Taller 2

Métodos Numéricos

Departamento de Computación, FCEyN, Universidad de Buenos Aires

2 de noviembre de 2018

Hasta ahora...

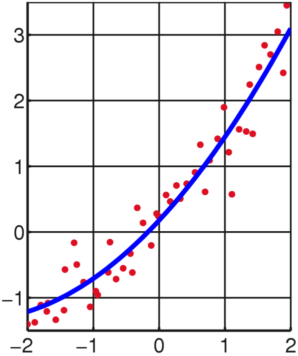
Sabemos resolver sistemas de la forma $Ax = b$ (con $A \in \mathbb{R}^{m\times n}$ cuadrada).

Sirve para muchos problemas y aplicaciones, pero siempre asumimos que A era cuadrada e inversible.

¿Qué pasa si quiero resolver sistemas con $m \neq n$? Ya no se garantiza única solución. ¿En qué casos tiene sentido plantearse esto?

Motivación

Producto de algún experimento obtengo varias medidas que se traducen en un conjunto de puntos (t_i, y_i).



Quiero encontrar una función que relacione "más o menos bien" mis datos. Dicho de otra manera, quiero una f tal que $f(t_i) \approx y_i$.

El problema de *data fitting*

Esto también se conoce como el problema de *data fitting*, donde lo que buscamos es aproximar m muestras a partir de n variables o *features*.

Además, entre más pares $(t_i, y_i) \implies$ más robusta es mi información sobre el experimento.

Nos vamos a restringir al caso donde $m \gg n$, y para resolver el problema vamos a usar el método de cuadrados mínimos lineales (CML).

Cuadrados mínimos **lineales**

Vamos a armar f combinando linealmente una familia de funciones $\mathcal{F} = \{\delta_1, \cdots, \delta_n\}$:
Las funciones δ_j no tienen **ninguna** restricción.

$$f(t_i) = \sum_{j=1}^{n} x_j \cdot \delta_j(t_i) = \begin{bmatrix} \delta_1(t_i) & \cdots & \delta_n(t_i) \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \approx y_i$$

Entonces, si tengo m pares de valores (t_i, y_i)

$$\begin{bmatrix} f(t_1) \\ \vdots \\ f(t_m) \end{bmatrix} = \underbrace{\begin{bmatrix} \delta_1(t_1) & \cdots & \delta_n(t_1) \\ \vdots & \ddots & \vdots \\ \delta_1(t_m) & \cdots & \delta_n(t_m) \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x} \approx \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

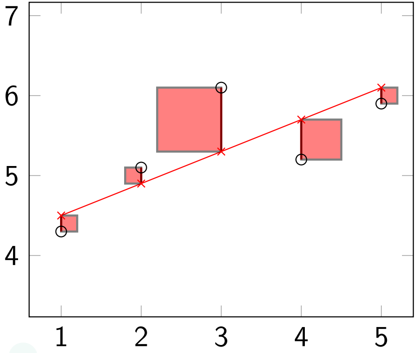
Cuadrados mínimos lineales

¿Cómo sé cuál es la mejor función? Más específicamente, ¿qué solución del sistema anterior estoy buscando?

En el caso de CML buscamos minimizar la suma de los cuadrados de la diferencia entre los datos reales y los que se obtienen con nuestra función.
Es decir que buscamos una función f que garantice minimizar:

$$\arg\min_{f} \sum_{i=1}^{m} (y_i - f(t_i))^2 = \arg\min_{x} \|b - Ax\|_2^2$$

Lo mismo gráficamente



Dados una serie de m puntos (t_i, y_i) Queremos combinar linealmente n funciones para obtener $f(t)$ Y minimizar la suma de los errores... Y minimizar la suma de los errores **cuadrados**

En este caso tomé $\delta_1(t) = 1$ y $\delta_2(t) = t$ y $x = \begin{pmatrix} 4{,}1 \\ 0{,}4 \end{pmatrix}$.

Me queda entonces $f(t) = x_1 \cdot \delta_1(t) + x_2 \cdot \delta_2(t) = 4{,}1 \cdot 1 + 0{,}4 \cdot t$

Recapitulando

$$\mathcal{F} = \{\delta_1, \cdots, \delta_n\}$$

$$f(t_i) = \sum_{i=1}^{n} x_i \cdot \delta_i(t_i)$$

$$\underbrace{\begin{bmatrix} \delta_1(t_1) & \cdots & \delta_n(t_1) \\ \vdots & \ddots & \vdots \\ \delta_1(t_m) & \cdots & \delta_n(t_m) \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x} \approx \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

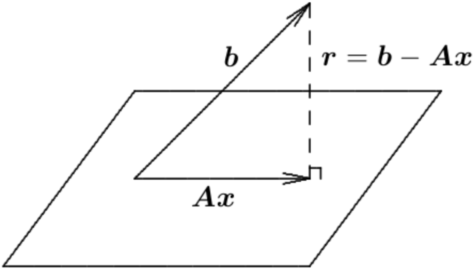
$$\arg\min_{f} \sum_{i=1}^{m} (y_i - f(t_i))^2 = \arg\min_{x} \|b - Ax\|_2^2$$

¿Cómo podemos resolver esto? Seguramente el sistema ni tenga solución exacta porque tenemos más ecuaciones que incógnitas.

Intuición

El problema se traduce en minimizar la norma cuadrada del vector residual $r = b - Ax$

¿Qué vector x hace que Ax sea lo más parecido posible a b?



Lo que buscamos entonces es el x tal que Ax sea la proyección ortogonal de b sobre $Im(A)$

Ecuaciones normales

En base a lo anterior se puede ver que la solución a nuestro problema de cuadrados mínimos está dada por:

$$A^T A x = A^T b \tag{1}$$

Además si A tiene rango columna completo:

- $A^T A$ es no singular. De esta manera resolviendo (1) obtenemos la solución al problema de cuadrados mínimos.
- $A^T A$ es simétrica y definida positiva, por lo que podemos resolver por Cholesky.
- Ventajas de resolver CML usando ecuaciones normales:
    - Si $m \gg n$ (cosa muy común), resolvemos un sistema en $n \times n$
    - Fácil de implementar: calcular $A^T A$ y resolver por Cholesky