Apéndice D

Procedimientos de cálculo.

D.1. Introducción

La resolución de las ecuaciones normales,

$$(X'X)\vec{\beta} = X'\vec{Y}$$

requiere, en su aproximación más directa, la obtención de la inversa (ordinaria o generalizada) de $(X'X)$. Hay procedimientos mucho menos costosos desde el punto de vista del cálculo que, además, permiten en algunos casos intuiciones interesantes y demostraciones de gran simplicidad.

En lo que sigue se presenta uno de los métodos de cálculo más utilizados, y la construcción en que se basa (la *factorización QR*). Se detalla también la correspondencia entre la notación empleada y los resultados de algunas funciones de `S` que hacen uso de dicha factorización.

D.2. Transformaciones ortogonales.

Sea el problema,

$$\min_{\vec{x}} ||D\vec{x} - \vec{c}||^2 \tag{D.1}$$

Podemos ver el problema como el de encontrar la combinación lineal de las columnas de D que mejor aproxima $\vec{c}$, en términos de norma de la discrepancia. Dicho problema queda inalterado cuando realizamos una misma transformación ortogonal de las columnas de D y del vector $\vec{c}$. En efecto,

$$\begin{aligned}
\min_{\vec{x}} ||Q(D\vec{x} - \vec{c})||^2 &= \min_{\vec{x}} < Q(D\vec{x} - \vec{c}), Q(D\vec{x} - \vec{c}) > \\
&= \min_{\vec{x}} (D\vec{x} - \vec{c})'Q'Q(D\vec{x} - \vec{c}) \\
&= \min_{\vec{x}} ||D\vec{x} - \vec{c}||^2
\end{aligned}$$

al ser Q ortogonal.

Definición D.1 *Sea D una matriz de orden $n \times m$. Supongamos que puede expresarse del siguiente modo:*

$$D = HRK'$$

en que:

(i) H es $n \times n$ y ortogonal.

(ii) R es $n \times m$ de la forma,

$$\begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

con R_{11} cuadrada de rango completo $k \leq$ mín(m, n).

(iii) K es $m \times m$ ortogonal.

Se dice que HRK' es una descomposición ortogonal de D.

En general, hay más de una descomposición ortogonal, dependiendo de la estructura que quiera imponerse a R. Si requerimos que R sea diagonal, tenemos la *descomposición en valores singulares.* Podemos también requerir que R sea triangular superior, o triangular inferior, obteniendo diferentes descomposiciones de D.

La elección de una descomposición ortogonal adecuada simplifica enormemente la solución de (D.1). Los resultados fundamentales vienen recogidos en el siguiente teorema.

Teorema D.1 *Sea D una matriz de orden $n \times m$ y rango k, admitiendo la descomposición ortogonal,*

$$D = HRK'. \tag{D.2}$$

Sea el problema

$$\min_{\vec{x}} ||D\vec{x} - \vec{y}\,||^2 \tag{D.3}$$

y definamos,

$$\begin{aligned} H'\vec{y} &= \vec{g} = \begin{pmatrix} \vec{g}_1 \\ \vec{g}_2 \end{pmatrix} \begin{matrix} k \\ n-k \end{matrix} \\ K'\vec{x} &= \vec{\gamma} = \begin{pmatrix} \vec{\gamma}_1 \\ \vec{\gamma}_2 \end{pmatrix} \begin{matrix} k \\ m-k \end{matrix}. \end{aligned}$$

Sea $\tilde{\gamma}_1$ *la solución (única) del sistema,*

$$R_{11}\tilde{\gamma}_1 = \vec{g}_1.$$

Entonces, todas las posibles soluciones del problema (D.3) son de la forma

$$\vec{x} = K \begin{pmatrix} \tilde{\gamma}_1 \\ \vec{\gamma}_2 \end{pmatrix},$$

con γ_2 *arbitrario. Cualquiera de esas soluciones da lugar al vector de residuos*

$$\vec{r} = \vec{y} - D\vec{x} = H \begin{pmatrix} \vec{0} \\ \vec{g}_2 \end{pmatrix}$$

y en consecuencia, $||\vec{r}\,|| = ||\vec{g}_2||$.

