overline{a}_1\ \overline{a}_2\ \cdots\ \overline{a}_n] = [H_2\overline{a}_1\ H_2\overline{a}_2\ \cdots\ H_2\overline{a}_n] = \begin{bmatrix} \alpha_1 & * & * & \dots & * \\ 0 & \alpha_2 & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix} = A^{(2)}$$

donde $\overline{a}_i$ con $i = 1, 2, \cdots n$ representan las columnas de la matriz $A^{(1)}$. Note que $\overline{a}_1 = (\alpha_1, 0, \cdots, 0)^t$, con lo cual H_2 no debe modificar la primera columna de $A^{(1)}$ para no perder los ceros colocados en el primer paso. Más aún, si H_2 tampoco modifica la primera fila de $A^{(1)}$, entonces es posible hallar $\widetilde{H}_2 \in I\!R^{(n-1)\times(n-1)}$ tal que

$$\widetilde{H}_2 \widetilde{a}_1 = \begin{bmatrix} \alpha_2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(n-1)\times(n-1)} ,$$

donde $\widetilde{a}_1$ es la primera columna de $\widetilde{A}^{(1)}$ (ver **NOTA** escrita en rojo en el Primer paso). Dicho de otra forma, es posible usar el procedimiento escrito en el **recuadro del primer paso** para hallar $\widetilde{H}_2$ tal que

$$\widetilde{H}_2 \widetilde{a}_1 = \alpha_2 e_1, \tag{12}$$

donde e_1 es el primer canónico de tamaño $n-1$. Finalmente la matriz H_2 buscada es de la forma

$$H_2 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & \widetilde{H}_2 & & \\ & & & & \\ 0 & & & & \end{bmatrix} = A^{(2)}$$

NOTA: Siguiendo con la notación, la matriz $\widetilde{A}^{(2)} \in I\!R^{(n-2)\times(n-2)}$ será aquella submatriz de $A^{(2)}$ que no contiene ni las dos primeras filas ni las dos primeras columna de $A^{(2)}$.

Paso $n-1$:

La matriz proveniente del paso $n-2$, es decir, $A^{(n-2)}$ tiene la forma

$$A^{(n-2)} = \begin{bmatrix} \alpha_1 & * & * & \dots & * & * & * \\ 0 & \alpha_2 & * & \dots & * & * & * \\ 0 & 0 & \alpha_3 & \dots & * & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{n-2} & * & * \\ 0 & 0 & 0 & \dots & 0 & * & * \\ 0 & 0 & 0 & \dots & 0 & * & * \end{bmatrix},$$

además se tiene a la matriz $\widetilde{A}^{(n-2)} \in I\!\!R^{2\times 2}$ que es aquella submatriz de $A^{(n-2)}$ que no posee las $(n-2)$ primeras filas ni las dos $(n-2)$ primeras columnas de $A^{(n-2)}$. El objetivo es hallar una matriz de Householder H_{n-1} tal que

$$H_{n-1}A^{(n-2)} = \begin{bmatrix} \alpha_1 & * & * & \dots & * & * & * \\ 0 & \alpha_2 & * & \dots & * & * & * \\ 0 & 0 & \alpha_3 & \dots & * & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{n-2} & * & * \\ 0 & 0 & 0 & \dots & 0 & \alpha_{n-1} & * \\ 0 & 0 & 0 & \dots & 0 & 0 & * \end{bmatrix} = A^{(n-1)} = R.$$

Siguiendo con la filosofía de que H_{n-1} no debe modificar los ceros colocados en los $(n-2)$ pasos anteriores, H_{n-1} debe ser de la forma

$$H_{n-1} = \begin{bmatrix} I_{(n-2)} & 0_{(n-2)\times 2} \\ 0_{2\times(n-2)} & \widetilde{H}_{n-1} \end{bmatrix},$$

donde $\widetilde{H}_{n-1} \in I\!\!R^{2\times 2}$ es tal que

$$\widetilde{H}_{n-1}\widetilde{a}_1 = \alpha_{n-1}e_1, \tag{13}$$

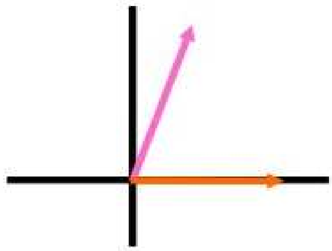
siendo $\widetilde{a}_1$ la primera columna de $\widetilde{A}^{(n-2)}$. Note que $\widetilde{a}_1$ es un vector de dos componentes.

Al finalizar este paso, y recordando lo discutido en relación a la factorización QR se tiene que

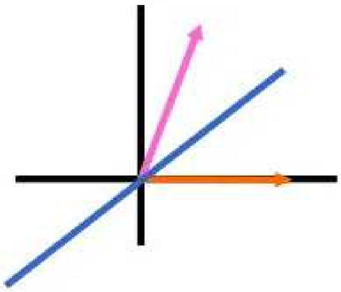
$$\begin{aligned}
\underbrace{H_{n-1}H_{n-2}\cdots H_2H_1}_{Q^{-1}}A &= R \\
Q^{-1}A &= R \\
A &= QR
\end{aligned}$$

Note que $Q^{-1} = H_{n-1}H_{n-2}\cdots H_2H_1$ es una matriz ortogonal pues ella es el producto de $n-1$ matrices ortogonales de Householder. Más aún

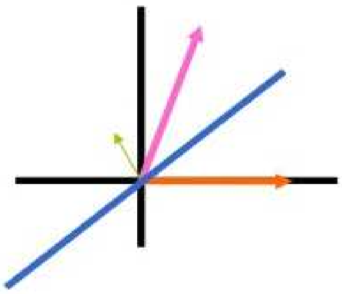
$$\begin{aligned}
Q &= (H_{n-1}H_{n-2}\cdots H_2H_1)^{-1} \\
&= H_1^{-1}H_2^{-1}\cdots H_{n-2}^{-1}H_{n-1}^{-1} \\
&= H_1^{T}H_2^{T}\cdots H_{n-2}^{T}H_{n-1}^{T} \\
&= H_1H_2\cdots H_{n-2}H_{n-1} \quad \text{ya que cada } H_i \text{ es simétrica}
\end{aligned}$$



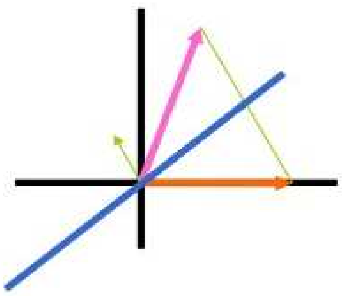
Inicialmente se tienen dos vectores x e y de igual tamaño, es decir $\|x\| = \|y\|$



Dados esos vectores es posible definir un plano P (plano azul) tal que x sea el reflejo de y a través de P (y viceversa).



Una vez definido el plano P, se puede encontrar un vector u (vector verde) ortogonal al plano P.



Note que el vector u de la figura anterior no es más que la traslación del vector $x-y$ (línea verde) al origen. Por o tanto $u = x - y$, y con este vector se puede hallar una matriz de Householder tal que $Hx = y$