

FUNDAMENTOS MATEMÁTICOS

Ya son conocidas de cursos anteriores las factorizaciones de matrices cuadradas (LU, Cholesky, QR).

Ejemplo:

$$A = \begin{pmatrix} 0.4087 & 0.1593 & 0.6593 \\ 0.3515 & 0.9665 & 0.6245 \\ 0.6590 & 0.9342 & 0.9039 \end{pmatrix}$$

$$m = n = 3, \quad \text{rango}(A) = 3, \quad Q^t A = R.$$

$$Q = \begin{pmatrix} -0.4800 & 0.6615 & -0.5762 \\ -0.4129 & -0.7499 & -0.5169 \\ -0.7740 & -0.0103 & 0.6331 \end{pmatrix}$$

$$R = \begin{pmatrix} -0.8514 & -1.1986 & -1.2740 \\ 0 & -0.6290 & -0.0415 \\ 0 & 0 & -0.1305 \end{pmatrix}$$

Veremos ahora el caso de matrices rectangulares.

1 Factorizaciones de matrices rectangulares:

1.1 Factorización QR

Sea A una matriz $m \times n$ de rango n con $m > n$ (i.e. rango máximo en columnas). Después de n transformaciones de Householder se obtiene una matriz triangular superior:

$$H_n \dots H_2 H_1 A = Q^t A = \begin{pmatrix} R \\ 0 \end{pmatrix}$$

donde :

- H_i : matrices de Householder diseñadas para anular los elementos desde $i + 1$ hasta m de la i -ésima columna de la matriz parcialmente triangularizada
- R : matriz $n \times n$ triangular superior no singular
- Q : matriz ortogonal ($Q^t Q = I$), $Q = H_1 \dots H_n$

La expresión anterior es conocida como **factorización QR** de A .

$$A = \begin{pmatrix} 0.4087 & 0.1594 \\ 0.4302 & 0.3516 \\ 0.6246 & 0.3384 \end{pmatrix}$$

$$m = 3, \quad n = 2, \quad \text{rango}(A) = 2, \quad Q^t A = \begin{pmatrix} R \\ 0 \end{pmatrix}.$$

$$Q = \begin{pmatrix} -0.4744 & 0.5838 & -0.6589 \\ -0.4993 & -0.7948 & -0.3448 \\ -0.7250 & 0.1654 & 0.6686 \end{pmatrix}$$

$$\begin{pmatrix} R \\ 0 \end{pmatrix} = \begin{pmatrix} -0.8615 & -0.4965 \\ 0 & -0.1304 \\ 0 & 0 \end{pmatrix}$$

Si A es de rango $r < n$ pueden hacerse intercambios de columnas para asegurar que r columnas linealmente independientes se procesan primero. El resultado de r transformaciones por la izquierda es

$$Q^t AP = \begin{pmatrix} T \\ 0 \end{pmatrix}$$

donde P es una matriz de permutación y T es una matriz $r \times n$ trapezoidal superior:

$$T = \begin{pmatrix} t_{11} & \dots & \dots & \dots & t_{1r} & \dots & t_{1n} \\ 0 & t_{22} & \dots & \dots & t_{2r} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & \dots & \dots & t_{3n} \\ \vdots & \dots & \ddots & \ddots & \vdots & \dots & \vdots \\ 0 & \dots & \dots & 0 & t_{rr} & \dots & t_{rn} \end{pmatrix}$$

Puede aplicarse una sucesión de transformaciones de Householder \bar{H}_i por la derecha para anular las últimas $n - r$ columnas de T mientras se preserva la estructura triangular de las primeras r columnas. De esta forma se obtiene

$$Q^t AP \bar{H}_r \dots \bar{H}_1 = Q^t AV = \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}$$

donde

- R : es una matriz $r \times r$ triangular superior
- V : es una matriz ortogonal $n \times n$

Esta forma se llama **factorización ortogonal completa** de A .

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 7 & 6 & 10 \\ 4 & 4 & 6 \\ 1 & 0 & 1 \end{pmatrix}$$

$$m = 4, \quad n = 3, \quad \text{rango}(A) = 2, \quad Q^t AV = \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}.$$

$$Q = \begin{pmatrix} -0.1925 & 0.7180 & 0.6689 & 0 \\ -0.8340 & -0.2602 & 0.0393 & -0.4851 \\ -0.5132 & 0.2289 & -0.3934 & 0.7276 \\ -0.0642 & -0.6036 & 0.6295 & 0.4851 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.3333 & 0.6667 & -0.6667 \\ -0.6667 & -0.6667 & -0.3333 \\ -0.6667 & 0.3333 & 0.6667 \end{pmatrix}$$

$$\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 15.5885 & -4.4264 & 0 \\ 0 & -1.1863 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Si se busca una factorización similar para una matriz A con rango total m en filas (es decir $m < n$), es más conveniente elegir la factorización de forma que :

$$AQ = (L \ 0)$$

con L triangular inferior, que es conocida como **factorización LQ**.

$$A = \begin{pmatrix} 0.4087 & 0.4301 & 0.6246 \\ 0.1594 & 0.3515 & 0.3384 \end{pmatrix}$$

$$m = 2, \quad n = 3, \quad \text{rango}(A) = 2, \quad AQ = (L \ 0).$$

$$Q = \begin{pmatrix} -0.4744 & -0.5839 & -0.6588 \\ -0.4993 & -0.7948 & -0.3449 \\ -0.7250 & 0.1653 & 0.6686 \end{pmatrix}$$

$$(L \ 0) = \begin{pmatrix} -0.8615 & 0 & 0 \\ -0.4965 & -0.1304 & 0 \end{pmatrix}$$

NOTA:

Sea $Q = (Q_r \ Q_{m-r})$ con Q_r matriz $m \times r$ y Q_{m-r} matriz $m \times (m-r)$. Las columnas de Q_r forman una base ortonormal para el espacio columna de A y las columnas de Q_{m-r} forman una base ortonormal para el correspondiente espacio de vectores ortogonales a las filas de A^t .

1.2 Factorización LU

La factorización LU se hace como para el caso de matrices cuadradas inversibles, utilizando eliminación gaussiana con pivote total.

Dada una matriz A de tamaño $m \times n$ y de rango r se pueden encontrar matrices:

- L : matriz $m \times m$ triangular inferior con unos en la diagonal
- U : matriz $m \times n$ trapezoidal superior y de rango r

tales que

$$\tilde{P}AP = LU$$

siendo \tilde{P} y P matrices de permutación.

Además las matrices L y U tienen la forma :

$$L = \begin{pmatrix} L_{11} & 0 \\ L_{21} & I_{m-r} \end{pmatrix}$$

con

- L_{11} : matriz $r \times r$ triangular inferior con unos en la diagonal
- L_{21} : matriz $(m - r) \times r$
- I_{m-r} : matriz identidad $(m - r) \times (m - r)$

y

$$U = \begin{pmatrix} U_{11} & U_{12} \\ 0 & 0 \end{pmatrix}$$

con

- U_{11} : matriz $r \times r$ triangular superior inversible
- U_{12} : matriz $r \times (n - r)$

Esta factorización admite una forma reducida:

$$\tilde{P}AP = \begin{pmatrix} L_{11} \\ L_{21} \end{pmatrix} (U_{11} \ U_{12}) = \tilde{L}\tilde{U}$$

donde \tilde{L} y \tilde{U} tienen rango r pues $LU = \tilde{L}\tilde{U}$.

1.3 Actualizaciones de factorizaciones de matrices

1. *Ejemplo:*

Consideremos la factorización QR de una matriz \bar{A} generada añadiendo una columna a al final de la matriz $m \times n$ A .

Entonces,

$$Q^t \bar{A} = (Q^t A \quad Q^t a) = \begin{pmatrix} R & v_1 \\ 0 & v_2 \end{pmatrix}$$

donde v_1 y v_2 son las particiones correspondientes del vector $v = Q^t a$.