Existe un resultado interesante que muestra cómo es posible encontrar una transformación ortogonal que rota (y quizá refleja) un vector $\vec{v}$ hasta abatirlo sobre el subespacio generado por otro, $\vec{e}_1$. Se denomina *transformación de Householder*, y se obtiene de manera muy cómoda y simple como muestra el teorema siguiente.

Teorema D.2 *Sea* $\vec{v}$ *cualquier vector* $m \times 1$ *distinto de* $\vec{0}$. *Existe una matriz ortogonal* P $m \times m$ *tal que:*

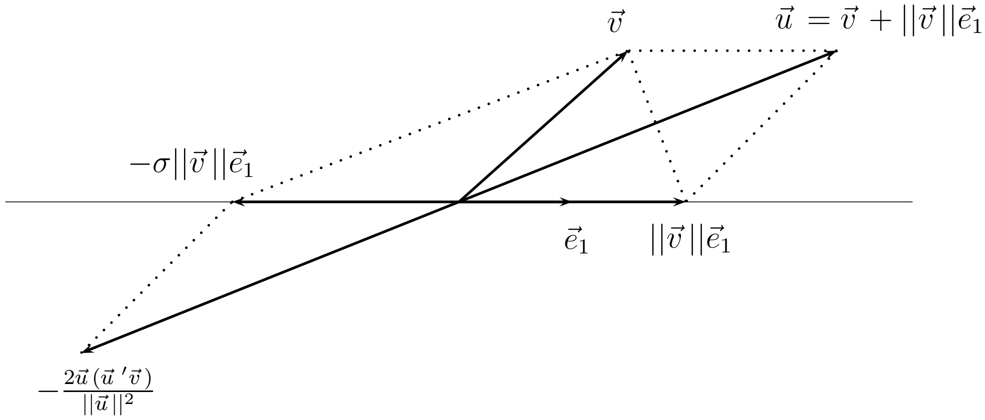
$$P\vec{v} = -\sigma||\vec{v}\,||\vec{e}_1 \tag{D.4}$$

siendo

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{D.5}$$

$$\sigma = \begin{cases} +1 & si\ v_1 \geq 0 \\ -1 & si\ v_1 < 0. \end{cases} \tag{D.6}$$

Figura D.1: Visualización de la transformación de Householder.



Esta matriz tiene por expresión,

$$P = I - 2\frac{\vec{u}\,\vec{u}\,'}{||\vec{u}||^2} \tag{D.7}$$

con $\vec{u} = \vec{v} + \sigma||\vec{v}||\vec{e}_1$.

DEMOSTRACIÓN:

Entonces (ver Figura D.1),

$$\vec{u} = \vec{v} + \sigma||\vec{v}||\vec{e}_1 \tag{D.8}$$
$$\vec{z} = \vec{v} - \sigma||\vec{v}||\vec{e}_1 \tag{D.9}$$

son ortogonales y $\vec{v} = \frac{1}{2}\vec{u} + \frac{1}{2}\vec{z}$. Tenemos en consecuencia,

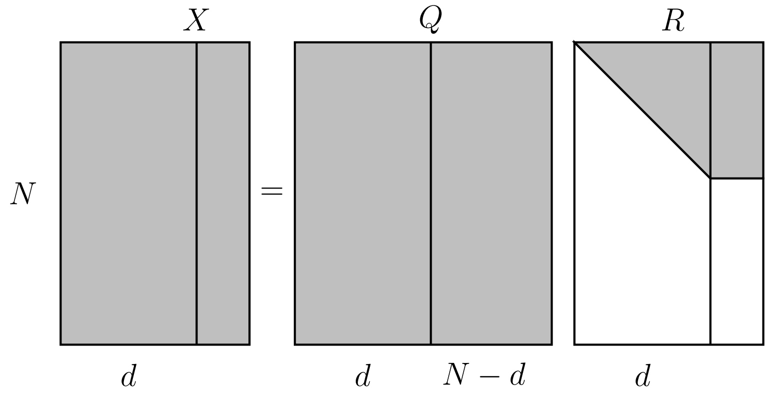
$$\begin{aligned}
P\vec{v} &= \left(I - 2\frac{\vec{u}\,\vec{u}\,'}{||\vec{u}||^2}\right)\left(\frac{1}{2}\vec{u} + \frac{1}{2}\vec{z}\right) && \text{(D.10)}\\
&= \frac{1}{2}\vec{u} - \vec{u} + \frac{1}{2}\vec{z} && \text{(D.11)}\\
&= -\frac{1}{2}\vec{u} + \vec{v} - \frac{1}{2}\vec{u} && \text{(D.12)}\\
&= \vec{v} - \vec{u} && \text{(D.13)}\\
&= -\sigma||\vec{v}||\vec{e}_1 && \text{(D.14)}
\end{aligned}$$

