CURSO DE METODOS NUMERICOS

CUARTA PARTE

METODOS DE MINIMO CUADRADOS

CAPITULO XIX. EL PROBLEMA DE LOS MINIMOS CUADRADOS: PRELIMINARES

1. SISTEMAS LINEALES DE ECUACIONES SOBREDETERMINADOS

La discusión de los problemas algebráicos de la parte anterior se había centrado exclusivamente en la ecuación $A\,\mathbf{x} = \mathbf{b}$, en la cual se ha intentado hacer coincidir exactamente el vector $\mathbf{b}$ que aparece a la derecha con el vector $A\,\mathbf{x}$ que aparece a la izquierda, es decir hacer que $\mathbf{b}$ sea exactamente una combinación lineal de las columnas de A. En muchos casos, el vector $\mathbf{b}$ no se puede expresar de la forma $A\,\mathbf{x}$, y el sistema $A\,\mathbf{x} = \mathbf{b}$ es **incompatible**. El problema lineal de mínimos cuadrados se acerca a la optimización si se encuentra un vector $\mathbf{x}$ que en algún sentido da la **mejor** (aunque no la perfecta) aproximación entre $A\,\mathbf{x}$ y $\mathbf{b}$.

Considérese un sistema de m ecuaciones con n incógnitas escrito en la forma

$$A\,\mathbf{x} = \mathbf{b}\ . \tag{XIX.1}$$

Aquí, A es de $m \times n$, $\mathbf{x}$ es de $n \times 1$ y $\mathbf{b}$ es de $m \times 1$. Supondremos que el rango de A es n, de lo que resulta que $m \geq n$. Si $m > n$, se dice que el sistema lineal de ecuaciones es **sobredeterminado**. Los sistemas sobredeterminados son generalmente no compatibles.

Ejemplo.

a) Sea dado el sistema de ecuaciones lineales

$$\begin{array}{rcrcr} x_1 & + & x_2 & = & 1\ , \\ x_1 & - & x_2 & = & 3\ , \\ -\,x_1 & + & 2\,x_2 & = & -2\ . \end{array}$$

Las tres ecuaciones del sistema representano rectas en el plano. Las primeras dos se intersecan en el punto $(2, -1)$, mientras que la tercera no pasa por este punto. Entonces, no existe ningun punto común entre las tres líneas. El sistema es incompatible.

b) Sea dado el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & + & x_3 & = & 1\ , \\ 2\,x_1 & - & x_2 & + & x_3 & = & 2\ , \\ 4\,x_1 & + & 3\,x_2 & + & 3\,x_3 & = & 4\ , \\ 2\,x_1 & - & x_2 & + & 3\,x_3 & = & 5\ , \end{array}$$

Usando eliminación Gaussiana se deduce que el sistema tiene exactamente una única solución $(0.1, -0.3, 1.5)$. El sistema es compatible determinado.

c) Sea dado el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & + & x_3 & = & 1\ , \\ 2\,x_1 & - & x_2 & + & x_3 & = & 2\ , \\ 4\,x_1 & + & 3\,x_2 & + & 3\,x_3 & = & 4\ , \\ 3\,x_1 & + & x_2 & + & 2\,x_3 & = & 3\ , \end{array}$$

Usando eliminación Gaussiana y resolviendo x_2 y x_1 en función de x_3 se deduce que el sistema tiene una infinidad de soluciones de la forma $(1 - 0.6\alpha, -0.2\alpha, \alpha)$. El sistema es compatible indeterminado.

2. EL VECTOR RESIDUAL Y EL PROBLEMA DE LOS MINIMOS CUADRADOS

Por lo general, el sistema $(XIX.1)$ no tiene solución debido a que $\mathbf{b}$ no pertenece al subespacio de $\mathcal{R}^n$ de dimensión n, generado por las columnas de A. Es frecuente en tales casos que se requiera encontrar un $\mathbf{x}$ tal que $\mathbf{b}$ esté cerca de $A\ \mathbf{x}$, es decir que minimice una norma del vector residual $\mathbf{r} = \mathbf{b} - A\ \mathbf{x}$. El primer problema será determinar el sentido de cercanía, es decir la norma que mide la diferencia entre $\mathbf{b}$ y $A\ \mathbf{x}$. La interpretación y la naturaleza de la solución de este problema varía dependiendo de la norma que se use. La norma más comúnmente usada es la norma euclidiana, la norma l_2. Entonces, la *solución* en mínimos cuadrados de $(XIX.1)$ es el vector $\mathbf{x}$ que hace de $||\mathbf{r}||_2 = ||\mathbf{b} - A\ \mathbf{x}||_2$ un mínimo

$$\min_{\mathbf{x}\in\mathcal{R}^n} ||\mathbf{b} - A\ \mathbf{x}||_2^2\ , \tag{XIX.2}$$

Dado que las normas no pueden ser negativas, minimizar una norma es equivalente a minimizar su cuadrado; hableremos entonces libremente de normas cuadradas y no. Según lo que se ha supuesto acerca del rango de A, la solución $\mathbf{x}$ será única.

La formulación de los mínimos cuadrados es popular sobre todo por el hecho de que resolver un problema de mínimos cuadrados lineal es más fácil que minimizar otras normas de $\mathbf{r}$. Veremos que el problema de mínimos cuadrados se puede resolver transformando la matriz A y resolviendo un sistema compatible a ella relacionado. Al contrario, minimizar la norma l_1 o la norma l_∞ es equivalente a un problema de programación lineal no diferenciable, cuya solución necesita una técnica de iteración.

Es útil pensar en el problema de los mínimos cudrados en términos de los subespacios definidos por la matriz A. El subespacio rango de A, rango(A), consiste en todos los vectores m-dimensionales que son combinaciones lineales de las columnas de A y el subespacio complementario, el subespacio nulo de A^t, nulo(A), contiene todos los vectores m-dimensionales que son $\mathbf{z}$-ortogonales a las columnas de A, es decir tales que $A^t\ \mathbf{z} = 0$. Si r denota el rango de la matriz A, entonces r es la dimensión del subespacio rango de A y $(m - r)$ es la dimensión del subespacio nulo de A^t.

Dada una matriz no nula A, cualquier vector m-dimensional $\mathbf{c}$ se puede expresar como la suma de un vector $\mathbf{c_R}$ en el rango de A, $\mathbf{c_R} \in$ rango(A), y de un vector $\mathbf{c_N}$ en el espacio nulo de A^t, $\mathbf{c_N} \in$ nulo(A^t), es decir

$$\mathbf{c} = \mathbf{c_R} + \mathbf{c_N}\ . \tag{XIX.3}$$

Los vectores $\mathbf{c_R}$ y $\mathbf{c_N}$ son únicos y satisfacen

$$\mathbf{c_R} = A\ \mathbf{c_A}\ , \qquad A^t\ \mathbf{c_N} = 0\ , \qquad \mathbf{c_R}^t \cdot \mathbf{c_N} = 0\ , \tag{XIX.4}$$

donde $\mathbf{c_A}$ se refiere a un vector n-dimensional cualquiera tal que $\mathbf{c_R} = A\ \mathbf{c_A}$. La unicidad de $\mathbf{c_R}$ y $\mathbf{c_N}$ implica que las componentes en el espacio rango y en el espacio nulo de vectores iguales tienen que ser iguales. En efecto,

$$\mathbf{c} = \mathbf{d} \quad \text{si y sólo si} \quad \mathbf{c_R} = \mathbf{d_R} \quad \text{y} \quad \mathbf{c_N} = \mathbf{d_N} \ . \qquad (XIX.5)$$

Aunque $\mathbf{c_R}$ es único, el vector $\mathbf{c_A}$ es único sólo si las columnas de A son linealmente independientes.

Debido al hecho de que la norma euclídea está definida en términos del producto escalar, las relaciones $(XIX.3)$ y $(XIX.4)$ tienen una consecuencia muy importante:

$$||\mathbf{c}||_2^2 = \mathbf{c^t} \cdot \mathbf{c} = ||\mathbf{c_R}||_2^2 + ||\mathbf{c_N}||_2^2 \ , \qquad (XIX.6)$$

de manera que la representación de un vector en términos de sus componentes en el espacio rango y nulo separa el cuadrado de la norma euclídea en dos partes independientes. Esta propiedad no vale para otras normas.