Sea H_{n+1} la matriz de Householder que anula los elementos de v que van desde $n + 2$ hasta m y deja los primeros n elementos de v sin cambiar, así que:

$$H_{n+1} v = \begin{pmatrix} v_1 \\ \gamma \\ 0 \end{pmatrix}$$

con $|\gamma| = \|v_2\|_2$.

Cuando H_{n+1} se aplica a $Q\bar{A}$ se obtiene:

$$H_{n+1} Q^t \bar{A} = \begin{pmatrix} R & v_1 \\ 0 & \gamma \\ 0 & 0 \end{pmatrix}$$

y así llamando $\bar{Q}^t = H_{n+1} Q^t$ y $\bar{R} = \begin{pmatrix} R & v_1 \\ 0 & \gamma \end{pmatrix}$, se tiene

$$\bar{Q}^t \bar{A} = \begin{pmatrix} \bar{R} \\ 0 \end{pmatrix}$$

que define la factorización QR de \bar{A} . Observemos que esa factorización se ha obtenido calculando únicamente una matriz de Householder.

2. Ejemplo:

Otras modificaciones comunes que suelen considerarse aparecen cuando se necesita la factorización de Cholesky de una matriz \bar{B} obtenida a partir de una matriz B simétrica definida positiva:

$$\bar{B} = B + vv^t$$

Entonces

$$\bar{B} = LDL^t + vv^t = L(D + pp^t)L^t$$

donde p es la solución del sistema $Lp = v$.

La especial naturaleza de la matriz $D + pp^t$ implica que sus factores Cholesky (que denotaremos por \tilde{L} y \tilde{D}) pueden calcularse directamente. Entonces:

$$\bar{B} = L\tilde{L}\tilde{D}\tilde{L}^tL^t = \bar{L}\bar{D}\bar{L}^t$$

con $\bar{L} = L\tilde{L}$ y $\bar{D} = \tilde{D}$. El vector p puede calcularse al tiempo que se hace la multiplicación de L y \tilde{L} .

2 Descomposición en valores singulares:

2.1 Descomposición espectral de una matriz simétrica

Cuando una matriz A es **simétrica** sus autovalores $\{\lambda_1, \dots, \lambda_n\}$ son todos **reales** y puede encontrarse una base ortonormal $\{u_1, \dots, u_n\}$ de \mathbb{R}^n formada por autovectores de A (i.e., $Au_i = \lambda_i u_i$).

Definiendo

$$U = (u_1 | \dots | u_n)$$

y $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ se tiene

$$AU = U\Lambda$$

ó, equivalentemente, al ser U ortogonal:

$$A = U\Lambda U^t$$

Explícitamente,

$$A = \sum_{i=1}^n \lambda_i u_i u_i^t$$

Esto es lo que se llama **descomposición espectral** de A .

2.2 Valores singulares

Cualquier matriz real $m \times n$ puede escribirse como:

$$A = U\Sigma V^t$$

donde U es una matriz ortogonal $m \times m$, V es una matriz ortogonal $n \times n$ Σ es una matriz diagonal $m \times n$ con $\sigma_i \geq 0, \forall i$:

- si $m > n$:

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_n \\ 0 & \dots & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

- si $m \leq n$:

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \ddots & \vdots & \vdots & \dots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \dots & \vdots \\ 0 & \dots & 0 & \sigma_m & 0 & \dots & 0 \end{pmatrix}$$

Los números σ_i son llamados **valores singulares de A** y se ordenan $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$.

Si A es de rango r entonces $\sigma_r > 0$ y $\sigma_{r+1} = 0$ (i.e. $\sigma_{r+i} = 0, \forall i \geq 1$), de manera que una matriz de rango r tiene r valores singulares no nulos.

Los valores singulares de A satisfacen:

- $\sigma_i^2(A) = \lambda_i(A^t A)$ si $m \geq n$
- $\sigma_i^2(A) = \lambda_i(AA^t)$ si $m \leq n$

Además $\sigma_1(A) = \|A\|_2$, es decir, la norma 2 de A es el mayor valor singular. Si A es simétrica sus valores singulares coinciden con los valores absolutos de los autovalores de A , por tanto $\sigma_1(A)$ es el radio espectral de A .

2.3 La pseudo-inversa

Cuando A es no singular A^{-1} es la única matriz tal que el vector $x = A^{-1}b$ resuelve el sistema $Ax = b$.

Si A es singular o rectangular la inversa no existe. Además el sistema $Ax = b$ puede ser incompatible. Sin embargo puede generalizarse el concepto de transformación inversa en términos de un problema de mínimos cuadrados:

La **pseudo-inversa** de la matriz $m \times n$, A se denota por A^+ y es la única matriz $n \times m$ tal que $x = A^+b$ es el vector de menor norma euclídea que minimiza $\|Ax - b\|_2$.

Cuando A tiene **rango total en columnas** la pseudo-inversa puede escribirse como:

$$A^+ = (A^t A)^{-1} A^t$$

Sin embargo, no debe calcularse con esta expresión. Si hemos calculado la descomposición QR de A :

$$A^+ = R^{-1} Q_n^t$$

donde $Q = (Q_n \ Q_{m-n})$.

Cuando A es de **rango deficiente** la forma más conveniente de calcular la pseudo-inversa está basada en la descomposición en valores singulares. Si $A = U \Sigma V^t$ con r valores singulares no nulos, entonces

$$A^+ = V \Omega U^t$$

donde $\Omega = \text{diag}(\omega_i)$ con $\omega_i = 1/\sigma_i$, $i = 1, \dots, r$ y $\omega_i = 0$ si $i \geq r + 1$.

3 Sistemas compatibles:

Sea A una matriz $m \times n$ con **rango total en filas** (así $m \leq n$).

En consecuencia las columnas de A generan todo \mathbb{R}^m y por tanto el sistema $Ax = b$ es compatible sea cual sea $b \in \mathbb{R}^m$.

Además A debe contener al menos un subconjunto de m columnas linealmente independientes. Estas columnas constituyen una submatriz regular $m \times m$ de A que puede ser usada para resolver $Ax = b$:

*Sea B una matriz no singular compuesta por columnas de A especificada por un **conjunto básico** de m índices distintos $\mathcal{B} = \{\beta_1, \dots, \beta_m\}$ donde la columna j de B es la columna β_j de A .*

Dado un conjunto básico \mathcal{B} y su matriz asociada B podemos resolver $Ax = b$ como sigue:

Como B es no singular el vector $x_B \in \mathbb{R}^m$ que cumple $Bx_B = b$ es único.

Una solución x del sistema original puede definirse de manera que x_B rellene las componentes de x designadas por índices de \mathcal{B} y todas las demás componentes de x sean cero, es decir:

- $x_i = (x_B)_j$ si $i = \beta_j$ para algún $j \in \{1, \dots, m\}$
- $x_i = 0$ en otro caso

*Así, $Ax = Bx_B = b$. La solución x construida de esta forma se llama **solución básica** de $Ax = b$.*

• **Ejemplo:**

Sea

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

una matriz 2×4 de rango 2 (rango total en filas).

El conjunto $\mathcal{B}_1 = \{1, 2\}$ es básico con:

$$B_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Otro conjunto básico es $\mathcal{B}_2 = \{1, 3\}$ con:

$$B_2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

El conjunto $\{1, 4\}$ no es básico pues la columna 1 y la 4 son dependientes.

Vamos a buscar soluciones básicas del sistema $Ax = b$, siendo A la matriz anterior y $b = \begin{pmatrix} 12 \\ 5 \end{pmatrix}$ Si elegimos como

conjunto básico $\mathcal{B}_1 = \{1, 2\}$ y resolvemos $B_1 x_{B_1} = b$ entonces $x_{B_1} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$ y una solución (básica) del sistema

$Ax = b$ es $x_1 = \begin{pmatrix} 7 \\ 5 \\ 0 \\ 0 \end{pmatrix}$ Si elegimos como conjunto básico

$\mathcal{B}_2 = \{1, 3\}$ y resolvemos $B_2 x_{B_2} = b$ entonces $x_{B_2} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$

y una solución (básica) del sistema $Ax = b$ es $x_2 = \begin{pmatrix} 2 \\ 0 \\ 5 \\ 0 \end{pmatrix}$

Sea A una matriz $m \times n$ con **rango total en columnas** (así $m \geq n$). Las n columnas son entonces linealmente independientes.