D.3. Factorización QR.

Teorema D.3 *Sea una matriz X de orden $(N \times p)$ y rango $d \leq \text{mín}(N, p)$. Existe siempre una matriz ortogonal Q de orden $(N \times N)$ y una matriz R trapezoidal superior verificando:*

$$X = QR \tag{D.15}$$

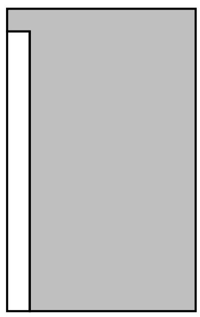
Esquemáticamente,



Demostración:

La prueba es constructiva, y reposa en la aplicación reiterada de la transformación de Householder a las columna de la matriz X. Sea $\vec{x}_1$ la primera de dichas columnas. Existe una transformación de Householder, de matriz ortogonal P_1 que abate dicha primera columna sobre el $\vec{e}_1$ de la base canónica de R^n. Es decir,

$P_1X =$

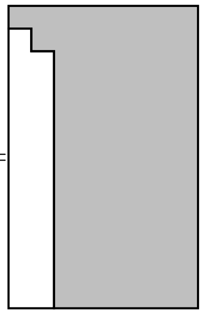


Llamemos X_1 a la matriz así obtenida, y consideremos su segunda columna eliminado su primer elemento. Los restantes, pueden verse como un vector en R^{N-1}, que puede tambien abatirse sobre el primer vector $\vec{e}_1$ de la base canónica de dicho subespacio multiplicando por una matriz de Householder P_2^*. Entonces,

$$\begin{pmatrix} 1 & \vec{0}\,' \\ \vec{0} & P_2^* \end{pmatrix} P_1 \tag{D.16}$$

reduce la matriz X de la forma que esquemáticamente se muestra a continuación:

$$\begin{pmatrix} 1 & \vec{0}\,' \\ \vec{0} & P_2^* \end{pmatrix} P_1 X =$$



Por consiguiente, si llamamos

$$P_2 = \begin{pmatrix} 1 & \vec{0}\,' \\ \vec{0} & P_2^* \end{pmatrix}$$

el producto P_2P_1 reduce las dos primeras columnas de X a forma escalonada. Como tanto P_1 como P_2 son ortogonales, su producto también lo es. Fácilmente se comprueba que el proceso puede continuarse hasta obtener un producto de matrices ortogonales $Q\,' = P_d P_{d-1} \dots P_1$ que deja X con sus d primeras columnas "escalonadas". Además, como el rango de X era d, necesariamente las últimas $N - d$ filas de R son de ceros.

En definitiva, $Q\,'X = R$ y por tanto $X = QR$, lo que prueba el teorema.

D.4. Bibliografía

Hay abundante literatura sobre la factorización QR y procedimientos similares de aplicación al problema (D.1). Casi cualquier texto de Cálculo Numérico contiene una discusión de la factorización QR. Una referencia fundamental que continúa vigente es Lawson and Hanson (1974). Una exposición breve, clara, y con abundantes referencias a la literatura más reciente puede encontrarse en Goodhall (1993). Ansley (1985) muestra como, al margen y además de su utilidad como procedimiento numérico, la factorización QR arroja luz sobre, y simplifica la demostración de, bastantes resultados en regresión lineal.