Estas observaciones son relevantes para el problema de los mínimos cuadrados dado que permiten determinar la más pequeña posible norma euclídea del vector residual $\mathbf{r} = \mathbf{b} - A\ \mathbf{x}$ para todos vectores $\mathbf{x}$. Dado que ambos $\mathbf{b}$ y el vector residual $\mathbf{r}$ son vectores m-dimensionales, ambos pueden escribirse en la representación $(XIX.3)$:

$$\begin{aligned} \mathbf{b} &= \mathbf{b_R} + \mathbf{b_N} \quad \text{con} \quad \mathbf{b_R} = A\ \mathbf{b_A} \ , \\ \mathbf{r} &= \mathbf{r_R} + \mathbf{r_N} \quad \text{con} \quad \mathbf{r_R} = A\ \mathbf{r_A} \ . \end{aligned} \qquad (XIX.7)$$

Combinando la definición de $\mathbf{r}$ como $\mathbf{b} - A\ \mathbf{x}$ con estas relaciones, se obtiene

$$\begin{aligned} \mathbf{r} = \mathbf{r_R} + \mathbf{r_N} &= \mathbf{b} - A\ \mathbf{x} = \mathbf{b_R} + \mathbf{b_N} - A\ \mathbf{x} \\ &= \mathbf{b_R} - A\ \mathbf{x} + \mathbf{b_N} \end{aligned} \qquad (XIX.8)$$

Un punto obvio pero crucial en esta expresión es que el vector $A\ \mathbf{x}$ está enteramente en el rango de A. Cuando se resta $A\ \mathbf{x}$ del vector $\mathbf{b}$ para crear el residual, sigue de $(XIX.5)$ que la componente en el espacio rango de $\mathbf{b} - A\ \mathbf{x}$ tiene que ser $\mathbf{b_R} - A\ \mathbf{x}$. En contraste, el restar $A\ \mathbf{x}$ no puede eliminar nada a la componente de b en el espacio nulo, de manera que la componente en el espacio nulo de $\mathbf{b} - A\ \mathbf{x}$ es igual a $\mathbf{b_N}$. Entonces, las componentes en el espacio rango y en el nulo del vector residual satisfacen

$$\mathbf{r_R} = \mathbf{b_R} - A\ \mathbf{x} \quad \text{y} \quad \mathbf{r_N} = \mathbf{b_N} \ . \qquad (XIX.9)$$

Como ya dicho, el problema de mínimos cuadrados consiste en minimizar la norma l_2 del vector residual. De $(XIX.6)$ y $(XIX.9)$ sigue que la norma euclídea del vector residual $\mathbf{r}$ para cualquier vector $\mathbf{x}$ satisface

$$||\mathbf{r}||_2^2 = ||\mathbf{b_R} - A\ \mathbf{x}||_2^2 + ||\mathbf{b_N}||_2^2 \geq ||\mathbf{b_N}||_2^2 \ . \qquad (XIX.10)$$

Dado que $\mathbf{b_N}$ se mantiene en el residual, se encontrará el mínimo de $||\mathbf{b} - A\ \mathbf{x}||_2$ cuando la norma de la componente en el subespacio rango, $||\mathbf{b_R} - A\ \mathbf{x}||_2$, sea la más pequeña posible. Además, por definición $\mathbf{b_R}$ está en el rango de A, luego tiene que existir un vector $\mathbf{x}$ tal que $A\ \mathbf{x} = \mathbf{b_R}$. Para este vector particular $\mathbf{x}$, se elimina la componente de $\mathbf{b}$ en el espacio rango cuando se le resta $A\ \mathbf{x}$, que significa que $\mathbf{b} - A\ \mathbf{x} = \mathbf{b_N}$ y que la norma euclídea del vector residual es igual a su cota inferior $||\mathbf{b_N}||_2$.

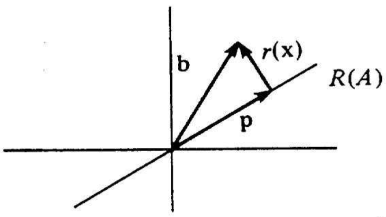
Escoger $\mathbf{x}$ de esta manera, $A\ \mathbf{x} = \mathbf{b_R}$, no sólo minimiza la norma del vector residual, sino que además fuerza al residual a estar enteramente en el espacio nulo de A^t. De esta manera se llega a dos caracterizaciones equivalentes de la solución optimal del problema de mínimos cuadrados:

$$\mathbf{x} \quad \text{minimiza} \quad ||\mathbf{b} - A\ \mathbf{x}||_2 \quad \Longleftrightarrow \quad A^t\ (\mathbf{b} - A\ \mathbf{x}) = \mathbf{0}\ , \qquad (XIX.11)$$

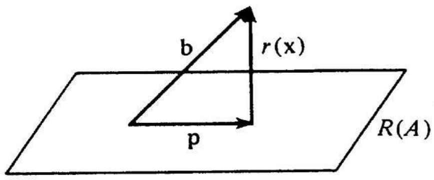
$$\mathbf{x} \quad \text{minimiza} \quad ||\mathbf{b} - A\ \mathbf{x}||_2 \quad \Longleftrightarrow \quad \begin{cases} A\ \mathbf{x} = \mathbf{b_R}\ , \\ \mathbf{b} - A\ \mathbf{x} = \mathbf{b_N}\ . \end{cases} \qquad (XIX.12)$$

Un problema de mínimos cuadrados se dice **compatible** si su vector residual optimal es cero, en este caso $\mathbf{b_N} = 0$ y $\mathbf{b} = \mathbf{b_R}$; si éste no es el caso se dice **incompatible**.

La ortogonalidad del vector residual respecto de las filas de la matriz A corresponde al principio geométrico familiar de que la distancia más corta de un punto a un plano es la longitud de la perpendicular que los une. Aquí el punto es el vector $\mathbf{b}$ y el plano es el subespacio rango de A.



(a) $\mathbf{b} \in R^2$ and A is a 2×1 matrix of rank 1.



(b) $\mathbf{b} \in R^3$ and A is a 3×2 matrix of rank 2.

La unicidad del vector $\mathbf{b_N}$ implica que el vector residual optimal para el problema de mínimos cuadrados es único. El vector $\mathbf{b_R}$ es también único, y el sistema $A\ \mathbf{x} = \mathbf{b_R}$ es compatible por definición. Sin embargo, el vector $\mathbf{x}$ que satisface $A\ \mathbf{x} = \mathbf{b_R}$ es único si y sólo si las columnas de la matriz A son linealmente independientes. Dado que cualquier vector $\mathbf{x}$ que satisface $A\ \mathbf{x} = \mathbf{b_R}$ es una solución de mínimos cuadrados, sigue una importante conclusión: **la solución del problema de mínimos cuadrados es única si y sólo si la matriz A es de rango n.**

Ejemplo. Sean

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix}$$

con las columnas de A linealmente independientes. Por definición, cualquier vector $\mathbf{y}$ en el rango de A tiene la forma

$$\mathbf{y} = A\mathbf{x} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 \\ x_2 \end{pmatrix} .$$

El espacio nulo de A^t consiste de vectores $\mathbf{z}$ que satisfacen

$$A^t\mathbf{z} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} z_1 + z_2 \\ z_1 + z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \ .$$

Las relaciones $z_1 + z_2 = 0$ y $z_1 + z_3 = 0$ implican que cualquier vector $\mathbf{z}$ del espacio nulo de A^t tiene que ser de la forma

$$\mathbf{z} = \begin{pmatrix} x_3 \\ -x_3 \\ -x_3 \end{pmatrix}$$

para algún escalar x_3. Entonces, para buscar la representación del vector $\mathbf{b}$ como suma de vectores en el subespacio rango de A y en el subespacio nulo de A^t tiene que ser

$$\mathbf{b} = \mathbf{y} + \mathbf{z} = \begin{pmatrix} x_1 + x_2 \\ x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ -x_3 \\ -x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix}$$

que nos da $x_1 = x_3 = 2$ y $x_2 = -3$.

