Métodos Numéricos-2024

Cuadrados Mínimos Lineales

DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Cuadrados Mínimos Lineales

Dado un conjunto de pares ordenados de valores (x_i, y_i) para $i = 1, \ldots, m$, buscamos una función $f(x)$ perteneciente a una familia $\mathcal{F}$ tal que "mejor aproxime" a los datos.

Cuadrados Mínimos Lineales

Dado un conjunto de pares ordenados de valores (x_i, y_i) para $i = 1, \ldots, m$, buscamos una función $f(x)$ perteneciente a una familia $\mathcal{F}$ tal que "mejor aproxime" a los datos.
¿Cómo podemos expresar matemáticamente "mejor aproxime"?

- $\min\limits_{f \in \mathcal{F}} \max\limits_{i=1,\ldots,m} |f(x_i) - y_i|$

Cuadrados Mínimos Lineales

Dado un conjunto de pares ordenados de valores (x_i, y_i) para $i = 1, \ldots, m$, buscamos una función $f(x)$ perteneciente a una familia $\mathcal{F}$ tal que "mejor aproxime" a los datos.
¿Cómo podemos expresar matemáticamente "mejor aproxime"?

- $\min\limits_{f\in\mathcal{F}} \max\limits_{i=1,\ldots,m} |f(x_i) - y_i|$
- $\min\limits_{f\in\mathcal{F}} \sum\limits_{i=1}^{m} |f(x_i) - y_i|$

Cuadrados Mínimos Lineales

Dado un conjunto de pares ordenados de valores (x_i, y_i) para $i = 1, \ldots, m$, buscamos una función $f(x)$ perteneciente a una familia $\mathcal{F}$ tal que "mejor aproxime" a los datos.
¿Cómo podemos expresar matemáticamente "mejor aproxime"?

- $\min\limits_{f \in \mathcal{F}} \max\limits_{i=1,\ldots,m} |f(x_i) - y_i|$
- $\min\limits_{f \in \mathcal{F}} \sum\limits_{i=1}^{m} |f(x_i) - y_i|$
- $\min\limits_{f \in \mathcal{F}} \sum\limits_{i=1}^{m} (f(x_i) - y_i)^2$

Cuadrados Mínimos Lineales

Dado un conjunto de pares ordenados de valores (x_i, y_i) para $i = 1, \ldots, m$, buscamos una función $f(x)$ perteneciente a una familia $\mathcal{F}$ tal que "mejor aproxime" a los datos.
¿Cómo podemos expresar matemáticamente "mejor aproxime"?

- $\displaystyle\min_{f\in\mathcal{F}} \max_{i=1,\ldots,m} |f(x_i) - y_i|$
- $\displaystyle\min_{f\in\mathcal{F}} \sum_{i=1}^{m} |f(x_i) - y_i|$
- $\displaystyle\min_{f\in\mathcal{F}} \sum_{i=1}^{m} (f(x_i) - y_i)^2$ $\leftarrow$ Método de cuadrados mínimos

Ejemplo 1

$$\mathcal{F} = \{f(x) : f(x) = ax + b\}$$

$$\min_{f \in \mathcal{F}} \sum_{i=1}^{m} (f(x_i) - y_i)^2 = \min_{f(x)=ax+b} \sum_{i=1}^{m} (ax_i + b - y_i)^2$$

Ejemplo 2

$$\mathcal{F} = \{f(x) : f(x) = be^{ax}\}$$

$$\min_{f \in \mathcal{F}} \sum_{i=1}^{m} (f(x_i) - y_i)^2 = \min_{f(x)=be^{ax}} \sum_{i=1}^{m} (be^{ax_i} - y_i)^2$$

Cuadrados Mínimos Lineales

Dada un conjunto de funciones $\{\phi_1, \ldots \phi_n\}$ linealmente independientes, definimos $\mathcal{F} = \{f(x) = \sum_{j=1}^{n} c_j\phi_j\}$.
El problema de Cuadrados mínimos lineales (CML) es:

$$\min_{f\in\mathcal{F}} \sum_{i=1}^{m}(f(x_i) - y_i)^2 = \min_{c_1,\ldots,c_n} \sum_{i=1}^{m}(\sum_{j=1}^{n} c_j\phi_j(x_i) - y_i)^2$$

