Resumen Aproximación

Mínimos cuadrados

Teorema (Principio del mínimo): Si $A \in \mathbb{R}^{N \times N}$ es una matriz simétrica y definida positiva, $b \in \mathbb{R}^N$ y $f: \mathbb{R}^N \longrightarrow \mathbb{R}$ es una función cuadrática definida en cada $x \in \mathbb{R}^N$ como

$$f(x) = \frac{1}{2} x^T A x - b^T x$$

entonces f alcanza un mínimo en un único vector, la solución del sistema $Ax = b$, y su valor mínimo es

$$-\frac{1}{2} b^T A^{-1} b$$

Proposición

- A es simétrica y definida positiva $\Leftrightarrow$ A admite factorización LL Cholesky
- condición suficiente: A regular $\Rightarrow$ A^TA simétrica y definida positiva $(A \in \mathbb{R}^{N \times N})$
- condición suficiente (más general): $rg(A) = N$ $\Leftrightarrow$ columnas de A son linealmente independientes $\Rightarrow$ A^TA es simétrica y definida positiva. $(A \in \mathbb{R}^{M \times N})$ $(y\ N \leq M)$
- condición suficiente (aun más general): $C \in \mathbb{R}^{M \times M}$ simétrica y definida positiva, $A \in \mathbb{R}^{M \times N}$: $rg(A) = N$ $\Leftrightarrow$ columnas de A linealmente independientes $\Rightarrow$ A^TCA simétrica y definida positiva $(y\ N \leq M)$

Aproximación por mínimos cuadrados discreta

Veamos cuál es la mejor aproximación euclídea en dimensión finita.
Sea S un subespacio vectorial de $\mathbb{R}^M$

$$\left\{ \begin{bmatrix} a_{11} \\ \vdots \\ a_{M1} \end{bmatrix}, \dots, \begin{bmatrix} a_{1N} \\ \vdots \\ a_{MN} \end{bmatrix} \right\} \text{ base de } S$$

Luego, rango $(A) = N$, $S = \{ Ax : x \in \mathbb{R}^N \}$
Por lo que si tenemos el vector de soluciones $b \in \mathbb{R}^M$, entonces:

i) $b \in S \Rightarrow Ax = b$ compatible
ii) $b \notin S \Rightarrow Ax = b$ incompatible

Por el principio del mínimo, como A^TA es simétrica y definida positiva,

$$f(x) = \frac{1}{2} x^T A^T A x - b^T A x$$

Luego, es cierto que existen tales vectores y la solución es:

$$A^T A x = A^T b$$

Esta solución recibe el nombre de ecuaciones normales. Además, Ax es la mejor aproximación de b en S en el sentido de los mínimos cuadrados (discretos, $(\mathbb{R}^M, \|\cdot\|)$)

Podemos darle una interpretación geométrica a esta aproximación

Sea S un subespacio vectorial de $\mathbb{R}^M$ $\left\{\begin{bmatrix} a_{11} \\ \vdots \\ a_{M1} \end{bmatrix}, \dots, \begin{bmatrix} a_{1N} \\ \vdots \\ a_{MN} \end{bmatrix}\right\}$ base de $S \Rightarrow S = \{Ax : x \in \mathbb{R}^N\}$

Denotamos por $\langle \cdot, \cdot \rangle$ al producto escalar euclídeo usual en $\mathbb{R}^M$, entonces las ecuaciones normales se pueden expresar usando el producto escalar:

$$A^T A x = A^T b \Leftrightarrow \left\langle \begin{bmatrix} a_{1j} \\ \vdots \\ a_{Mj} \end{bmatrix}, b - Ax \right\rangle = 0, \text{ con } j = 1, \dots, N$$

Geométricamente, el vector $b - Ax$ es perpendicular a los vectores de la base S, equivalentemente a todos los vectores de S ($b - Ax \in S^\perp$), luego Ax es la proyección ortogonal de b sobre S.