Se deduce entonces, que la representación del vector $\mathbf{b}$ como suma de vectores en el subespacio rango de A y en el subespacio nulo de A^t viene dada por

$$\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix} = \mathbf{b_R} + \mathbf{b_N} = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ -2 \end{pmatrix} =$$

$$= A\ \mathbf{b_A} + \mathbf{b_N} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ -2 \end{pmatrix} \ .$$

Aplicando $(XIX.12)$, se obtiene que la solución de mínimos cuadrados y el vector residual optimal son

$$\mathbf{x} = \begin{pmatrix} 2 \\ -3 \end{pmatrix} \quad \text{y} \quad \mathbf{r} = \begin{pmatrix} 2 \\ -2 \\ -2 \end{pmatrix} \ .$$

En este ejemplo, la solución de mínimos cuadrados es única porque A es de rango lleno.

Un aspecto crucial de la relación $(XIX.12)$ es que revela cómo cambios en el vector $\mathbf{b}$ afectan a la solución $\mathbf{x}$ y al problema de mínimos cuadrados $\min ||\mathbf{b} - A\ \mathbf{x}||_2^2$. Para sistemas de ecuaciones lineales no singulares y compatibles, cambios en $\mathbf{b}$ implicaban encontrar una solución $\mathbf{x}$ diferente. Esta propiedad no se verifica para el problema de los mínimos cuadrados, cuyo analisis es más complicado.

Se supone por simplicidad que A tiene rango lleno, entonces $\mathbf{x}$ es la única solución del problema de mínimos cuadrados $\min ||\mathbf{b} - A\ \mathbf{x}||_2^2$ y $\mathbf{r}$ es el vector residual optimal, $\mathbf{r} = \mathbf{b} - A\ \mathbf{x}$. Si $\mathbf{b}$ viene perturbado por un vector que está enteramente en el subespacio rango de A, es decir $\tilde{\mathbf{b}} = \mathbf{b} + \delta\mathbf{b_R}$, donde $\delta\mathbf{b_R} = A\ \delta\mathbf{x}$ para un único n-vector $\delta\mathbf{x}$, entonces

$$\tilde{\mathbf{b}}_\mathbf{R} = \mathbf{b_R} + \delta\mathbf{b_R} = \mathbf{b_R} + A\ \delta\mathbf{x} \quad \text{y} \quad \tilde{\mathbf{b}}_\mathbf{N} = \mathbf{b_N} \ . \qquad (XIX.13)$$

Sigue entonces de $(XIX.12)$ que la solución de mínimos cuadrados $\tilde{\mathbf{x}}$ y el residual $\tilde{\mathbf{r}}$ correspondientes a $\tilde{\mathbf{b}}$ satisfacen

$$\tilde{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x} \quad \text{y} \quad \tilde{\mathbf{r}} = \mathbf{r} \ , \tag{XIX.14}$$

mostrando que cambios en $\mathbf{b}$ en el espacio rango modifican el vector solución $\mathbf{x}$ pero no el vector residual $\mathbf{r}$. Por el contrario, si $\mathbf{b}$ viene perturbado por un vector que está enteramente en el subespacio nulo de A^t, es decir $\overline{\mathbf{b}} = \mathbf{b} + \mathbf{z}$, donde $A^t\ \mathbf{z} = \mathbf{0}$, entonces

$$\overline{\mathbf{x}} = \mathbf{x} \quad \text{y} \quad \overline{\mathbf{r}} = \mathbf{r} + \mathbf{z} \ . \tag{XIX.15}$$

mostrando que cambios en $\mathbf{b}$ en el espacio nulo modifican el vector residual $\mathbf{r}$, mientras que el vector solución $\mathbf{x}$ permanece inalterado.

3. LAS ECUACIONES NORMALES

La propiedad más significativa del vector residual optimal del problema de mínimos cuadrados es que está enteramente en el espacio nulo de la matriz A^t. En términos algebráicos, una solución $\mathbf{x}$ satisface

$$A^t\ (\mathbf{b} - A\ \mathbf{x}) = \mathbf{0} \ , \tag{XIX.16}$$

que es equivalente a

$$A^t A\ \mathbf{x} = A^t\ \mathbf{b} \ . \tag{XIX.17}$$

Este sistema de ecuaciones se conoce como el de las **ecuaciones normales**.

La matriz simétrica A^tA, conocida como la **matriz de la ecuación normal**, es semidefinida positiva para una matriz cualquiera A, y es definida positiva si y sólo si las columnas de A son linealmente independientes. Las ecuaciones normales son siempre compatibles, es decir existe siempre una solución de $(XIX.17)$ aunque A^tA sea singular. Cualquier solución $\mathbf{x}$ tiene que satisfacer las ecuaciones normales, y cualquier vector $\mathbf{x}$ que satisface las ecuaciones normales tiene que ser una solución de mínimos cuadrados.

Las ecuaciones normales son de gran importancia práctica por el hecho de que ofrecen una manera directa para calcular la solución de mínimos cuadrados. Si A^tA es definida positiva, es decir si A tiene rango lleno, las ecuaciones normales $(XIX.17)$ tienen solución única, que se puede encontrar usando el algoritmo de factorización de Cholesky. Si A^tA es singular, se puede hallar una solución de las ecuaciones normales usando una versión de la factorización de Cholesky que incluya un intercambio simétrico. Entonces, el problema de mínimos cuadrados se puede resolver siguiendo los siguientes pasos:
(i) formar la matriz de la ecuación normal A^tA y el vector $A^t\ \mathbf{b}$;
(ii) calcular la factorización de Cholesky $A^tA = L^tL$, con L triangular superior;
(iii) resolver los dos sistemas $L^t\ \mathbf{y} = A^t\ \mathbf{b}$ y $L\ \mathbf{x} = \mathbf{y}$. El vector $\mathbf{x}$ es la solución deseada.

Ejemplo. Como ejemplo especifico de matriz de rango lleno, reconsideremos el ejemplo anterior:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix}$$

con las columnas de A linealmente independientes. La matriz de la ecuación normal $A^t A$ y el vector $A^t\ \mathbf{b}$ para el vector asociado $\mathbf{b}$ son:

$$A^t\ A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{y} \quad A^t\ \mathbf{b} = \begin{pmatrix} 1 \\ -4 \end{pmatrix} ,$$

y la solución de las ecuaciones normales

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

es

$$\mathbf{x} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

que es la solución de mínimos cuadrados obtenida anteriormente.

Ejemplo. Para ver un ejemplo especifico de matriz de rango deficiente, consideremos una matriz variación de la matriz usada en el ejemplo anterior, creada añadiendo una columna dependente a la matriz A, y el mismo vector $\mathbf{b}$:

$$\overline{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix}$$

ahora las columnas de $\overline{A}$ son linealmente dependientes. La matriz de la ecuación normal $\overline{A}^t\overline{A}$ y el vector $\overline{A}^t\ \mathbf{b}$ para el vector asociado $\mathbf{b}$ son:

$$\overline{A}^t\overline{A} = \begin{pmatrix} 2 & 1 & \frac{3}{2} \\ 1 & 2 & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} & \frac{3}{2} \end{pmatrix} \quad \text{y} \quad \overline{A}^t\mathbf{b} = \begin{pmatrix} 1 \\ -4 \\ -\frac{3}{2} \end{pmatrix} .$$

Aunque $\overline{A}^t\overline{A}$ es obviamente singular (la última columna es la media de las primeras dos columnas), las ecuaciones normales son compatibles, y el vector $\mathbf{x} = (2, -3, 0)^t$ (la solución anterior ampliada con una componente nula correspondiente a la última columna de $\overline{A}$) resuelve la ecuaciones normales. La solución de mínimos cuadrados ya no es única, y cualquiera del conjunto infinito de vectores que satisfacen las ecuaciones normales singulares per compatibles es una solución de mínimos cuadrados. Por ejemplo, $\mathbf{x_1} = (3, -2, -2)^t$ y $\mathbf{x_2} = (1, -4, 2)^t$ resuelven las ecuaciones normales y son consecuentemente soluciones de mínimos cuadrados.

Desde el punto de vista computacional, resolver el problema de mínimos cuadrados con las ecuaciones normales es eficiente: para formar $A^t A$ y $A^t\mathbf{b}$ se necesitan aproximadamente $m \cdot n^2/2$ operaciones, calcular la factorización de Cholesky usa del orden de $n^3/6$ operaciones y resolver los dos sistemas triangulares implica aproximadamente n^2 operaciones. Desafortunadamente, el uso de las ecuaciones normales presenta el problema de la estabilidad numérica, dado que el número de condición de la matriz $A^t A$ es el cuadrado del número de condición de A. Consecuentemente, la matriz de la ecuación normal está seriamente mal-condicionada si la matriz A misma está ligeramente mal-condicionada.