La estrategia de encontrar una submatriz no singular también se utiliza.

En este caso, sin embargo, el sistema no tiene que ser necesariamente compatible. *Si es compatible, la solución es única.*

Como A es una matriz con rango total en columnas contiene al menos un subconjunto de n filas linealmente independientes que constituyen una submatriz B no singular de tamaño $n \times n$.

El conjunto básico \mathcal{B} denota el conjunto de n índices distintos

$$\mathcal{B} = \{\beta_1, \dots, \beta_n\}$$

donde la fila i de B es la fila β_i de A .

Supongamos que el sistema es **compatible** entonces b está en la imagen de A .

Definimos el subvector b_B como:

$$(b_B)_j = b_l \text{ si } l = \beta_j, \quad 1 \leq j \leq n$$

La única solución x de $Ax = b$ se obtiene resolviendo el sistema no singular $Bx = b_B$.

- **Ejemplo:**

Sea

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

una matriz 3×2 de rango 2 (rango total en columnas).

y sea

$$b = \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix}$$

El sistema $Ax = b$ es compatible y las filas 2 y 3 de la matriz A constituyen una matriz no singular. Entonces:

$$\mathcal{B}_1 = \{2, 3\}, \quad B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, \quad b_B = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

la solución x satisface $Bx = b_B$ y así $x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

3.1 Solución de norma mínima de $Ax = b$

Sea A una matriz $m \times n$ de rango $r < n$ (si $r = n$ la solución es única).

Definimos x_+ **solución de norma mínima** de $Ax = b$ como aquella que cumple:

- $Ax_+ = b$
- $\|x_+\|_2 < \|x\|_2$ para cualquier otra x que satisfice $Ax = b$.

Cualquier vector no nulo de \mathbb{R}^n puede expresarse de manera única en términos de las componentes en el rango de A^t y el núcleo de A .

En particular, para cualquier vector tal que $Ax = b$, podemos escribir

$$x = x_R + x_N$$

con $\|x\|_2^2 = \|x_R\|_2^2 + \|x_N\|_2^2$ donde $x_R = A^t x_A$, $x_R \neq 0$ si $b \neq 0$ y $Ax_N = 0$ y $x_R^t x_N = 0$.

Un vector x que satisfice $Ax = b$ y que está enteramente en el rango de A^t (i.e. $x = x_R, x_N = 0$) tiene norma euclídea mínima.

Esto nos conducirá a un método para construir la solución de norma mínima:

Cualquier matriz A $m \times n$ de rango r puede escribirse de la forma:

$$A = GH$$

con

- G , matriz $m \times r$ de rango r
- H , matriz $r \times n$ de rango r

De esta forma G^tG y HH^t son no singulares, las columnas de G constituyen una base para el rango de A , y además H y A tienen el mismo núcleo. Por supuesto, la descomposición anterior no es única.

Suponemos por un momento la existencia de x_+ que satisface

$$Ax_+ = b$$

y que está en el rango de A^t , es decir

$$x_+ = A^t v$$

Como b está en el rango de A entonces b está en el rango de G . Puesto que G tiene rango total en columnas, $b = Gs$ para un vector $s \in \mathbb{R}^r$ único.

Sustituyendo $A = GH$, $x_+ = A^t v$, $b = Gs$, se tiene:

$$Ax_+ = AA^t v = GHH^t G^t v = Gs = b$$

De esta manera $x_+ = H^t z$ donde $HH^t z = s$ y $Gs = b$ (pues $z = G^t v$).

ALGORITMO RESUMEN:

- Paso 1: Calcular s que resuelve: $Gs = b$
- Paso 2: Hacer HH^t y calcular z tal que $HH^t z = s$
- Paso 3: Calcular $x_+ = H^t z$

Para obtener las matrices G y H pueden utilizarse como factorizaciones la LU, la QR, ...

EJEMPLO:

Utilizando la factorización QR de A y considerando:

$$Q = (Q_r \ Q_{m-r})$$

se tiene,

$$AP = Q_r R$$

y por tanto puede elegirse $G = Q_r$ y $H = RP^t$ y el algoritmo será:

- Paso 1: Calcular s que resuelve: $Q_r s = b$ o, equivalentemente, $s = Q_r^t b$ pues $Q_r^t Q_r = I$.
- Paso 2: Hacer RR^t y calcular z tal que $RR^t z = s$
- Paso 3: Calcular $x_+ = PR^t z$

Esta expresión es **numéricamente insatisfactoria** pues si A es mal condicionado RR^t tiene como condicionamiento el cuadrado del condicionamiento de A (pues $cond(R) = cond(A)$).

Para evitar esto, si $m < n$, usamos los factores QR de A^t :

$$A^t = Q \begin{pmatrix} R \\ 0 \end{pmatrix} P^t$$

con $R = (R_{11} \ R_{12})$, siendo R_{11} triangular superior de rango r .

Entonces

$$A^t = Q_r R P^t = Q_r (R_{11} \ R_{12}) P^t$$

y de esta forma:

$$A = P R^t Q_r^t = P \begin{pmatrix} R_{11}^t \\ R_{12}^t \end{pmatrix} Q_r^t$$

Entonces $G = P R^t$ y $H = Q_r^t$ y el algoritmo es:

- Paso 1: Calcular s que resuelve: $R_{11}^t s = (P^t b)_r$ (a causa de la unicidad s es solución de $R^t s = P^t b$)
- Paso 2: Como $H H^t = Q_r^t Q_r = I_r$ se tiene que $z = s$
- Paso 3: Calcular $x_+ = Q_r z = Q_r s$

APLICACIÓN:

Sean

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad y \quad \equiv \begin{pmatrix} 12 \\ 5 \end{pmatrix}$$

$$A^t = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 1 \\ 3 & 0 \end{pmatrix} = Q \begin{pmatrix} R \\ 0 \end{pmatrix} P^t$$

En este caso $P = I$ $P^t b = b$ y:

$$Q_r = \begin{pmatrix} -.2582 & .1690 \\ -.2582 & -.6761 \\ -.5164 & -.5071 \\ -.7746 & .5071 \end{pmatrix}$$

$$R_{11} = \begin{pmatrix} -3.873 & -.7746 \\ 0 & -1.183 \end{pmatrix}$$

En consecuencia resolviendo $R_{11}^t s = b_r$, obtenemos $s = \begin{pmatrix} -3.098 \\ -2.197 \end{pmatrix}$

y la solución de norma mínima es $x_+ = \begin{pmatrix} .4286 \\ 2.286 \\ 2.714 \\ 1.286 \end{pmatrix}$

Para evitar el mal condicionamiento en el caso $m > n$ el procedimiento de elección es usualmente la factorización ortogonal completa:

$$A = Q \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^t = Q_r (R \ 0) V^t$$

De esta forma $G = Q_r$ y $H = (R \ 0) V^t$ y el algoritmo de cálculo de la solución de norma mínima es:

- Paso 1: Calcular s que resuelve: $Q_r s = b$ o, equivalentemente, $s = Q_r^t b$.
- Paso 2: Calcular v tal que $Rv = s$ (pues $HH^t z = \begin{pmatrix} RR^t \\ 0 \end{pmatrix} z = s$, y hacemos $v = R^t z$)
- Paso 3: Calcular $x_+ = V \begin{pmatrix} v \\ 0 \end{pmatrix}$

3.2 Condicionamiento de un sistema compatible

El condicionamiento de un problema refleja la sensibilidad de su solución exacta a los cambios en los datos.

Vamos a fijarnos solo en perturbaciones que conserven la compatibilidad.