Cuadrados Mínimos Lineales

Sean $A \in \mathbb{R}^{m\times n}$, $b \in R^m$ y $x \in \mathbb{R}^n$ definidos como

$$\begin{array}{cc} f & \\ & \\ & \\ & \end{array} A = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_n(x_2) \\ \vdots & \vdots & \cdots & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \cdots & \phi_n(x_m) \end{bmatrix} \quad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad \underset{c}{\underline{x}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

El problema de Cuadrados Mínimos Lineales se formula como

$$\min_{x \in \mathbb{R}^n} ||Ax - b||_2^2$$

Cuadrados Mínimos Lineales

Como $Im(A) \oplus Nucleo(A^t) = \mathbb{R}^m$, entonces $b = b^1 + b^2$ con $b^1 \in Im(A)$ y $b^2 \in Nucleo(A^t)$

$$\min_{x \in \mathbb{R}^n} ||Ax - b||_2^2 = \min_{y \in Im(A)} ||y - b||_2^2$$

$$||y-b||_2^2 = ||(y-b^1) - b^2||_2^2 = ||(y-b^1)||_2^2 + ||b^2||_2^2 - 2(y-b^1)^t b^2$$

Pero $(y - b^1) \in Im(A)$ y $b^2 \in Nucleo(A^t) = Im(A)^{\perp}$, entonces el último término se anula.
Buscamos entonces

$$\min_{y \in Im(A)} ||(y - b^1)||_2^2 + ||b^2||_2^2$$

Cuadrados Mínimos Lineales

$$\min_{y \in Im(A)} = ||(y - b^1)||_2^2 + ||b^2||_2^2$$

EL segundo término es constante y lo mínimo que podría valer el primer término es cero, que se alcanzaría cuando $y = b^1$. Esto siempre es posible ya que $b^1 \in Im(A)$. Entonces $x^* \in \mathbb{R}^n$ es solución de cuadrados mínimos lineales si

$$Ax^* = b^1$$

Cuadrados Mínimos Lineales

El problema de Cuadrados Mínimos Lineales siempre tiene solución ya que $b^1 \in Im(A)$.

Cuadrados Mínimos Lineales

El problema de Cuadrados Mínimos Lineales siempre tiene solución ya que $b^1 \in Im(A)$.

La solución es única $\Leftrightarrow$ las columnas de A son linealmente independientes.

Cuadrados Mínimos Lineales

El problema de Cuadrados Mínimos Lineales siempre tiene solución ya que $b^1 \in Im(A)$.

La solución es única $\Leftrightarrow$ las columnas de A son linealmente independientes.

Si $x^* \in R^n$ es solución y $r = b - Ax^*$ entonces

$$A^t r = 0$$

$$A^t A x^* = A^t b$$

(ecuaciones normales)

Sea $A \in \mathbb{R}^{m\times n}$ con $rango(A) = n$. Entonces A^tA es inversible y la solución de MCL es $x^* = (A^tA)^{-1}A^tb$

Cuadrados Mínimos Lineales

Sea $A \in \mathbb{R}^{m \times n}$ con $rango(A) = n$. Entonces $A^t A$ es inversible y la solución de MCL es $x^* = (A^t A)^{-1} A^t b$

Sea $A \in \mathbb{R}^{m \times n}$ con $rango(A) = n$. Sean $b, \bar{b} \in \mathbb{R}^m$ y $b^1, \bar{b}^1$ las proyecciones en $Im(A)$. Si $b^1 \neq 0$ entonces

$$\frac{||(A^t A)^{-1} A^t b - (A^t A)^{-1} A^t \bar{b}||_2}{||(A^t A)^{-1} A^t b||_2} \leq \chi(A) \frac{||b^1 - \bar{b}^1||_2}{||b^1||_2}$$

donde $\chi(A) = ||A||_2 ||(A^t A)^{-1} A^t||_2$

Propiedad: $\chi(A)^2 = \chi(A^t A)$ usando $||.||_2$

Cuadrados Mínimos Lineales

Usando QR-$rango(A) = n$

Sean $A \in \mathbb{R}^{m\times n}$, $rango(A) = n$, $Q \in \mathbb{R}^{m\times m}$ matriz ortogonal y $R_1 \in \mathbb{R}^{n\times n}$ triangular superior con $rango(R_1) = n$ tal que

$$A = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \min_{x\in\mathbb{R}^n} ||Ax - b||_2^2 &= \min_{x\in\mathbb{R}^n} ||Q^t Ax - Q^t b||_2^2 \\ &= \min_{x\in\mathbb{R}^n} ||R_1 x - b^{(n)}||_2^2 + ||b^{(m-n)}||_2^2 \end{aligned}$$