El mal condicionamiento de la matriz de la ecuación normal asociada puede conducir no solo a una aproximación no precisa de la solución calculada de las ecuaciones normales, sino también a una perdida de información cuando el rango numérico de A es marginal.

4. APLICACIONES

Los científicos a menudo coleccionan datos e intentan encontrar una relación funcional entre las variables. Si los datos son $n+1$ puntos del plano, es posible encontrar un polinomio de grado n o inferior que pasa por todos los puntos. Este polinomio se llama **polinomio de interpolación**. Dado que los datos tienen en general error experimental, no hay razón para pedir que las funciones pasen por todos los puntos. De hecho, polinomios de grado inferior que no pasan por los puntos de manera exacta, dan una mejor descripción de la relación entre variables. Por ejemplo, si la relación entre variables es actualmente lineal y los datos tienen un pequeño error, sería desastroso usar un polinomio de interpolación.

Dada una tabla de datos

$$\begin{array}{c|c|c|c|c} x & x_1 & x_2 & \dots & x_m \\ \hline y & y_1 & y_2 & \dots & y_m \end{array} \qquad (XIX.18)$$

deseamos encontrar una función lineal

$$y = c_0 + c_1\ x \qquad (XIX.19)$$

que mejor aproxima los datos en el sentido de mínimos cuadrados. Si se pide que

$$y_i = c_0 + c_1\ x_i \qquad \text{para} \quad i = 1, 2, \dots, m \qquad (XIX.20)$$

obtenemos un sistema de m ecuaciones en dos incognitas

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} . \qquad (XIX.21)$$

La función lineal cuyos coeficientes son la solución de mínimos cuadrados de $(XIX.21)$ viene a ser la aproximación de mínimos cuadrados a los datos con una función lineal.

Ejemplo. Queremos encontrar la mejor aproximación de mínimos cuadrados con una función lineal a los siguientes datos:

x	0	3	6
y	1	4	5

Para estos datos el sistema $(XIX.21)$ es $A\ \mathbf{c} = \mathbf{y}$ con

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 6 \end{pmatrix} , \qquad \mathbf{c} = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} \quad \text{y} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} .$$

Las ecuaciones normales, $A^t A\ \mathbf{c} = A^t\ \mathbf{y}$, se reducen a

$$\begin{pmatrix} 3 & 9 \\ 9 & 45 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 10 \\ 42 \end{pmatrix} .$$

La solución a este sistema es $(\frac{4}{3}, \frac{2}{3})$. Entonces la mejor aproximación de mínimos cuadrados linear es

$$y = \frac{4}{3} + \frac{2}{3}\ x .$$

Podriamos reconsiderar este ejemplo usando el vector residual $\mathbf{r}(\mathbf{c}) = \mathbf{y} - A\ \mathbf{c}$ y

$$\begin{aligned} ||\mathbf{r}(\mathbf{c})||_2^2 &= ||\mathbf{y} - A\ \mathbf{c}||_2^2 = \\ &= [1 - (c_0 + 0\ c_1)]^2 + [4 - (c_0 + 3\ c_1)]^2 + [5 - (c_0 + 6\ c_1)]^2 = \\ &= f(c_0, c_1) . \end{aligned}$$

Entonces $||\mathbf{r}(\mathbf{c})||_2^2$ puede pensarse como una función $f(c_0, c_1)$ de dos variables. El minímo de tale función se presentará cuando sus derivadas parciales son nulas:

$$\frac{\partial f}{\partial c_0} = -2(10 - 3\ c_0 - 9\ c_1) = 0$$

$$\frac{\partial f}{\partial c_1} = -6(14 - 3\ c_0 - 15\ c_1) = 0$$

que resolviendo da el resultado anterior.

Si los datos no aparecen en relación lineal, se podria usar un polinomio de grado mayor. Es decir, para encontrar los coeficientes c_0, $c_1, \ldots$, c_n de la mejor aproximación por mínimos cuadrados a los datos

x	x_1	x_2	$\ldots$	x_m
y	y_1	y_2	$\ldots$	y_m

$(XIX.18)$

con un polinomio de grado n, tenemos que encontrar la solución de mínimos cuadrados al sistema

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & \ldots & x_m^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} . \qquad (XIX.22)$$

CAPITULO XX. LOS METODOS DE TRANSFORMACION ORTOGONAL

1. LAS TRANSFORMACIONES DE HOUSEHOLDER

Debido a las posibles dificultades numéricas del uso de las ecuaciones normales, los modernos métodos de mínimos cuadrados se han desarrollado basándose en las **transformaciones ortogonales**, que preservan las distancias euclídeas y no empeoran las condiciones de la matriz A. La idea es transformar un problema de mínimos cuadrados de manera que sea fácil de resolver, reduciendo la matriz A a la forma que revela el rango. El término forma triangular que revela el rango denota una matriz genérica $m \times n$ de rango r correspondiente a la matriz $\tilde{T} = \begin{pmatrix} T \\ 0 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ 0 & 0 \end{pmatrix}$, con T_{11} matriz $r \times r$ no singular triangular superior y T_{12} matriz $r \times (n-r)$. Cuando $\tilde{T}$ tiene rango lleno de filas entonces no aparecen ambos bloques de ceros; si $\tilde{T}$ es de rango lleno de columna, la matriz T_{12} y el bloque de ceros derecho no aparecerán. Más en general, el rango de una matriz $m \times n$, $\tilde{F}$, es r si $\tilde{F}$ es de la forma $\tilde{F} = \begin{pmatrix} F \\ 0 \end{pmatrix}$, con F matriz $r \times n$ y las filas de F son linealmente independientes.

Dado que ya no estamos resolviendo ecuaciones, el conjunto de transformaciones que se puede aplicar a A sin cambiar la solución está restringido. Las transformaciones ortogonales son obvias en este contexto dado que estas no alteran la norma l_2.

Sea A una matriz no nula $m \times n$, con rango igual a r. Supóngase que se pueda encontrar una matriz $m \times m$ ortogonal Q que produzca la forma que revela el rango

$$Q^t\ A = \begin{pmatrix} F \\ 0 \end{pmatrix}\ , \tag{XX.1}$$

donde F es una matriz $r \times n$ y tiene las filas linealmente independientes; el bloque de ceros no aparece si $r = m$.

La forma $(XX.1)$ nos permite resolver el problema de mínimos cuadrados usando las matrices Q y F. Sea $\mathbf{d}$ el vector transformado $Q^t\ \mathbf{b}$, descompuesto en

$$\mathbf{d} = Q^t\ \mathbf{b} = \begin{pmatrix} \mathbf{d_r} \\ \mathbf{d_{m-r}} \end{pmatrix}\ . \tag{XX.2}$$

Usando esta definición y la estructura mostrada en $(XX.1)$, se puede escribir el vector residual transformado como

$$Q^t\ (\mathbf{b} - A\ \mathbf{x}) = \begin{pmatrix} \mathbf{d_r} \\ \mathbf{d_{m-r}} \end{pmatrix} - \begin{pmatrix} F\ \mathbf{x} \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{d_r} - F\ \mathbf{x} \\ \mathbf{d_{m-r}} \end{pmatrix}\ . \tag{XX.3}$$

Dado que la matriz Q^t es ortogonal, su aplicación al vector residual no altera la norma euclídea, y

$$||\mathbf{b} - A\ \mathbf{x}||_2^2 = ||Q^t\ (\mathbf{b} - A\ \mathbf{x})||_2^2\ . \tag{XX.4}$$

Combinando esta relación con la forma especial del vector residual transformado, se concluye que

$$||\mathbf{b} - A\ \mathbf{x}||_2^2 = ||Q^t\ (\mathbf{b} - A\ \mathbf{x})||_2^2 = ||\mathbf{d_r} - F\ \mathbf{x}||_2^2 + ||\mathbf{d_{m-r}}||_2^2 \geq ||\mathbf{d_{m-r}}||_2^2\ , \qquad (XX.5)$$

que significa que $||\mathbf{b} - A\ \mathbf{x}||_2$ no puede ser menor que $||\mathbf{d_{m-r}}||_2$ para cualquier vector $\mathbf{x}$. El valor más pequeño posible para $||\mathbf{b} - A\ \mathbf{x}||_2^2$ es la cota inferior. Igualdades con la cota inferior se obtienen si y sólo si el vector $\mathbf{x}$ satisface

$$F\ \mathbf{x} = \mathbf{d_r}\ . \qquad (XX.6)$$

Dado que el rango de la matriz F es igual a r, el sistema $F\ \mathbf{x} = \mathbf{d_r}$ tiene que ser compatible. Cuando $F\ \mathbf{x} = \mathbf{d_r}$, el vector residual transformado satisface $Q^t\ (\mathbf{b} - A\ \mathbf{x}) = \begin{pmatrix} 0 \\ \mathbf{d_{m-r}} \end{pmatrix}$ y $||\mathbf{b} - A\ \mathbf{x}||_2 = ||\mathbf{d_{m-r}}||_2$. Esto demuestra que cualquier vector $\mathbf{x}$ que satisfaga $F\ \mathbf{x} = \mathbf{d_r}$ será una solución de mínimos cuadrados. Suponiendo que los sistemas en los cuales aparecen las matrices F son fáciles de resolver, se sigue que el problema de mínimos cuadrados se puede resolver encontrando una matriz ortogonal Q^t que nos dé la forma $(XX.1)$.