En primer lugar consideraremos perturbaciones en el segundo miembro del sistema.

Sea x_+ la solución de norma mínima del sistema compatible

$$Ax = b$$

entonces

$$Ax_+ = b \quad \text{y} \quad \|b\| \leq \|A\| \|x_+\|$$

para cualquier norma subordinada.

Consideremos ahora la solución de norma mínima del sistema perturbado:

$$A(x_+ + \delta x_b) = b + \delta b$$

donde $b \neq 0$ y $\delta b \in \mathcal{R}(A)$.

Como $x_+ = A^+b$ la solución perturbada satisface:

$$x_+ + \delta x_b = A^+(b + \delta b)$$

de forma que

$$\delta x_b = A^+\delta b$$

Como consecuencia

$$\frac{\|\delta x_b\|}{\|x_+\|} \leq \|A\| \|A^+\| \frac{\|\delta b\|}{\|b\|} = \text{cond}(A) \frac{\|\delta b\|}{\|b\|}$$

El cambio relativo entre la solución exacta y la del sistema perturbado puede ser tan grande como la perturbación relativa en b multiplicada por el condicionamiento de A .

Cuando se perturba la matriz A por δA el sistema perturbado

$$(A + \delta A)x = b$$

sigue siendo compatible para todos los vectores b que estén en $\mathcal{R}(A)$ sólo si el espacio rango de la matriz de perturbación δA es un subespacio de $\mathcal{R}(A)$. Bajo esta suposición se tiene el siguiente resultado:

$$\frac{\|\delta x_A\|}{\|x_+ + \delta x_A\|} \leq \text{cond}(A) \frac{\|\delta A\|}{\|A\|}$$

Para ilustrar estos efectos , consideremos el siguiente ejemplo en el que A es de rango total en columnas pero mal condicionada.

- **Ejemplo:**

Sean

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1.00001 \\ 1 & 1.00001 \end{pmatrix} \quad \text{y} \quad b = \begin{pmatrix} 2 \\ 2.00001 \\ 2.00001 \end{pmatrix}$$

La única solución de $Ax = b$ es $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Los valores singulares de A son $\sigma_1 = 2.450$ y $\sigma_2 = 5.773 \times 10^{-6}$ y así $cond(A) = \frac{\sigma_1}{\sigma_2} = 4.243 \times 10^5$.

Sean b_1 y b_2 perturbaciones de b :

$$b_1 = b + \begin{pmatrix} 0.001 \\ 0.001 \\ 0.001 \end{pmatrix} = \begin{pmatrix} 2.001 \\ 2.00101 \\ 2.00101 \end{pmatrix}, \quad b_2 = b + \begin{pmatrix} -0.002 \\ 0.001 \\ 0.001 \end{pmatrix} = \begin{pmatrix} 1.998 \\ 2.00101 \\ 2.00101 \end{pmatrix}$$

donde $\frac{\|b_1 - b\|}{\|b\|} = 5 \times 10^{-4}$ y $\frac{\|b_2 - b\|}{\|b\|} = 7.07 \times 10^{-4}$. La

solución exacta y única x_1 de $Ax_1 = b_1$ es $x_1 = \begin{pmatrix} 1.001 \\ 1 \end{pmatrix}$ y la solución exacta y única x_2 de $Ax_2 = b_2$ es $x_2 = \begin{pmatrix} -299 \\ 301 \end{pmatrix}$.

Para b_1 el cambio relativo $\frac{\|x_1 - x\|}{\|x\|}$ es pequeño y similar en orden a $\frac{\|b_1 - b\|}{\|b\|}$ mientras que $\frac{\|x_2 - x\|}{\|x\|}$ es casi $cond(A)$ veces el valor de $\frac{\|b_2 - b\|}{\|b\|}$.

Vemos ahora que ocurre si perturbamos en este último sistema la matriz. Sea

$$\delta A = \begin{pmatrix} -0.001 & -0.001 \\ 0.0005 & 0.0005 \\ 0.0005 & 0.0005 \end{pmatrix} \text{ y así } A + \delta A = \begin{pmatrix} 0.999 & 0.999 \\ 1.0005 & 1.00051 \\ 1.0005 & 1.00051 \end{pmatrix}$$

con $\frac{\|\delta A\|}{\|A\|} = 7.07 \times 10^{-4}$. Como demostramos antes la

solución exacta de $Ax_2 = b_2$ es $x_2 = \begin{pmatrix} -299 \\ 301 \end{pmatrix}$. La

solución de $(A + \delta A)\hat{x}_2 = b_2$ es $\hat{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. En este caso,

$\frac{\|\hat{x}_2 - x_2\|}{\|\hat{x}_2\|}$ es $\text{cond}(A)$ veces $\frac{\|\delta A\|}{\|A\|}$.

4 Mínimos cuadrados lineales:

El problema de mínimos cuadrados lineales es un problema computacional de importancia primordial que, originalmente, apareció por la necesidad de ajustar a un modelo matemático lineal observaciones dadas. Con el fin de reducir la influencia de los errores en las observaciones deseáramos usar un número más grande de medidas que el número de parámetros desconocidos del modelo. El problema resultante es resolver un sistema de ecuaciones lineales sobredeterminado.

DOS EJEMPLOS:

1. Supongamos que un proceso tiene una entrada simple $t \in \mathbb{R}$ y una salida simple $y \in \mathbb{R}$. Realizamos un experimento con el proceso resultando un número de medidas que son mostradas en la tabla:

i	0	1	2
t_i	2	3	4
y_i	3	4	15

Queremos encontrar una recta dada por

$$y = mt + c$$

que ajuste los datos experimentales. En otras palabras queremos encontrar dos números m y c tales que:

$$y_i = mt_i + c, \quad i = 0, 1, 2.$$

Sin embargo, no tienen por qué existir m y c que satisfagan estas ecuaciones. Podemos intentar representar nuestro problema con un sistema de tres ecuaciones lineales de la forma:

$$\begin{cases} 2m + c = 3 \\ 3m + c = 4 \\ 4m + c = 15 \end{cases}$$

Escribamos este sistema como:

$$Ax = b$$

donde

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 4 \\ 15 \end{pmatrix} \quad x = \begin{pmatrix} m \\ c \end{pmatrix}$$

Como el rango de A es menor que el rango de la matriz ampliada $[A|b]$ el sistema es inconsistente y no existe solución.

Podemos buscar la recta que mejor ajuste los datos en el sentido de los **mínimos cuadrados** (es decir, minimizando la distancia euclídea). El problema es, entonces, encontrar

$$\min_{x \in \mathbb{R}^2} \|Ax - b\|_2^2 = \sum_{i=0}^2 (mt_i + c - y_i)^2$$

La solución en este caso es:

$$x = \begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} 6 \\ -10/3 \end{pmatrix}$$

Este método puede ser usado para ajustar los datos a otras funciones modelo si estas son de la forma:

$$f(t) = c_1g_1(t) + c_2g_2(t) + \dots + c_mg_m(t)$$

esto es, si la función modelo es lineal en los parámetros.

En nuestro ejemplo anterior las funciones componentes que definen $f(t)$ son $g_1(t) = t$ y $g_2(t) = 1$.

Si quisieramos utilizar un modelo cuadrático serían $g_1(t) = t^2$, $g_2(t) = t$ y $g_3(t) = 1$.

