Semana del 21/10

Monday, October 21, 2024 10:41 PM

Mínimos Cuadrados.

La idea es que tenemos un montón de datos $\{(t_i, y_i)\}_{i=1,\dots,m} \subset \mathbb{R}^2$ y queremos una curva que "ajuste" los datos.

La diferencia con interpolación es que acá asumimos un modelo de como se comportan los datos.

Vamos a asumir que existe $f : \mathbb{R} \to \mathbb{R}$ tq

$$f(t_i) \approx y_i$$

No necesariamente debe ser así (Por ejemplo el entregable). (Si $f(t) = x_1 e^{-\alpha_1 t} + x_2 e^{-\alpha_2 t} + \dots + x_n e^{-\alpha_n t}$ No es lineal en los α_i's)

Nos vamos a concentrar en problemas lineales: Tenemos $\phi_1(t), \dots, \phi_n(t)$ n funciones LI, y queremos

$$f(t) = \sum_{j=1}^{n} x_j \phi_j(t)$$

Donde $\{x_j\}$ se elige para que f "ajuste" lo mejor posible.

Al igual que en interpolación ponemos

$$\begin{cases} y_1 = x_1 \phi_1(t_1) + x_2 \phi_2(t_1) + \dots + x_n \phi_n(t_1) \\ \vdots \qquad \ddots \qquad \vdots \end{cases}$$

$$\begin{cases} \vdots \\ y_m = x_1\phi_1(t_m) + x_2\phi_2(t_m) + \dots + x_n\phi_n(t_m) \end{cases}$$

A la matriz $a_{ij} = \phi_j(t_i)$ le llamamos matriz de diseño. $(A \in M_{m\times n}(\mathbb{R}))$

Si A es la matriz de diseño, $y = (y_1, \dots, y_m)^t$

$\Rightarrow$ el sistema se puede escribir como $Ax = y$

Notar que el sistema no es cuadrado y además esperamos que $m > n$, y en ese caso esperamos que sea incompatible. $(\underbrace{\dim(\operatorname{Im} A)}_{= Rg(A)} \leq n)$

Def. $x \in \mathbb{R}^n$ es solución del sistema $Ax = y$ en el sentido de mínimos cuadrados si minimiza la norma euclídea de $r(x) = Ax - y$, i.e.,

$$x \in \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - y\|_2^2 .$$

$$\phi(x) = \|Ax - y\|_2^2$$

Ecuaciones normales

x^* es sol. de $Ax = y$ en el sentido de mínimos cuadrados sii.

$$\underbrace{A^TAx^* = A^Ty}_{(*)}, \iff A^T(Ax^* - y) = 0$$

- No hay unicidad! Salvo que A sea de rango completo $(Rg(A) = n)$.

- Notar que cumplir $(*)$ es que $\underline{Ax^* - y}$ sea

• Notar que cumplir (**) es que $\underline{Ax^* - y}$ sea ortogonal al espacio generado por las columnas de A.

$$\underset{m\times n}{\begin{pmatrix} A \end{pmatrix}} = \underset{m\times n}{\begin{pmatrix} Q \end{pmatrix}} \underset{n\times n}{\begin{pmatrix} R \end{pmatrix}}$$

Descomposición QR.

A de rango completo $\Rightarrow A = QR$ con (Desc. Económica)

$Q \in \mathbb{R}^{m\times n}$ tq sus columnas forman un conjunto ortonormal en $\mathbb{R}^m$

$R \in \mathbb{R}^{n\times n}$ triangular superior e invertible.

Q es el resultado de ortonormalizar las columnas de A con Gram-Schmidt, $Q = (u_1 | u_2 | \dots | u_n)$.

$R = (r_{ij}) \in \mathbb{R}^{n\times n}$ como

$$r_{ij} = \begin{cases} \langle A^{(i)}, u_j \rangle & \text{si } i \leq j \\ 0 & \text{si } i > j. \end{cases}$$

Tambien está la QR completa, donde $Q \in \mathbb{R}^{m\times m}$ y es ortogonal, y $R \in \mathbb{R}^{m\times n}$

Completa:

$$\underset{m\times n}{\begin{pmatrix} A \end{pmatrix}} = \underset{m\times m}{\begin{pmatrix} Q \end{pmatrix}} \underset{m\times n}{\begin{pmatrix} R \end{pmatrix}}$$

Esto se hace completando las columnas de la Q económica a una base ortonormal y a R agregandole ceros despues de la fila $n+1$:

$$A = QR = [\overbrace{Q^{(1:n)}}^{Q \text{ económica}} \underbrace{Q^{(n+1:m)}}_{\text{Vectores que completo}}] \begin{pmatrix} \overbrace{R_{(1:n)}}^{R \text{ económica}} \\ \underbrace{0_{(m-n)\times n}}_{\text{Matriz}} \end{pmatrix} = Q^{(1:n)} R^{(1:n)} + \cancel{Q^{(n+1:n)} 0_{(m-n)}} \to 0$$

$$= Q^{(1:n)} R_{(1:n)}$$

Vectores que completo

Matriz de ceros.