Antes de presentar la **factorización QR** para resolver el problema de mínimos cuadrados, es necesario introducir las **transformaciones de Householder**.

La técnica más popular para construir una matriz ortogonal que reduzca la matriz A en forma triangular usa una clase especial de matrices que son simultaneamente simétricas, elementales y ortogonales. Para cualquier vector no nulo $\mathbf{u}$, la correspondiente **transformación de Householder** (o **matriz de Householder**, o **reflector de Householder**) es una matriz de la forma

$$H = H(u) = I - \frac{2\mathbf{u}\mathbf{u}^t}{||\mathbf{u}||_2^2} = I - \frac{\mathbf{u}\mathbf{u}^t}{\beta}\ , \qquad \text{con}\ \ \beta = \frac{||\mathbf{u}||_2^2}{2}\ , \qquad (XX.7)$$

donde el vector $\mathbf{u}$ es el **vector de Householder**

Teorema XX.1

Si H es la matriz definida en $(XX.7)$, entonces

1) $H = H^t$,

2) $H = H^{-1}$,

que es lo mismo que decir que la matriz H es simétrica y ortogonal.

Demostración: Antes de todos, volvemos a escribir la definición de la matriz de Householder usando el vector normalizado de norma uno

$$\mathbf{w} = \frac{\mathbf{u}}{||\mathbf{u}||_2}\ .$$

Entonces, 1) usando las propiedades básicas de las matrices:

$$H^t = I^t - 2\ (\mathbf{w}\ \mathbf{w}^t)^t = I - 2\ (\mathbf{w}^t)^t\ \mathbf{w}^t = I - 2\ \mathbf{w}\ \mathbf{w}^t = H\ .$$

2) usando el hecho que $\mathbf{w}^t\ \mathbf{w} = 1$, se sigue que

$$H^t\ H = H^2 = (I - 2\ \mathbf{w}\ \mathbf{w}^t)\ (I - 2\ \mathbf{w}\ \mathbf{w}^t) = I - 4\ \mathbf{w}\ \mathbf{w}^t + 4\ \mathbf{w}\ (\mathbf{w}^t\ \mathbf{w})\ \mathbf{w}^t = I\ .$$

c.q.d.

Entonces, las matrices de Householder son simétricas y ortogonales, y dependen sólo de la dirección del vector $\mathbf{u}$.

En el contexto de las reducciones a matrices triangulares, las matrices de Householder poseen dos propiedades cruciales:

- para cualquier par de vectores distintos de igual norma l_2, existe una transformación de Householder que transforma el uno en el otro,

$$H\ \mathbf{a} = (I - \frac{\mathbf{u}\mathbf{u}^t}{\beta})\ \mathbf{a} = \mathbf{b}\ , \qquad (XX.8)$$

con $||\mathbf{a}||_2 = ||\mathbf{b}||_2$. Esto implica que el vector $\mathbf{u}$ tiene que satisfacer la condición

$$-\frac{\mathbf{u}^t\ \mathbf{a}}{\beta}\ \mathbf{u} = \mathbf{b} - \mathbf{a}\ , \qquad (XX.9)$$

es decir, $\mathbf{u}$ es un múltiplo de $\mathbf{b} - \mathbf{a}$;

- cualquier vector $\mathbf{c}$ transformado por una matriz de Householder posee una forma especial:

$$H\ \mathbf{c} = (I - \frac{\mathbf{u}\mathbf{u}^t}{\beta})\ \mathbf{c} = \mathbf{c} - \frac{\mathbf{u}^t\mathbf{c}}{\beta}\ \mathbf{u}\ , \qquad (XX.10)$$

de manera que $H\ \mathbf{c}$ es la diferencia entre el vector original $\mathbf{c}$ y un múltiplo especial del vector de Householder $\mathbf{u}$.

Claramente el vector $\mathbf{c}$ no varía si $\mathbf{u}^t\ \mathbf{c} = 0$. Además, calcular el producto $H\ \mathbf{c}$ no necesita los elementos explícitos de H, sino sólo el vector $\mathbf{u}$ y el escalar β.

2. LA FACTORIZACION QR

Para una matriz genérica A, $n \times n$, no singular, las propiedades que se acaban de describir permiten construir una sucesión de $n - 1$ matrices de Householder tales que

$$H_{n-1}\ \ldots\ H_2\ H_1\ A = R\ , \qquad (XX.11)$$

donde R es una matriz $n \times n$ no singular y triangular superior.

El primer paso de este proceso es construir una matriz de Householder H_1 que transforma $\mathbf{a_1}$ (la primera columna de A) en un múltiplo de $\mathbf{e_1}$, es decir, se desean crear ceros en las componentes 2 hasta n del vector $\mathbf{a_1}$. La norma euclídea se conserva bajo transformaciones ortogonales, de manera que

$$H_1\ \mathbf{a_1} = (I - \frac{\mathbf{u_1}\mathbf{u_1}^t}{\beta_1})\ \mathbf{a_1} = \pm||\mathbf{a_1}||_2\ \mathbf{e_1} = \begin{pmatrix} r_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}\ , \qquad (XX.12)$$

donde $|r_{11}| = ||\mathbf{a_1}||_2$. De la expresión $(XX.9)$ sabemos que el vector $\mathbf{u_1}$ tiene que ser un múltiplo de $||\mathbf{a_1}||_2\ \mathbf{e_1} - \mathbf{a_1}$, y dado que H_1 depende sólo de la dirección de $\mathbf{u_1}$, podemos elegir el vector $\mathbf{u_1}$ como

$$\mathbf{u_1} = \begin{pmatrix} a_{11} - r_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} . \qquad (XX.13)$$

Al definir $\mathbf{u_1}$, el signo de r_{11} se puede eligir positivo o negativo (excepto cuando $\mathbf{a_1}$ es ya un múltiplo de $\mathbf{e_1}$), y para evitar el problema de la cancelación de términos parecidos, usualmente se escoge el signo opuesto al signo de a_{11}, de manera que

$$\text{signo}(r_{11}) = -\text{signo}(a_{11}) . \qquad (XX.14)$$

Después de la primera aplicación de las transformaciones de Householder, la primera columna de la matriz parcialmente reducida $A^{(2)} = H_1\ A$ es un múltiplo de $\mathbf{e_1}$, y los demás elementos han sido alterados

$$A^{(2)} = H_1\ A = \begin{pmatrix} r_{11} & a_{12}^{(2)} & \ldots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \ldots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \ldots & a_{nn}^{(2)} \end{pmatrix} . \qquad (XX.15)$$

En muy importante notar que a diferencia de la eliminación Gaussiana, la primera fila de la matriz A viene modificada por efecto de la transformación de Householder H.

Al construir la segunda transformación de Householder, el objetivo principal es reducir la segunda columna a la forma correcta, sin alterar la primera fila y la primera columna de la matriz parcialmente reducida. Debido a la propiedad $(XX.10)$, se puede obtener este resultado definiendo el segundo vector de Householder $\mathbf{u_2}$ con la primera componente nula. Habiendo escogido así el vector $\mathbf{u_2}$, la aplicación de la matriz de Householde H_2 a un vector genérico no cambia la primera componente, y la aplicación a un múltiplo de $\mathbf{e_1}$ deja el vector entero como está.