Vamos a considerar ahora otro ejemplo donde la función modelo es de tipo trigonométrico:

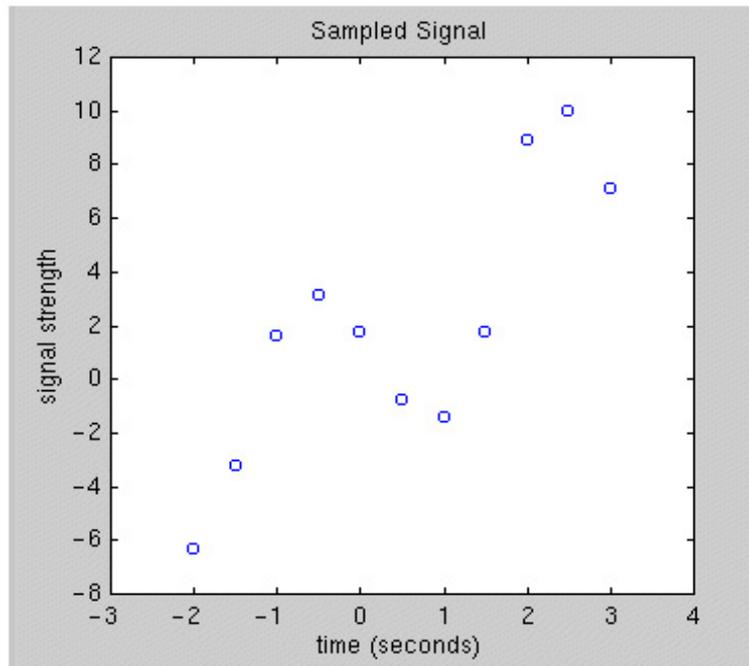
2. Los datos son medidas de una señal emitida por un dispositivo electrónico. La señal es muestreada cada medio segundo durante un cierto intervalo de tiempo. Los datos vienen dados en la siguiente tabla:

Tiempo (s)	Señal (mv)
t_i	y_i
-2.0	-6.32
-1.5	-3.23
-1.0	1.62
-0.5	3.13
0.0	1.74
0.5	-0.75
1.0	-1.41
1.5	1.78
2.0	8.88
2.5	9.98
3.0	7.10

Queremos encontrar una función modelo sinusoidal que suministre un buen ajuste de los datos observados. Debido a consideraciones teóricas basadas sobre propiedades físicas del dispositivo electrónico pensamos que un modelo adecuado es un polinomio trigonométrico de la forma

$$f(t) = a_0 + a_1 \text{sen}(t) + b_1 \text{cos}(t) + a_2 \text{sen}(2t) + b_2 \text{cos}(2t).$$

Queremos calcular los parámetros a_0 , a_1 , b_1 , a_2 y b_2 de forma que minimicen la suma de los cuadrados de los residuos.



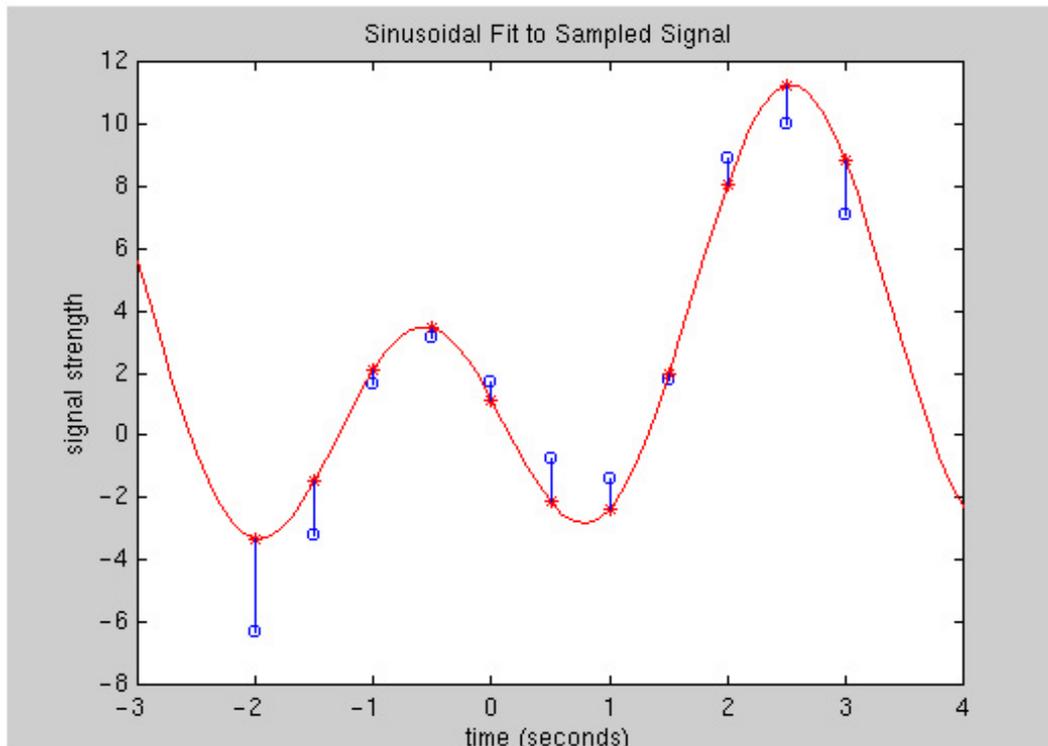
Nuestro problema es, entonces, de la forma

$$\min_{x \in \mathbb{R}^5} \|Ax - b\|_2^2 = \sum_{i=0}^{10} (f(t_i) - y_i)^2$$

donde

$$A = \begin{pmatrix} 1 & \text{sen}(t_1) & \text{cos}(t_1) & \text{sen}(2t_1) & \text{cos}(2t_1) \\ 1 & \text{sen}(t_2) & \text{cos}(t_2) & \text{sen}(2t_2) & \text{cos}(2t_2) \\ 1 & \text{sen}(t_3) & \text{cos}(t_3) & \text{sen}(2t_3) & \text{cos}(2t_3) \\ 1 & \text{sen}(t_4) & \text{cos}(t_4) & \text{sen}(2t_4) & \text{cos}(2t_4) \\ 1 & \text{sen}(t_5) & \text{cos}(t_5) & \text{sen}(2t_5) & \text{cos}(2t_5) \\ 1 & \text{sen}(t_6) & \text{cos}(t_6) & \text{sen}(2t_6) & \text{cos}(2t_6) \\ 1 & \text{sen}(t_7) & \text{cos}(t_7) & \text{sen}(2t_7) & \text{cos}(2t_7) \\ 1 & \text{sen}(t_8) & \text{cos}(t_8) & \text{sen}(2t_8) & \text{cos}(2t_8) \\ 1 & \text{sen}(t_9) & \text{cos}(t_9) & \text{sen}(2t_9) & \text{cos}(2t_9) \\ 1 & \text{sen}(t_{10}) & \text{cos}(t_{10}) & \text{sen}(2t_{10}) & \text{cos}(2t_{10}) \\ 1 & \text{sen}(t_{11}) & \text{cos}(t_{11}) & \text{sen}(2t_{11}) & \text{cos}(2t_{11}) \end{pmatrix}, \quad b = \begin{pmatrix} -6.32 \\ -3.23 \\ 1.62 \\ 3.13 \\ 1.74 \\ -0.75 \\ -1.41 \\ 1.78 \\ 8.88 \\ 9.98 \\ 7.10 \end{pmatrix}, \quad x = \begin{pmatrix} a_0 \\ a_1 \\ b_1 \\ a_2 \\ b_2 \end{pmatrix}$$

En la figura se puede observar la función sinusoidal obtenida, así como los residuos correspondientes.



En términos matriciales, dado un vector $b \in \mathbb{R}^m$ y una matriz $A \in \mathbb{R}^{m \times n}$, con $m > n$ queremos encontrar un vector $x \in \mathbb{R}^n$ tal que Ax sea la “mejor aproximación” a b . Existen distintos modos de definir la “mejor” solución. Una elección que puede motivarse por razones estadísticas y que conduce a un problema computacional simple es elegir x como solución del problema de minimización:

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$

donde $\|\cdot\|_2$ es la norma euclídea.

Este problema se conoce como problema de mínimos cuadrados lineal y x se llama **solución de mínimos cuadrados** del sistema lineal $Ax = b$. Llamaremos **residuo** a $r = b - Ax$.