$+ \cancel{Q^{(n+1:n)} 0_{(m-n)}}$

El comando qr de Octave produce la factorización completa. $qr(A) = [Q, R]$

Mínimos Cuadrados y QR.

El residuo $\underbrace{r = Ax - y}$ se puede escribir como

$$\|r\|_2^2 = \underbrace{\|R_{(1:n)}x - (Q^T y)_{(1:n)}\|_2^2}_{\text{Tengo que minimizar esto.}} + \underbrace{\|(Q^T y)_{n+1:m}\|_2^2}_{\text{Es cte.}}$$

Entonces podemos minimizar el primer término de la izquierda, resolviendo el sistema:

$$\overbrace{Q^T}^{\in M_{m\times m}(\mathbb{R})} y \in \mathbb{R}^m$$

$$R_{(1:n)} x = (Q^T y)_{(1:n)} \qquad \Rightarrow Q^T y = \begin{bmatrix} (Q^T y)_{(1:n)} \\ (Q^T y)_{(n+1:m)} \end{bmatrix}$$

En resumen, dada la factorización $A = QR$, para hallar la solución de mínimos cuadrados del sistema $A\mathbf{x} = \mathbf{y}$:

- computamos el vector $Q^t\mathbf{y}$;
- resolvemos el sistema $R\mathbf{x} = (Q^t\mathbf{y})_{(1:n)}$ para hallar la solución de mínimos cuadrados;
- si nos interesa el valor mínimo de la norma euclídea del residuo, ésta es igual a $\|(Q^t\mathbf{y})_{(n+1:m)}\|_2$.

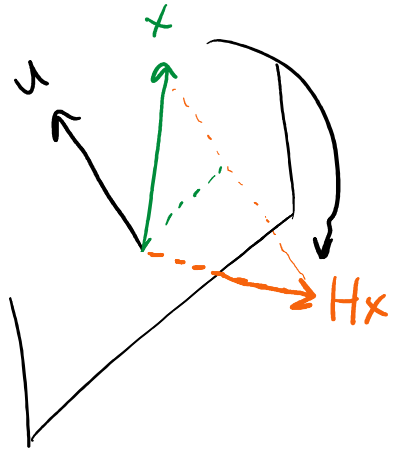
Transformaciones de Householder.

Dado $u \in \mathbb{R}^m \setminus \{0\}$ la reflexión de Householder asociada a u es la matriz $H \in \mathbb{R}^{m\times m}$

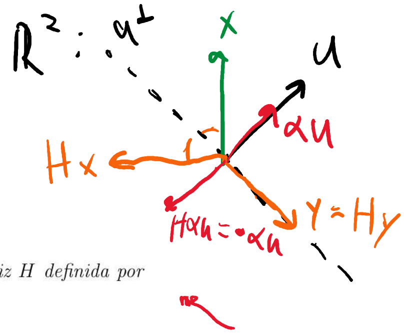
$$H = Id - 2\frac{uu^T}{\|u\|_2^2} \leftarrow (uu^T)_{ij} = u_i u_j$$

$(u_1 u_2 \ldots u_n) u^T$

$H = Id - 2\frac{uu^T}{\|u\|_2^2} \longleftarrow (uu^T)_{ij} = u_i u_j$



$$\underbrace{\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}}_{u} \overbrace{(u_1\, u_2 \dots u_n)}^{u^T} \begin{pmatrix} u_1u_1 & u_1u_2 & \cdots & u_1u_n \\ \vdots & & \ddots & \vdots \\ u_nu_1 & u_nu_2 & \cdots & u_nu_n \end{pmatrix}$$



Lema 5.3.2 (simetría y ortogonalidad). *Dado* $\mathbf{u} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$, *la matriz* H *definida por* (5.8) *es simétrica y ortogonal.*

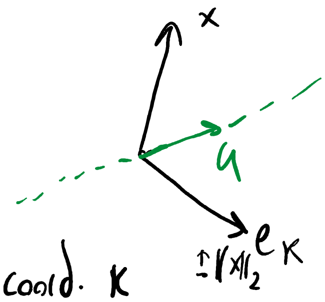
Se pueden usar para computar la desc. QR de forma estable.

Primero, si dado $x \in \mathbb{R}^m$ queremos $H = Id - \frac{uu^T}{\|u\|_2^2}$ t.q. $Hx = \alpha e_k = \begin{pmatrix} 0 \\ \vdots \\ \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k$ H es ortogonal:

$\Rightarrow |\alpha| = \|x\|_2$

Moraleja (Mirar las notas) basta tomar

$$\boxed{u = x \pm \|x\|_2 e_k}$$



Se elige + ó − dependiendo si la coord. k de x $(\langle x, e_k \rangle)$ es positiva o negativa. Para que la entrada k de u sea $\approx 2x_k$.