Si A es una matriz no singular, se pueden efectuar $n-1$ pasos de reducción de Householder, para obtener $H_{n-1}\ \ldots\ H_2\ H_1\ A = R$, con R una matriz triangular superior no singular. Si se denota con Q^t la matriz ortogonal $n \times n$

$$Q^t = H_{n-1}\ \ldots\ H_2\ H_1 \qquad \Longrightarrow \qquad Q = H_1\ H_2\ \ldots\ H_{n-1} . \qquad (XX.16)$$

Cualquiera de las dos formas

$$Q^t\ A = R \qquad \text{o} \qquad A = Q\ R \qquad (XX.17)$$

se conoce como la factorización QR de la matriz A.

Una vez conocida la factorización QR de A, ecuación $(XX.17)$, la solución al sistema $A\ \mathbf{x} = \mathbf{b}$ se obtiene resolviendo el sistema triangular superior $R\ \mathbf{x} = Q^t\ \mathbf{b}$.

Ejemplo. Sean

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \\ 3 & -1 & -1 \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 3 \\ 2 \\ 6 \end{pmatrix} .$$

Aplicando la $(XX.13)$, obtenemos

$$||\mathbf{a_1}||_2 = \sqrt{14} , \qquad r_{11} = -\sqrt{14} , \qquad \text{y} \qquad \mathbf{u_1} = \begin{pmatrix} 1+\sqrt{14} \\ 2 \\ 3 \end{pmatrix} .$$

Entonces al primer paso, en aritmética de cuatro dígitos,

$$H_1 = \begin{pmatrix} -0.2673 & -0.5345 & -0.8018 \\ -0.5345 & 0.7745 & -0.3382 \\ -0.8018 & -0.3382 & 0.4927 \end{pmatrix} \quad \text{y} \quad A^{(2)} = \begin{pmatrix} -0.2673 & -1.069 & -0.2673 \\ & 2.127 & 0.04368 \\ & -2.309 & -2.434 \end{pmatrix} .$$

Efectuando un segundo paso, se obtiene

$$R = \begin{pmatrix} -3.742 & -1.069 & -0.2673 \\ & -3.140 & -1.820 \\ & & -1.617 \end{pmatrix} \quad \text{y} \quad Q^t\ \mathbf{b} = \begin{pmatrix} -6.682 \\ 1.320 \\ -1.617 \end{pmatrix} .$$

Con sustitución regresiva se obtiene la solución $\mathbf{x} = (2, -1, 1)^t$.

En general, es necesario un intercambio de columnas para asegurar que el rango está plenamente revelado. Sin el intercambio de columnas la reducción de Householder terminaría inmediatamente con una matriz no nula cuya primera columna es cero. Si las demás columnas son también ceros, la reducción termina. De otra manera, existe por lo menos una columna no nula, la columna pivote, que se puede eligir como candidata para la siguiente reducción. Como en la eliminación gaussiana, una estrategia de pivoteo pide que se escoja la columna pivote como la primera columna de norma máxima (otra posibilidad es escoger la columna "menos reducida").

En general, si A es $m \times n$ de rango r, se necesitarán r permutaciones $\{P_k\}$ y r matrices de Householder $\{H_k\}$, $k = 1, \ldots, r$. Después de estos r pasos, la configuración final será

$$H_r\ \ldots\ H_1\ A\ P_1\ \ldots\ P_r = \tilde{R} , \tag{XX.18}$$

donde

$$\tilde{R} = \begin{pmatrix} R \\ 0 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix} \tag{XX.19}$$

es triangular superior que nos revela el rango, y R_{11} es una matriz $r \times r$ no singular triangular superior. Con una correcta estrategia de pivoteo, dependiente del problema original, y aritmética exacta, este procedimiento de Householder terminará después de r pasos, cuando la matriz restante se transformará en cero. Combinando los intercambios de columnas en una sola matriz de permutación P y las transformaciones de Householder en una sola matriz ortogonal Q^t, se obtiene

$$Q^t\ A\ P = \tilde{R} , \tag{XX.20}$$

o equivalentemente,

$$A\ P = Q\ \tilde{R}\ , \qquad (XX.21)$$

y

$$Q^t\ A = \tilde{R}\ P^t = \begin{pmatrix} R \\ 0 \end{pmatrix}\ P^t = \begin{pmatrix} R\ P^t \\ 0 \end{pmatrix}\ . \qquad (XX.22)$$

Las filas de la matriz $R\ P^t$ son linealmente independientes, de manera que la matriz transformada $Q^t\ A$ tiene la forma deseada $(XX.1)$, con $F = R\ P^t$. El problema de mínimos cuadrados se puede entonces resolver con el siguiente algoritmo:
(i) Calcular la factorización QR de la matriz A, usando la reducción de Householder (que determina el rango r, y las matrices P, Q y R);
(ii) Formar el vector $\mathbf{d} = Q^t\ \mathbf{b}$, y denotar con $\mathbf{d_r}$ las primeras r componentes de $\mathbf{d}$;
(iii) Calcular cualquier solución $\mathbf{y}$ del sistema $R\ \mathbf{y} = \mathbf{d_r}$;
(iv) Formar $\mathbf{x} = P\ \mathbf{y}$.

El número de operaciones requerido para resolver el problema de mínimos cuadrados de esta manera es del orden de $2 \cdot m \cdot n \cdot r - r^2\ (m+n) + \frac{2}{3}n^3$.

Si la matriz A posee columnas linealmente independientes ($r = n$), la matriz R es no singular, la solución de $R\ \mathbf{y} = \mathbf{d_r}$ es única, y la solución del problema de mínimos cuadrados es única. En este caso resolver el problema con el método QR es aproximadamente dos veces más costoso que resolviendolo con las ecuaciones normales. Cuando las columnas de A son linealmente dependientes, de manera que $r < n$, el sistema $r \times n\ R\ \mathbf{y} = \mathbf{d_r}$ posee un número infinito de soluciones. Dado que R es de la forma $R = (R_{11} R_{12})$, con R_{11} triangular superior no singular, es posible escoger un vector $\mathbf{y} = (\mathbf{y_B}, 0)^t$ tal que $R\ \mathbf{y} = \mathbf{d_r}$ y con la propiedad de que $R_{11}\ \mathbf{y_B} = \mathbf{d_r}$. Si $\mathbf{y}$ posee esta forma, la solución $\mathbf{x} = P\ \mathbf{y}$ es sencillamente una reordenación de $\mathbf{y}$, y es llamada una **solución básica** del problema de mínimos cuadrados.

Ejemplo. Sean

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 7 & 6 & 10 \\ 4 & 4 & 6 \\ 1 & 0 & 1 \end{pmatrix} \qquad \text{y} \qquad \mathbf{b} = \begin{pmatrix} 6 \\ 6 \\ 8 \\ 3 \end{pmatrix}\ .$$

La representación única de $\mathbf{b}$ en términos de los espacios rango y nulo de A es

$$\mathbf{b} = \mathbf{b_R} + \mathbf{b_N}\ , \qquad \text{con} \qquad \mathbf{b_R} = \begin{pmatrix} 4 \\ 8 \\ 6 \\ -1 \end{pmatrix} \qquad \text{y} \qquad \mathbf{b_N} = \begin{pmatrix} 2 \\ -2 \\ 2 \\ 4 \end{pmatrix}\ ,$$

con $||\mathbf{b_N}||_2 = 5.292$. Aplicando la reducción de Householder con intercambio de columnas en A, obtenemos

$$R = \begin{pmatrix} -11.87 & -7.411 & -8.169 \\ & 1.038 & -0.5191 \end{pmatrix} \qquad \text{y} \qquad P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\ .$$

El vector transformado $\mathbf{d} = Q^t\ \mathbf{b}$ es

$$\mathbf{d} = Q^t\ \mathbf{b} = \begin{pmatrix} d_r \\ d_{m-r} \end{pmatrix} = \begin{pmatrix} -10.36 \\ 3.115 \\ 3.419 \\ 4.038 \end{pmatrix} ,$$

con $||\mathbf{d_{m-r}}||_2 = ||\mathbf{b_N}||_2 = 5.292$. Ahora, el vector $\mathbf{y_B}$ que satisface $R_{11}\ \mathbf{y_B} = \mathbf{d_r}$ es

$$\mathbf{y_B} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} ,$$

de manera que la solución básica del problema de mínimos cuadrados $\mathbf{x}$ es

$$\mathbf{x} = P\ \begin{pmatrix} y_B \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} .$$

Finalmente, el vector residual optimal es

$$\mathbf{r} = \mathbf{b} - A\ \mathbf{x} = \begin{pmatrix} 2 \\ -2 \\ 2 \\ 4 \end{pmatrix} .$$

3. LAS ROTACIONES DE GIVENS

Vamos a presentar ahora un método más efectivo que los estudiados anteriormente para resolver el problema de mínimos cuadrados. Como el método de Householder, el **método de las rotaciones de Givens** consiste en factorizar la matriz A como producto de una matriz ortogonal Q y una matriz triangular R, $A = Q\,R$. La razón de estos esquemas es que las matrices ortogonales cumplen unas propiedades que permiten ser manipuladas computacionalmente sin que los errores de redondeo crezcan sin control.