4.1 Caracterización de las soluciones de los problemas de mínimos cuadrados

Llamaremos S al conjunto de las soluciones del problema de mínimos cuadrados:

$$S = \{x \in \mathbb{R}^n / \|Ax - b\|_2 = \min\}$$

entonces $x \in S$ si y sólo si se verifica la siguiente condición de ortogonalidad:

$$A^t(b - Ax) = 0$$

o equivalentemente, las **ecuaciones normales**:

$$A^tAx = A^tb$$

Como $r = b - Ax \in \mathcal{N}(A^t)$ para x solución de mínimos cuadrados de $Ax = b$ se tiene que ésta descompone (de manera única) el vector b en dos componentes ortogonales:

$$b = Ax + r, \quad Ax \in \mathcal{R}(A).$$

Señalemos que:

- La matriz $A^tA \in \mathbb{R}^{n \times n}$ es simétrica y semidefinida positiva.
- las ecuaciones normales tienen siempre solución pues $A^tb \in \mathcal{R}(A^t) = \mathcal{R}(A^tA)$
- La matriz A^tA es definida positiva si y sólo si el rango de A es n y en consecuencia si A tiene rango máximo en columnas la única solución de mínimos cuadrados es:

$$x = (A^tA)^{-1}A^tb$$

$$r = b - A(A^tA)^{-1}A^tb$$

- Si el rango de A es $< n$, entonces A tiene un núcleo no trivial y la solución de mínimos cuadrados no es única. Si \hat{x} es una solución particular, el conjunto de soluciones de mínimos cuadrados es:

$$S = \{x = \hat{x} + z / z \in \mathcal{N}(A)\}$$

Además si \hat{x} es ortogonal al $\mathcal{N}(A)$ entonces

$$\|x\|_2^2 = \|\hat{x}\|_2^2 + \|z\|_2^2$$

y en consecuencia \hat{x} es la única solución de norma mínima del problema de mínimos cuadrados.

NOTA:

El problema de calcular la solución de norma mínima $y \in \mathbb{R}^m$ para un sistema de ecuaciones indeterminado:

$$\min \|y\|_2, \text{ sujeto a } A^t y = c$$

*donde $A \in \mathbb{R}^{m \times n}$, aparece como un subproblema en los algoritmos de optimización. Si el rango de la matriz es n entonces el sistema $A^t y = c$ es consistente y la única solución del problema anterior está dado por las **ecuaciones normales de segundo tipo**:*

$$A^t A z = c, \quad y = A z$$

es decir, $y = A(A^t A)^{-1} c$

4.2 Cálculo de las soluciones de mínimos cuadrados

El método clásico para resolver las ecuaciones normales está basado en la factorización de Cholesky: Si $C \in \mathbb{R}^{n \times n}$ es simétrica y definida positiva, existe una única matriz triangular inferior B con elementos diagonales positivos tal que $C = BB^t$.

Desde el punto de vista de trabajo computacional utilizar las ecuaciones normales es eficiente, sin embargo, **su uso puede tener consecuencias indeseables respecto a la estabilidad numérica** ya que el número de condición de $A^t A$ es el cuadrado del número de condición de A . En consecuencia $A^T A$ está severamente mal condicionada aunque A sea moderadamente mal condicionada.

Para evitar esto se utilizan otros métodos para el cálculo de las soluciones de mínimos cuadrados que comentaremos brevemente.

Una herramienta poderosa para resolver el problema de mínimos cuadrados es la DESCOMPOSICIÓN EN VALORES SINGULARES:

Consideremos el problema de encontrar la solución de norma mínima del problema de mínimos cuadrados:

$$\min_{x \in S} \|x\|_2, \quad S = \{x \in \mathbb{R}^n / \|Ax - b\|_2 = \min\}$$

Este problema tiene una única solución que puede encontrarse usando la descomposición en valores singulares ($A = U\Sigma V^t$):

$$x_+ = V \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^t b = A^+ b$$

donde r es el rango de A .

APLICACIÓN DE LA DESCOMPOSICIÓN QR:

- Caso rango total:

Sea $A \in \mathbb{R}^{m \times n}$, $m \geq n$, $b \in \mathbb{R}^m$. Si el rango de A es n y $A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$ entonces la solución de norma mínima del problema de mínimos cuadrados se calcula según el algoritmo:

- Paso 1: $\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = Q^t b$

- Paso 2: Calcular x como solución de $Rx = d_1$. El residuo es $r = Q \begin{pmatrix} 0 \\ d_2 \end{pmatrix}$

- Caso rango $r < \min(m, n)$:

Para este caso se hace la descomposición ortogonal completa (que es equivalente a la descomposición en valores singulares):

$$A = Q \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^t$$

y se razona igual que para la descomposición en valores singulares. Entonces la solución de norma mínima del problema de mínimos cuadrados se obtiene como:

$$x = V \begin{pmatrix} R^{-1} & 0 \\ 0 & 0 \end{pmatrix} Q^t b$$

Si $Q = (Q_r \ Q_{m-r})$ habrá que ejecutar el siguiente algoritmo:

- Paso 1: Hacer $s = Q_r^t b$

- Paso 2: Calcular v como solución de $Rv = s$

- Paso 3: Calcular $x = V \begin{pmatrix} v \\ 0 \end{pmatrix}$

Como el número de condición de R es el mismo que el de A no aparece el cuadrado del número de condición de A al utilizar esta técnica. Esta es una ventaja respecto a la utilización de las ecuaciones normales.

DOS EJEMPLOS:

- Ejemplo 1:

Sea $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $b = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix}$. Como A tiene rango total la solución del problema de mínimos cuadrados es única. Además x es solución si y sólo si:

$$\begin{aligned} b - Ax \in \mathcal{N}(A^t) &= \{z / A^t z = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\} = \\ &= \{z / \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\} = \\ &= \{z / z_1 + z_2 = 0, z_1 + z_3 = 0\} = \langle \left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\} \rangle \end{aligned}$$

Por tanto $x = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ es solución si $b - Ax = \begin{pmatrix} \alpha_3 \\ -\alpha_3 \\ -\alpha_3 \end{pmatrix}$,

es decir si:

$$\begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \alpha_3 \\ -\alpha_3 \\ -\alpha_3 \end{pmatrix}$$

Entonces $\alpha_1 = 2$, $\alpha_2 = -3$, $\alpha_3 = 2$ y $x = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ es la solución de mínimos cuadrados.

- Ejemplo 2:

Sea $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{pmatrix}$ y $b = \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix}$. Como A es rango deficiente la solución del problema de mínimos cuadrados no es única. Vamos a calcularla utilizando las ecuaciones normales:

$$A^t A = \begin{pmatrix} 2 & 1 & \frac{3}{2} \\ 1 & 2 & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} & \frac{5}{2} \end{pmatrix}$$

Esta matriz es singular pero el sistema $A^t A x = A^t b = \begin{pmatrix} 1 \\ -4 \\ -\frac{3}{2} \end{pmatrix}$ es compatible y una solución es $x = \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix}$.

Otras soluciones serían $\begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -4 \\ 2 \end{pmatrix}, \dots$

Para calcular la solución de norma mínima podemos utilizar la factorización ortogonal completa y aplicar el último algoritmo visto.

4.3 Algoritmo de mínimos cuadrados recursivo

(Recursive Least-Squares algorithm, RLS)

Consideramos el primer ejemplo de la sección. Tenemos puntos experimentales (t_0, y_0) , (t_1, y_1) , (t_2, y_2) y queremos encontrar los parámetros m y c de la línea recta que mejor ajusta estos datos en el sentido de los mínimos cuadrados.

Supongamos que tenemos una nueva medida (t_3, y_3) y, así un nuevo conjunto de datos experimentales (t_0, y_0) , (t_1, y_1) , (t_2, y_2) , (t_3, y_3) . Podemos repetir el procedimiento y obtener la línea que ajusta mejor los cuatro puntos. Sin embargo existe un modo **más eficiente**:

Podemos utilizar los cálculos previos de m y c para los tres primeros datos para obtener los nuevos valores para los cuatro datos. Este procedimiento es conocido como **algoritmo RLS**.