La idea va a ser aplicar $\underbrace{H_1, \dots, H_n}_{} \in \mathbb{R}^{m \times m}$ reflexiones t.q.

$$H_n \cdots H_1 A = R \text{ triangular superior.}$$

... como el col(A).

$$H_n \cdots H_1 A = R \text{ triangular superior.}$$

Voy a necesitar tantas como el rg(A).

Idea: Notamos como $A^{(1)}$ a la columna 1 de A.

• Definimos H_1 tq $\underbrace{H_1 A^{(1)}} = \pm \|A^{(1)}\|_2 e_1$, i.e.,

$$H_1 = Id - \frac{uu^T}{\|u\|_2^2} \quad \text{con} \quad u = A^{(1)} \pm \|A^{(1)}\|_2 e_1.$$

El signo lo tomamos para que sea el mismo que el de la primera entrada de $A^{(1)}$.

Entonces:

$$H_1 A = \begin{pmatrix} \pm\|A^{(1)}\|_2 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} = \left(\begin{array}{c|c} \pm\|A^{(1)}\|_2 & * \quad * \quad * \\ \hline 0 & \\ \vdots & \tilde{A}_2 \\ 0 & \end{array}\right)$$

$(m-1)\times(n-1)$

Y ahora hacemos lo mismo para $\tilde{A}_2 \in \mathbb{R}^{(m-1)\times(n-1)}$ obteniendo la reflexión $\tilde{H}_2 \in \mathbb{R}^{(m-1)\times(m-1)}$, y como no queremos deshacer lo anterior ponemos

$$H_2 = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \tilde{H}_2 \\ 0 & \end{array}\right) \in M_{m\times m}(\mathbb{R})$$

$$\Rightarrow H_2 H_1 A = \left(\begin{array}{cc|c} \pm\|A^{(1)}\|_2 & * & * \cdots * \\ 0 & \|A_2^{(1)}\|_2 & * \cdots * \\ \hline 0 & 0 & \\ \vdots & \vdots & \tilde{A}_3 \\ 0 & 0 & \end{array}\right)$$

$(m-2)\times(n-2)$

Y así seguimos (detalles en los notas)...

$\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$

Y así seguimos (detalles en los ...)...

$$\tilde{H}_3 \in M_{(m-2)\times(m-2)}(\mathbb{R}), \quad H_3 = \left(\begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \hline \vdots & 0 & & & \\ \vdots & \vdots & & \tilde{H}_3 & \\ 0 & 0 & & & \end{array}\right)$$

$$\underbrace{H_n \cdots H_1}_{=Q^{-1}} A = \underset{\substack{\uparrow \\ \text{Triag. sup}}}{R} \Rightarrow A = \underbrace{(H_n \cdots H_1)^{-1}}_{\text{Ortogonal}} R = QR$$

7) $\underbrace{(0,1), (\pi/2, 2), (\pi, 1), (3\pi/2, 1)}_{4 \text{ ptos}, \; m=4}$ y los queremos ajustar con

$$f(t) = \underbrace{a}\cos(x) + \underbrace{b}\operatorname{sen}(x) + \underbrace{c}\cos(4x)$$

Tres parámetros, $n=3$

a) $$\begin{cases} 1 = f(0) = a + 0 \cdot b + c \\ 2 = f(\pi/2) = 0 \cdot a + b + c \\ 1 = f(\pi) = -a + 0 \cdot b + c \\ 1 = f(3\pi/2) = 0 \cdot a - b + c \end{cases}$$

$a - b = -1$

Incompatible.

$a - b = 0$

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$$

b) $A^T A x = A^T y$

$$A^T A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \qquad A^T y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad A^Ty = \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}$$

$$\Rightarrow x = \begin{pmatrix} 0 \\ 1/2 \\ 5/4 \end{pmatrix}$$

$K(A) = 2 \Rightarrow$ Está bien condicionado.

c) $$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

¿Cuántas reflexiones hay que hacer? 3

$Rg(A) = 3$

$$A^{(1)} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad H_1 A^{(1)} = \pm \|A^{(1)}\|_2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \pm\sqrt{2} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$H_1 = Id - 2\frac{uu^T}{\|u\|_2^2}, \quad u = A^{(1)} \pm \|A^{(1)}\|_2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\Rightarrow H_1 \approx \begin{pmatrix} -0{,}7071 & 0 & -0{,}7071 & 0 \\ 0 & 1 & 0 & 0 \\ -0{,}7071 & 0 & 0{,}7071 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$