Consideramos antes el caso del plano R^2, y sea $\mathbf{v} = (v_1, v_2) \in R^2$. Si deseamos rotar el vector $\mathbf{v}$ de un ángulo θ, multiplicaremos el vector por una matriz Q del siguiente tipo:

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad Q\,\mathbf{v} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1\ \cos\theta - v_2\ \sin\theta \\ v_1\ \sin\theta + v_2\ \cos\theta \end{pmatrix}.$$

La multiplicación de un vector por la transpuesta de la matriz Q produce el efecto contrario, esto es rotar el vector en sentido opuesto al ángulo θ.

Como ya dicho en el método de Householder, el objetivo de los métodos de transformación ortogonales para resolver sistemas lineales es el de conseguir una factorización del tipo $A = Q\,R$, esto es $Q^t\,A = R$, donde R es triangular superior. Para ello se necesita que la multiplicación de la matriz Q^t por A rote o convierta el primer vector de esta última $\mathbf{v} = (v_1, v_2, \ldots, v_n)^t$ en un vector de la forma $Q^t\,\mathbf{v} = (w_1, 0, \ldots, 0)^t$.

En el caso de matrices de dimensión 2×2 se tiene

$$Q^t\,\mathbf{v} = \begin{pmatrix} v_1\,\cos\theta + v_2\,\sin\theta \\ -v_1\,\sin\theta + v_2\,\cos\theta \end{pmatrix} \quad \Rightarrow \quad -v_1\,\sin\theta + v_2\,\cos\theta = 0$$

y para que se cumpla esta última igualdad es suficiente que sea

$$\sin\theta = \frac{v_2}{\sqrt{v_1^2+v_2^2}}, \qquad \cos\theta = \frac{v_1}{\sqrt{v_1^2+v_2^2}}.$$

En el caso general de matrices de dimensión $n \times n$, a las matrices $Q_{j\,i}$ que, multiplicadas por una matriz A, convierten el elemento a_{ji} de esta última en 0, se les denomina **rotaciones de Givens**. Su estructura es la siguiente

$$Q_{ji} = \begin{cases} q_{kk} = 1 & \text{si} \quad k \neq i, j \\ q_{ii} = q_{jj} = c & \\ q_{ji} = -q_{ij} = s & \\ 0 & \text{en las otras posiciones} \end{cases}$$

donde los coeficientes c y s se definen como

$$c = \frac{a_{ii}}{\sqrt{a_{ii}^2 + a_{ji}^2}} \qquad \text{y} \qquad s = \frac{a_{ji}}{\sqrt{a_{ii}^2 + a_{ji}^2}}.$$

Al multiplicar la matriz Q_{ji}^t por la matriz A, se modifican sólo las componentes de las filas $i-$ésima y $j-$ésima, dado que sobre el resto de filas se aplica el mismo efecto que multiplicar por la matriz identidad. Se puede entonces pensar en un proceso de multiplicaciones ordenadas y sistematicas de matrices ortogonales Q_{ji}^t por la matriz A hasta reducir la matriz original A a una matriz triangular superior R.

La primera matriz Q_{21}^t que se multiplica por A, transforma el primer vector columna de A, $\mathbf{a}_1 = (a_{11}, a_{21}, a_{31}, \ldots, a_{n1})^t$, en $Q_{21}^t\,\mathbf{a}_1 = (a_{11}^{(1)}, 0, a_{31}^{(1)}, \ldots, a_{n1}^{(1)})^t$. Sucesivamente $Q_{31}^t\,Q_{21}^t\,\mathbf{a}_1 = (a_{11}^{(2)}, 0, 0, a_{41}^{(2)}, \ldots, a_{n1}^{(2)})^t$ y después de repetir $n-1$ veces el mismo proceso tenemos $Q_{n1}^t \ldots Q_{21}^t\,\mathbf{a}_1 = (a_{11}^{(n-1)}, 0, \ldots, 0)^t$. Como solamente se han utilizado rotaciones, la norma de este último vector es la misma que la del vector original $\mathbf{a}_1$.

Una vez acabado el proceso con la primera columna de A, se repite el proceso con la segunda, luego con todas las siguientes hasta acabar con la $(n-1)$-ésima. La multiplicación sucesiva de las $\frac{(n-1)\,n}{2}$ matrices Q_{ji}^t constituye una matriz ortogonal Q^t que completa la factorización de A en una matriz R triangular superior

$$Q_{n\,n-1}^t \cdots Q_{n1}^t \cdots Q_{21}^t = Q^t \quad \Rightarrow \quad Q^t\,M = R.$$

El coste operativo de la factorización $M = Q\,R$ es de $2\,\frac{n^3}{3}$ operaciones, esto es aproximadamente el doble de una facorización $L\,U$. La ventaja es que las descomposiciones $Q\,R$ no amplifican el error de redondeo surgidos en las operaciones.

Una vez obtenida la factorización $M = Q\,R$ se puede resolver el sistema lineal original $M\,\mathbf{x} = \mathbf{b}$, multiplicando Q^t por $\mathbf{b}$ y resoviendo directamente, mediante sustitución regresiva, el sistema triangular $R\,\mathbf{x} = Q^t\,\mathbf{b}$.

Ejemplo. Sean A la matriz y $\mathbf{b}$ el vector del problema $A\,\mathbf{x} = \mathbf{b}$

$$A = \begin{pmatrix} 3 & 15 & -1 \\ -4 & 1 & 2 \\ 9 & 9 & 5 \end{pmatrix} \qquad \text{y} \qquad \mathbf{b} = \begin{pmatrix} 17 \\ -1 \\ 23 \end{pmatrix}.$$

Empezamos definiendo la primera matriz ortogonal Q_{21} definida por

$$Q_{21} = \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

con

$$c = \frac{3}{\sqrt{3^2 + (-4)^2}} = \frac{3}{5} = 0.6 \qquad \text{y} \qquad s = -\frac{4}{\sqrt{3^2 + (-4)^2}} = -\frac{4}{5} = -0.8.$$

Esto es

$$Q_{21} = \begin{pmatrix} 0.6 & 0.8 & 0 \\ -0.8 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

y por lo tanto

$$Q_{21}^t\,A = \begin{pmatrix} 0.6 & -0.8 & 0 \\ 0.8 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 15 & -1 \\ -4 & 1 & 2 \\ 9 & 9 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 8.2 & -2.2 \\ 0 & 12.6 & 0.4 \\ 9 & 9 & 5 \end{pmatrix}.$$

Ahora pasamos a definir la segunda matriz ortogonal

$$Q_{31} = \begin{pmatrix} c & 0 & -s \\ 0 & 1 & 0 \\ s & 0 & c \end{pmatrix},$$

con

$$c = \frac{5}{\sqrt{5^2 + 9^2}} = 0.48564 \qquad \text{y} \qquad s = \frac{9}{\sqrt{5^2 + 9^2}} = 0.87416.$$

Esto es

$$Q_{31} = \begin{pmatrix} 0.48564 & 0 & -0.87416 \\ 0 & 1 & 0 \\ 0.87416 & 0 & 0.48564 \end{pmatrix},$$

y por lo tanto

$$Q_{31}^t\,Q_{21}^t\,A = \begin{pmatrix} 0.48564 & 0 & 0.87416 \\ 0 & 1 & 0 \\ -0.87416 & 0 & 0.48564 \end{pmatrix} \begin{pmatrix} 5 & 8.2 & -2.2 \\ 0 & 12.6 & 0.4 \\ 9 & 9 & 5 \end{pmatrix} =$$

$$= \begin{pmatrix} 10.2956 & 11.8497 & 3.30237 \\ 0 & 12.6 & 0.4 \\ 0 & -2.7973 & 4.35136 \end{pmatrix}.$$