$$P_S(b) = Ax$$

Teorema: Sea $A \in \mathbb{R}^{M \times N}$ con $rg(A) = N$, sea S el subespacio vectorial de $\mathbb{R}^M$ definido como

$$S := \{Ax : x \in \mathbb{R}^N\}$$

y sea $b \in \mathbb{R}^M$. Entonces existe un único vector $x \in \mathbb{R}^N$ de forma que

$$\|Ax - b\|_2 = \min \{\|Ay - b\|_2 : y \in \mathbb{R}^N\}$$

De hecho, el vector Ax (mejor aproximación de b en S, proyección ortogonal de b sobre S) viene caracterizado por las ecuaciones normales

$$A^T A x = A^T b$$

equivalentemente

$$\left\langle \begin{bmatrix} a_{1j} \\ \vdots \\ a_{Mj} \end{bmatrix}, b - \sum_{i=1}^{N} x_i \begin{bmatrix} a_{1i} \\ \vdots \\ a_{Mi} \end{bmatrix} \right\rangle = 0, \text{ con } j = 1, \dots, N$$

De este teorema se extrae que dados los datos $(x_0, y_0), (x_1, y_1), \dots, (x_M, y_M) \in \mathbb{R}^2$ tales que $x_i \neq x_j$ $i, j = 0, 1, \dots, M$. Entonces existe una única función polinómica de grado menor o igual que M, $p : \mathbb{R} \to \mathbb{R}$ tal que

$$p(x_i) = y_i, \quad i = 0, 1, \dots, M$$

Aproximación por mínimos cuadrados caso general: caso continuo

Dado un espacio vectorial real E, diremos que una aplicación $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ es un producto escalar en E siempre que se cumpla lo siguiente:

i) $\langle x, y \rangle = \langle y, x \rangle$

ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

iii) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$

iv) $\langle x, x \rangle \geq 0$

$\langle x, x \rangle = 0 \Leftrightarrow x = 0$

donde $x, y, z \in E$, $\lambda \in \mathbb{R}$.

E es un espacio euclídeo con producto escalar $\langle \cdot, \cdot \rangle$. El par $(E, \langle \cdot, \cdot \rangle)$ es un espacio euclídeo.

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio euclídeo, $x \in E$, $N, M \geq 1$, $\lambda_1, \dots, \lambda_N, \mu_1, \dots, \mu_M \in \mathbb{R}$ y $x_1, \dots, x_N, y_1, \dots, y_M \in E$, entonces:

- $\langle x, 0 \rangle = 0$
- $\langle \sum_{i=1}^N \lambda_i x_i, \sum_{j=1}^M \mu_j y_j \rangle = \sum_{i=1}^N \sum_{j=1}^M \lambda_i \mu_j \langle x_i, y_j \rangle$

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio euclídeo, la norma euclídea inducida es:

$$\|x\| := \sqrt{\langle x, x \rangle} \quad x \in E$$

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio euclídeo, $x, y \in E$, la norma inducida por $\langle \cdot, \cdot \rangle$

$$\operatorname{dist}(x, y) = \|x - y\|$$

Teorema (Mejor aproximación): Sean $(E, \langle \cdot, \cdot \rangle)$ un espacio euclídeo, S un subespacio vectorial finito dimensional con base $\{a_1, \dots, a_N\}$ y sea $v \in E$. Entonces existe un único vector $u \in S$ (mejor aproximación de v en S, proyección ortogonal de v sobre S) de forma que

$$\|u - v\| = \min \{ \|w - v\| : w \in S \}$$

De hecho, las coordenadas de $x \in \mathbb{R}^N$ del vector u en la base $a_1, \dots, a_N$ viene caracterizados por las ecuaciones normales $Ax = b$, donde

$$A := |\langle a_i, a_j \rangle|_{i,j=1}^{N} \in \mathbb{R}^{N \times N} \quad y \quad b := |\langle a_i, v \rangle|_{i=1}^{N} \in \mathbb{R}^N$$

equivalentemente,

$$\left\langle a_i, \ v - \sum_{j=1}^{N} x_j a_j \right\rangle = 0 \qquad i = 1, \dots, N$$