Vamos a derivar este algoritmo. Primero consideramos el problema:

$$\min_{x \in \mathbb{R}^n} \|A_0 x - b^{(0)}\|_2$$

donde $n=2$ y A_0 y $b^{(0)}$ corresponden a los 3 primeros datos. Sabemos que la solución está dada por:

$$x^{(0)} = G_0^{-1} A_0^t b^{(0)}$$

donde $G_0 = A_0^t A_0$.

Si tenemos más datos en forma de matriz A_1 y vector $b^{(1)}$, consideraremos el problema siguiente:

$$\min_{x \in \mathbb{R}^n} \left\| \begin{pmatrix} A_0 \\ A_1 \end{pmatrix} x - \begin{pmatrix} b^{(0)} \\ b^{(1)} \end{pmatrix} \right\|_2$$

En este caso la solución está dada por

$$x^{(1)} = G_1^{-1} \begin{pmatrix} A_0 \\ A_1 \end{pmatrix}^t \begin{pmatrix} b^{(0)} \\ b^{(1)} \end{pmatrix}$$

donde $G_1 = \begin{pmatrix} A_0 \\ A_1 \end{pmatrix}^t \begin{pmatrix} A_0 \\ A_1 \end{pmatrix}$.

Vamos a tratar de escribir $x^{(1)}$ como función de $x^{(0)}$, G_0 y los nuevos datos A_1 y $b^{(1)}$.

Para ello tenemos en cuenta que G_1 puede escribirse como

$$G_1 = \begin{pmatrix} A_0^t & A_1^t \end{pmatrix} \begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = A_0^t A_0 + A_1^t A_1 = G_0 + A_1^t A_1$$

Por otra parte:

$$\begin{pmatrix} A_0 \\ A_1 \end{pmatrix}^t \begin{pmatrix} b^{(0)} \\ b^{(1)} \end{pmatrix} = \begin{pmatrix} A_0^t & A_1^t \end{pmatrix} \begin{pmatrix} b^{(0)} \\ b^{(1)} \end{pmatrix} = A_0^t b^{(0)} + A_1^t b^{(1)}$$

Además:

$$A_0^t b^{(0)} = G_0 G_0^{-1} A_0^t b^{(0)} = G_0 x^{(0)} = (G_1 - A_1^t A_1) x^{(0)} = G_1 x^{(0)} - A_1^t A_1 x^{(0)}$$

Combinando estas fórmulas se tiene:

$$x^{(1)} = G_1^{-1} \begin{pmatrix} A_0 \\ A_1 \end{pmatrix}^t \begin{pmatrix} b^{(0)} \\ b^{(1)} \end{pmatrix} = G_1^{-1} (G_1 x^{(0)} - A_1^t A_1 x^{(0)} + A_1^t b^{(1)})$$

Y por tanto:

$$x^{(1)} = x^{(0)} + G_1^{-1} A_1^t \underbrace{(b^{(1)} - A_1 x^{(0)})}_{\text{innovación}}$$

Generalizando, en la $(k+1)$ -ésima iteración haremos:

$$G_{k+1} = G_k + A_{k+1}^t A_{k+1}$$

$$x^{(k+1)} = x^{(k)} + G_{k+1}^{-1} A_{k+1}^t \underbrace{(b^{(k+1)} - A_{k+1} x^{(k)})}_{\text{innovación}}$$

Para calcular G_{k+1}^{-1} puede utilizarse una fórmula de actualización (generalización de la fórmula de Sherman-Morrison):

Sea A no singular y U, V matrices tales que $I + VA^{-1}U$ es no singular, entonces $A + UV$ es no singular y

$$(A + UV)^{-1} = A^{-1} - A^{-1}U(I + VA^{-1}U)^{-1}VA^{-1}$$

Usando esta fórmula para calcular G_{k+1}^{-1} se obtiene:

$$G_{k+1}^{-1} = G_k^{-1} - G_k^{-1}A_{k+1}^t(I + A_{k+1}G_k^{-1}A_{k+1}^t)^{-1}A_{k+1}G_k^{-1}$$

Por simplicidad notamos $P_k = G_k^{-1}$ y el **algoritmo RLS** puede escribirse como:

$$P_{k+1} = P_k - P_kA_{k+1}^t(I + A_{k+1}P_kA_{k+1}^t)^{-1}A_{k+1}P_k$$

$$x^{(k+1)} = x^{(k)} + P_{k+1}A_{k+1}^t(b^{(k+1)} - A_{k+1}x^{(k)})$$

En el caso especial en que los nuevos datos son tales que A_{k+1} es una matriz de una única fila $A_{k+1} = a_{k+1}^t$ y $b^{(k+1)}$ es un escalar $b^{(k+1)} = b_{k+1}$ se tiene:

$$P_{k+1} = P_k - \frac{P_k a_{k+1} a_{k+1}^t P_k}{1 + a_{k+1}^t P_k a_{k+1}}$$

$$x^{(k+1)} = x^{(k)} + P_{k+1} a_{k+1} (b_{k+1} - a_{k+1}^t x^{(k)})$$

4.4 Condicionamiento de un problema de mínimos cuadrados

Vamos a dar algunos resultados sobre la sensibilidad de las pseudoinversas y las soluciones de mínimos cuadrados respecto a las perturbaciones en A y b .

4.4.1 Análisis de perturbación de pseudoinversas

Damos primero algunas cotas para las perturbaciones de las pseudoinversas. Sea $A \in \mathbb{R}^{m \times n}$ y $B = A + E$ la matriz perturbada. La teoría se complica por el hecho de que A^+ varía discontinuamente cuando el rango de A cambia (por ejemplo la

pseudoinversa de un escalar σ es $\sigma^+ = \begin{cases} \frac{1}{\sigma} & \text{si } \sigma \neq 0, \\ 0 & \text{en otro caso.} \end{cases}$)

Si el rango de $A + E$ es distinto del rango de A , entonces $\|(A + E)^+ - A^+\|_2 \geq \frac{1}{\|E\|_2}$.

Cuando el rango cambia, la perturbación en A^+ puede ser no acotada cuando $\|E\|_2 \rightarrow 0$ por ejemplo:

$$A = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad E = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon \end{pmatrix}$$

donde $\sigma > 0$ y $\epsilon \neq 0$. En este caso A tiene rango 1 mientras que el rango de $A + E$ es 2.

Las pseudoinversas son:

$$A^+ = \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad (A + E)^+ = \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & \epsilon^{-1} \end{pmatrix}$$

y, en consecuencia $\|(A + E)^+ - A^+\|_2 = \|\epsilon\|^{-1} = \frac{1}{\|E\|_2}$.

Si la perturbación E no cambia el rango de A entonces no puede ocurrir el crecimiento no acotado de $(A + E)^+$. En concreto si el rango de $A + E$ es igual al rango de A y $\eta =$

$\|A^+\|_2\|E\|_2 < 1$ entonces

$$\|(A + E)^+\|_2 \leq \frac{\|A^+\|_2}{1 - \eta}$$

Una condición necesaria y suficiente para que $\lim_{E \rightarrow 0} (A + E)^+ = A^+$ es que $\lim_{E \rightarrow 0} \text{rango}(A + E) = \text{rango}(A)$.

4.4.2 Análisis de perturbación de las soluciones de mínimos cuadrados

Consideramos ahora el efecto de las perturbaciones de A y b sobre la solución de norma mínima del problema de mínimos cuadrados ($x = A^+b$). En este análisis el número de condición de una matriz rectangular $A \in \mathbb{R}^{m \times n}$ juega un papel significativo. La siguiente definición generaliza el número de condición de una matriz cuadrada no singular:

$$\text{cond}(A) = \|A\|_2\|A^+\|_2 = \frac{\sigma_1}{\sigma_r}$$

donde $0 < r = \text{rango}(A) \leq \min(m, n)$ y $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ son los valores singulares distintos de cero de A .

Consideramos en primer lugar el efecto de las perturbaciones en el segundo miembro. Llamamos $\tilde{b} = b + \delta b$ al segundo miembro perturbado.