Finalmente la tercera matriz ortogonal es

$$Q_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix},$$

con

$$c = \frac{12.6}{\sqrt{12.6^2 + (-2.7973)^2}} = 0.97621 \quad \text{y} \quad s = \frac{-2.7973}{\sqrt{12.6^2 + (-2.7973)^2}} = -0.21674.$$

Esto es

$$Q_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.97621 & 0.21674 \\ 0 & -0.21674 & 0.97621 \end{pmatrix},$$

y por lo tanto

$$R = Q_{32}^t\, Q_{31}^t\, Q_{21}^t\, A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.97621 & -0.21674 \\ 0 & 0.21674 & 0.97621 \end{pmatrix} \begin{pmatrix} 10.2956 & 11.8497 & 3.30237 \\ 0 & 12.6 & 0.4 \\ 0 & -2.7973 & 4.35136 \end{pmatrix} =$$

$$= \begin{pmatrix} 10.2956 & 11.8497 & 3.30237 \\ 0 & 12.9068 & -0.552584 \\ 0 & 0 & 4.33463 \end{pmatrix}.$$

Además,

$$Q^t = Q_{32}^t\, Q_{31}^t\, Q_{21}^t = \begin{pmatrix} 0.291386 & -0.388514 & 0.874157 \\ 0.894659 & 0.434173 & -0.105254 \\ -0.338643 & 0.812743 & 0.4741 \end{pmatrix},$$

o que es lo mismo

$$Q = \begin{pmatrix} 0.291386 & 0.894659 & -0.338643 \\ -0.388514 & 0.434173 & 0.812743 \\ 0.874157 & -0.105254 & 0.4741 \end{pmatrix},$$

y resulta que

$$A = Q\, R.$$

Como dicho antes, una vez obtenida la factorización $A = Q\,R$ se puede resolver el sistema lineal original $A\,\mathbf{x} = \mathbf{b}$, multiplicando Q^t por $\mathbf{b}$ y resoviendo directamente, mediante sustitución regresiva, el sistema triangular $R\,\mathbf{x} = Q^t\,\mathbf{b}$. En este ejemplo,

$$Q^t\,\mathbf{b} = \begin{pmatrix} 0.291386 & -0.388514 & 0.874157 \\ 0.894659 & 0.434173 & -0.105254 \\ -0.338643 & 0.812743 & 0.4741 \end{pmatrix} \begin{pmatrix} 17 \\ -1 \\ 23 \end{pmatrix} = \begin{pmatrix} 25.4477 \\ 12.3542 \\ 4.33463 \end{pmatrix},$$

y por lo tanto el sistema a resolver es

$$\begin{pmatrix} 10.2956 & 11.8497 & 3.30237 \\ 0 & 12.9068 & -0.552584 \\ 0 & 0 & 4.33463 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 25.4477 \\ 12.3542 \\ 4.33463 \end{pmatrix}$$

que tiene solución $(1,1,1)^t$.

Ejemplo. Sean

$$A=\begin{pmatrix}1&2&2\\7&6&10\\4&4&6\\1&0&1\end{pmatrix}\qquad \text{y}\qquad \mathbf{b}=\begin{pmatrix}6\\6\\8\\3\end{pmatrix}$$

la matriz A y el vector $\mathbf{b}$ considerados en el ejemplo 3 donde se usaron las transformaciones de Householder. Usaremos ahora el método de las rotaciones de Givens para hallar una solución de mínimos cuadratos del problema.

Empezamos definiendo la primera matriz ortogonal Q_{21}, con $c=1/\sqrt{1^2+7^2}=0.141421$ y $s=7/\sqrt{1^2+7^2}=0.989949$, definida por

$$Q_{21}=\begin{pmatrix}c&-s&0&0\\s&c&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}=\begin{pmatrix}0.141421&-0.989949&0&0\\0.989949&0.141421&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},$$

y por lo tanto

$$Q_{21}^t\,A=\begin{pmatrix}0.141421&0.989949&0&0\\-0.989949&0.141421&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&2&2\\7&6&10\\4&4&6\\1&0&1\end{pmatrix}=$$

$$=\begin{pmatrix}7.07107&6.22254&10.1823\\0&-1.13137&-0.565685\\4&4&6\\1&0&1\end{pmatrix}.$$

Ahora pasamos a definir la segunda matriz ortogonal, siendo $c=0.870388$ y $s=0.492366$

$$Q_{31}=\begin{pmatrix}c&0&-s&0\\0&1&0&0\\s&0&c&0\\0&0&0&1\end{pmatrix}=\begin{pmatrix}0.870388&0&-0.492366&0\\0&1&0&0\\0.492366&0&0.870388&0\\0&0&0&1\end{pmatrix},$$

y por lo tanto

$$Q_{31}^t\,Q_{21}^t\,A=\begin{pmatrix}0.870388&0&0.492366&0\\0&1&0&0\\-0.492366&0&0.870388&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}7.07107&6.22254&10.1823\\0&-1.13137&-0.565685\\4&4&6\\1&0&1\end{pmatrix}=$$

$$=\begin{pmatrix}8.12404&7.38549&11.8168\\0&-1.13137&-0.565685\\0&0.417786&0.208893\\1&0&1\end{pmatrix}.$$

Siguiendo en la misma línea se calculan las demás matrices

$$Q_{41} = \begin{pmatrix} c & 0 & 0 & -s \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ s & 0 & 0 & c \end{pmatrix} = \begin{pmatrix} 0.992509 & 0 & 0 & -0.122169 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.122169 & 0 & 0 & 0.992509 \end{pmatrix},$$

$$Q_{41}^t\, Q_{31}^t\, Q_{21}^t\, A = \begin{pmatrix} 8.18535 & 7.33017 & 11.8504 \\ 0 & -1.13137 & -0.565685 \\ 0 & 0.417786 & 0.208893 \\ 0 & -0.902281 & -0.451141 \end{pmatrix};$$

$$Q_{32} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & -s & 0 \\ 0 & s & c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.938083 & -0.34641 & 0 \\ 0 & 0.34641 & -0.938083 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$Q_{32}^t\, Q_{41}^t\, Q_{31}^t\, Q_{21}^t\, A = \begin{pmatrix} 8.18535 & 7.33017 & 11.8504 \\ 0 & 1.20605 & 0.603023 \\ 0 & 0 & 0 \\ 0 & -0.902281 & -0.451141 \end{pmatrix};$$

$$Q_{42} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & 0 & -s \\ 0 & 0 & 1 & 0 \\ 0 & s & 0 & c \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.800717 & 0 & 0.599042 \\ 0 & 0 & 1 & 0 \\ 0 & -0.599042 & 0 & 0.800717 \end{pmatrix},$$

y finalmente

$$R = Q_{42}^t\, Q_{32}^t\, Q_{41}^t\, Q_{31}^t\, Q_{21}^t\, A = \begin{pmatrix} 8.18535 & 7.33017 & 11.8504 \\ 0 & 1.50621 & 0.753103 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Por otro lado

$$Q^t = Q_{42}^t\, Q_{32}^t\, Q_{41}^t\, Q_{31}^t\, Q_{21}^t = \begin{pmatrix} 0.122169 & 0.855186 & 0.488678 & 0.122169 \\ 0.733285 & -0.178367 & 0.277459 & -0.594555 \\ 0.408248 & 0.408248 & -0.816497 & 0 \\ 0.529813 & -0.264906 & 0.132453 & 0.794719 \end{pmatrix}$$

y

$$Q^t\, \mathbf{b} = \begin{pmatrix} 10.1401 \\ 3.76552 \\ -1.63299 \\ 5.03322 \end{pmatrix}.$$

Finalmente resolviendo las dos primeras ecuaciones del problemas $R\mathbf{x} = Q^t\,\mathbf{b}$ en las incognitas (x_1, x_2, x_3) se obtiene la solución de mínimos cuadratos

$$\mathbf{x} = (-1 - x_3, 2.5 - 0.5\, x_3, x_3)$$

cuyo vector residual optimal es, como ya visto en el ejemplo 3,

$$\mathbf{r} = \mathbf{b} - A\,\mathbf{x} = \begin{pmatrix} 2 \\ -2 \\ 2 \\ 4 \end{pmatrix}.$$