con $Q^t b = (b^{(n)}, b^{(m-n)})$

$$\boxed{\text{Solución } x^* = R_1^{-1} b^{(n)}}$$

Cuadrados Mínimos Lineales

Usando QR-$rango(A) < n$

Sean $A \in \mathbb{R}^{m\times n}$, $rango(A) = r < n$, $Q \in \mathbb{R}^{m\times m}$ matriz ortogonal y $R_1 \in \mathbb{R}^{r\times r}$ triang superior con $rango(R_1) = r$, $R_2 \in \mathbb{R}^{r\times(n-r)}$ tal que

$$\bar{A} = Q \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$$

con $\bar{A} = AP$ con P matriz de permutación. Definiendo $P\bar{x} = x$

$$\min_{x\in\mathbb{R}^n} ||Ax - b||_2^2 = \min_{\bar{x}\in\mathbb{R}^n} ||\bar{A}\bar{x} - b||_2^2 = \min_{\bar{x}\in\mathbb{R}^n} ||Q^t\bar{A}\bar{x} - Q^t b||_2^2$$
$$\min_{\bar{x}\in\mathbb{R}^n} ||\bar{A}\bar{x} - b||_2^2 = \min_{\bar{x}\in\mathbb{R}^n} ||R_1\bar{x}^r + R_2\bar{x}^{n-r} - b^r||_2^2 + ||b^{m-r}||_2^2$$

con $Q^t b = (b^{(r)}, b^{(m-r)})$

$$\boxed{\text{Soluciones: } R_1\bar{x}^{*r} + R_2\bar{x}^{*n-r} = b^r}$$

Usando SVD

Sean $A \in \mathbb{R}^{m \times n}$, $rango(A) = r < n$, $U \in \mathbb{R}^{m \times m}$ matriz ortogonal, $V \in \mathbb{R}^{n \times n}$ matriz ortogonal y $\Sigma \in \mathbb{R}^{m \times n}$ tal que

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

con $\Sigma = \begin{bmatrix} \sigma^1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^r & 0 & \cdots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$ $\sigma^1 \geq \sigma^2 \geq, \ldots, \geq \sigma^r > 0$

Cuadrados Mínimos Lineales

Usando SVD

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

$$\min_{x\in\mathbb{R}^n} ||Ax - b||_2^2 = \min_{x\in\mathbb{R}^n} ||U^t Ax - U^t b||_2^2 = \min_{x\in\mathbb{R}^n} ||\Sigma V^t x - U^t b||_2^2$$

definiendo $y = V^t x$

$$\min_{x\in\mathbb{R}^n} ||Ax - b||_2^2 = \min_{y\in\mathbb{R}^n} ||\Sigma y - U^t b||_2^2$$

$$\min_{x\in\mathbb{R}^n} ||Ax - b||_2^2 == \min_{y\in\mathbb{R}^n} \sum_{i=1}^{r} (\sigma^i y_i - (U^t b)_i)^2 + \sum_{i=r+1}^{m} ((U^t b)_i)^2$$

$$y_i^* = \frac{(U^t b)_i}{\sigma^i} \; \forall i = 1, \ldots, r \quad y_i^* \text{ cualquier valor } \forall i = r+1, \ldots, n$$

$$\boxed{\text{Soluciones: } x^* = Vy^*}$$

Cuadrados mínimos lineales: bibliografía

Recomendamos consultar la numerosa bibliografía existente sobre el tema. Algunas sugerencias:

- Análisis numérico, Richard L. Burden, J. Douglas Faires, International Thomson Editores, 2002.
- Applied Numerical Linear Algebra, James Demmel, SIAM, 1997.
- Numerical Linear Algebra and Optimization, Philip Gill, Walter Murray and Margaret Wright, Addison Wesley, 1991.
- Accuracy and Stability of Numerical Algorithms, Nicholas Higham, SIAM, 2002.
- Analysis of Numerical Methods, Eugene Isaacson and Herbert Keller, Dover Publications, 1994.
- Applied Linear Algebra, Peter J. Olver, Chehrzad Shakiban, Springer International Publishing, 2018.
- Numerical Analysis, Timothy Sauer, Pearson, 2018.
- Numerical Linear Algebra, Lloyd N. Trefethen, SIAM, 1997.
- Fundamentals of Matrix Computations, David Watkins, Wiley, 2010.