Un hecho crucial de la descomposición $b = b_R + b_N$, donde $b_R \in \mathcal{R}(A)$ y $b_N \in \mathcal{N}(A^t)$ es que indica como los cambios en b afectan a la solución y al residuo:

1. Si $\delta b \in \mathcal{R}(A)$ y $\tilde{b} = b + \delta b$ entonces la solución de

$$\min_{x \in \mathbb{R}^n} \|Ax - \tilde{b}\|_2$$

es $\tilde{x} = x + \delta x$ y el residuo $\tilde{r} = r$.

2. Si $\delta b \in \mathcal{N}(A^t)$ y $\tilde{b} = b + \delta b$ entonces la solución de

$$\min_{x \in \mathbb{R}^n} \|Ax - \tilde{b}\|_2$$

es $\tilde{x} = x$ y el residuo $\tilde{r} = r + \delta b$.

Para el problema

$$\min_{x \in \mathbb{R}^n} \|Ax - \tilde{b}\|_2$$

la solución de norma mínima es $x + \delta x = A^+(b + \delta b)$ y entonces $\delta x = A^+(\delta b_R)$, de donde se obtiene:

$$\|\delta x\| \leq \|A^+\| \|\delta b_R\|$$

pues $A^+(b_R + b_N + \delta b_R + \delta b_N) = A^+b_R + A^+\delta b_R$, ya que A^+ anula cualquier vector en el núcleo de A^t .

Por otra parte sabemos que si x es una solución de mínimos cuadrados $Ax = b_R$ y, por tanto,

$$\|b_R\| \leq \|A\| \|x\|$$

Si $\|b_R\| \neq 0$, podemos combinar estas dos desigualdades y obtener:

$$\frac{\|\delta x\|}{\|x\|} \leq \|A^+\| \|A\| \frac{\|\delta b_R\|}{\|b_R\|} = \text{cond}(A) \frac{\|\delta b_R\|}{\|b_R\|}$$

EJEMPLO:

Sean:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1.00001 \\ 1 & 1.00001 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ .00001 \\ 4.00001 \end{pmatrix}$$

Recordemos que $cond(A) = \frac{\sigma_1}{\sigma_2} = 4.243 \times 10^5$.

La única solución de norma mínima del problema de mínimos cuadrados lineales asociado a A y b es $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y su residuo

$$\text{es } r = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}.$$

Si $\delta b = (.00001, -.00001, .00001)^t$, la única solución de norma mínima del problema de mínimos cuadrados lineales asociado a A y $b + \delta b$ es $x + \delta x = \begin{pmatrix} 2.00001 \\ 0 \end{pmatrix}$ y $\delta x = \begin{pmatrix} 1.00001 \\ -1 \end{pmatrix}$

Para esta elección de δb , sus componentes en $\mathcal{R}(A)$ y $\mathcal{N}(A^t)$ son, respectivamente:

$$\delta b_R = \begin{pmatrix} 0.00001 \\ 0 \\ 0 \end{pmatrix} \quad \text{y} \quad \delta b_N = \begin{pmatrix} 0 \\ -0.00001 \\ 0.00001 \end{pmatrix}$$

De esta forma $\|\delta b_R\| = 10^{-5}$, $\|\delta b_N\| = 1.4 \times 10^{-5}$, y por tanto $\frac{\|\delta b_R\|}{\|b_R\|} = 2.9 \times 10^{-6}$.

Entonces, en este caso, el cambio δx alcanza la cota superior de la desigualdad anterior con:

$$\frac{\|\delta x\|}{\|x\|} \simeq 1.0 \simeq cond(A) \frac{\|\delta b_R\|}{\|b_R\|}$$

Vamos a considerar ahora el efecto de perturbaciones sobre la matriz. Sea $x + \delta x$ solución de un problema de mínimos cuadrados con matriz perturbada $A + E$, entonces:

$$x + \delta x = (A + E)^+ b$$

Suponemos que E no altera el rango, es decir $\text{rango}(A) = \text{rango}(A + E)$. La mayor parte del análisis conlleva cuantificar los cambios en los espacios imagen y núcleo de $A + E$ comparados con los de A .

Una medida de estos cambios puede obtenerse utilizando bases ortonormales para $\mathcal{R}(A)$ y $\mathcal{N}(A^t)$. Si M y N denotan bases ortonormales para $\mathcal{R}(A)$ y $\mathcal{N}(A^t)$, se definen:

$$E_R = M^t E, \quad E_N = N^t E, \quad e_R = \frac{\|E_R\|}{\|A\|}, \quad e_N = \frac{\|E_N\|}{\|A\|}$$

Hablando en sentido general, e_R mide el efecto relativo de la perturbación E sobre la imagen de A , y e_N mide el efecto análogo de E sobre el núcleo de A^t . Como antes, escribimos $b = b_R + b_N$ y suponemos que $b_R \neq 0$. Puede probarse que el cambio relativo en la solución exacta satisface:

$$\frac{\|\delta x\|}{\|x\|} \leq 2 \text{cond}(A) e_R + 4 \text{cond}(A)^2 \frac{\|b_N\|}{\|b_R\|} e_N + O(e_N^2)$$

Un punto esencial a notar sobre esta cota es la aparición del factor $\text{cond}(A)^2$ en el segundo miembro. Sin embargo, dado que este factor multiplica el producto de e_N por el cociente de $\|b_N\|$ y $\|b_R\|$, el cuadrado del número de condición no afectará a la solución de mínimos cuadrados si E no modifica $\mathcal{N}(A^t)$, o si b no tiene componente en $\mathcal{N}(A^t)$ (o es una componente despreciable).

EJEMPLO:

Sea, de nuevo, la matriz A del ejemplo anterior y consideremos la perturbación:

$$E_1 = \begin{pmatrix} .00001 & 0 \\ .00001 & 0 \\ .00001 & 0 \end{pmatrix}$$

Si notamos $A_1 = A + E_1$, la solución exacta del problema de mínimos cuadrados lineales asociado a A_1 y b es $x_1 = \begin{pmatrix} .99999 \\ 1. \end{pmatrix}$ con

$$\frac{\|x_1 - x\|}{\|x\|} = 0.71 \times 10^{-6} \simeq \frac{\|E_1\|}{\|A\|}$$

La segunda perturbación que consideraremos es:

$$E_2 = \begin{pmatrix} 0 & -.00002 \\ 0 & .00001 \\ 0 & .00001 \end{pmatrix}$$

y definimos $A_2 = A + E_2$. La solución exacta del problema de mínimos cuadrados lineales asociado a A_2 y b es $x_2 = \begin{pmatrix} 1.75 \\ .25 \end{pmatrix}$ con

$$\frac{\|x_2 - x\|}{\|x\|} = 0.60 \simeq \text{cond}(A) \frac{\|E_2\|}{\|A\|}$$

Lo que ocurre cuando aparece el número de condición al cuadrado se muestra en el siguiente caso. Consideremos:

$$E_3 = \begin{pmatrix} 0 & 0 \\ 0 & -.00001 \\ 0 & .00001 \end{pmatrix}$$

y notemos $A_3 = A + E_3$. La solución exacta del problema de mínimos cuadrados lineales asociado a A_3 y b es $x_3 = \begin{pmatrix} -1.5 \times 10^5 \\ 1.5 \times 10^5 \end{pmatrix}$

con

$$\frac{\|x_3 - x\|}{\|x\|} = 1.5 \times 10^5 \simeq \text{cond}(A)^2 \frac{\|E_3\|}{\|A\|}$$

Bibliografía básica:

- A. BJORCK : Numerical Methods for Least Squares Problems. SIAM, 1996.
- P.E. GILL - W. MURRAY - M. WRIGHT : Numerical linear Algebra and Optimization. Addison-Wesley, 1991.
- C. LAWSON - R. HANSON : Solving Least Squares Problems. Classics in Applied Mathematics 15, SIAM, 1995.