1. Método de los mínimos cuadrados

Se nos plantea el problema de resolver un Sistema de Ecuaciones Lineales (S.E.L.)

$$AX = B$$

Donde $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $B \in \mathbb{R}^m$ y $X\mathbb{R}^n$ el vector incógnita.

Notesé que el sistema no es de cramer puesto que la matriz NO es cuadrada.

definición. *Sea $AX = B$ un S.E.L. con $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ decimos entonces que el sistema $AX = B$ está sobredimensionado si $m > n$*

Supongamos que el sistema sobre el que trabajamos está sobredimensionado, entonces es probable que $rang(A) < Rang(A|B)$. En ese caso el sistema sería incompatible, pero ¿será posible encontrar un $X \in \mathbb{R}^n$ tal que $AX \approx B$?

Ver que efectivamente es posible, va a ser el objetivo de esta sección, junto con determinar cual va a ser el mejor de todos esas ”soluciones aproximadas”

definición. *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $B \in \mathbb{R}^m$, definimos el residuo del S.E.L. $AX = B$ como la función*

$$\begin{aligned} R:\ & \mathbb{R}^n \longrightarrow \mathbb{R}^m \\ & X \longrightarrow R(X) = B - AX \end{aligned}$$

Parece evidente que dada esta definición nuestro problema de encontrar tal X va a resumirse en encontrar el X que satisfaga $R(X) \approx 0_{\mathbb{R}^m}$

definición. *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $B \in \mathbb{R}^m$, definimos la función minimos cuadrados como:*

$$\begin{aligned} J:\ & \mathbb{R}^n \longrightarrow \mathbb{R}^+ \\ & X \longrightarrow \|R(X)\|^2 \end{aligned}$$

Observación:

(1). Si el sistema es compatible, entonces si X es solución de $AX = B$ entonces $J(X) = 0$

(2). Si el sistema es incompatible entonces $J(X) > 0$ para todo $X \in \mathbb{R}^n$

Teorema. *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $B \in \mathbb{R}^m$, J la función mínimos cuadrados del S.E.L. sobredimensionado $AX = B$. Entonces existe al menos un $\widehat{X}$ tal que $J(\widehat{X}) = \min_{X\in\mathbb{R}^n} J(X)$.*

Si además $rang(A) = n$ este $\widehat{X}$ es único y se llamará la solución en el sentido de los minimos cuadrados o pseudo-solución.

demostración. *Tomemos $A = (a_{ij})_{ij}$, $B = (b_i)_i$ y $X = (x_1 \dots x_n)$.*

$$J(x_1, x_2, \dots, x_n) = \|(b_i - \sum_{j=1}^{n} a_{ij}x_j)_i\|^2 = \sum_{i=1}^{m}(b_i - \sum_{j=1}^{n} a_{ij}x_j)^2$$

Esta función es regular, luego el mínimo, si existe estará entre los vectores $X = (x_1, x_2 \dots x_n)$ que verifiquen $\nabla J(X) = 0$.

Calculemos $\dfrac{\partial J}{\partial x_t}$

$$\frac{\partial J}{\partial x_t} = -2\sum_{i=1}^{m} a_{it}\Big(b_i - \sum_{j=1}^{n} a_{ij}x_j\Big)$$

tenemos que:

$$\nabla J(X) = 0 \quad \Longleftrightarrow \quad \sum_{t=1}^{m}\sum_{j=1}^{n} a_{tj}a_{ti}x_j = \sum_{t=1}^{m} a_{ti}b_t$$

pero

$$A^tAX = \sum_{t=1}^{m}\sum_{j=1}^{n} a_{tj}a_{ti}x_j = \sum_{t=1}^{m} a_{ti}b_t = A^tB$$

tenemos entonces que calcular los puntos críticos de J *se ha reducido a solucionar un S.E.L. de* n *ecuaciones con* n *incógnitas.* $A^tA \in \mathcal{M}_{m\times n}(\mathbb{R})$ *una matriz simétrica y definida no negativa.*

Veamos que el sistema $A^tAX = A^tB$ *siempre tiene solución.*

Se nos presentan dos casos:

(1). *rang*$(A) = n$ *entonces rang*$(A^tA) = n$ *por lo que el sistema tiene solución única.*

(2). *rang*$(A) < n$ *en ese caso veremos que* $A^tAX = a^tB$ *tiene un espacio de soluciones.*

(1). *Si rang*$(A) = n$ *entonces para ningun* $X \in \mathbb{R}^n - \{0\}$ *se tiene que* $AX = 0$ *luego la matriz* A^tA *es definida positiva, por lo que sabemos que es inversible y por tanto* $X = (A^tA)^{-1}A^tB$ *es la única solución del sistema.*

(2). *Si rang*$(A) < n$, *entonces construyo el subespacio* S*:*

$$S = \{X \in \mathbb{R}^m \| Z = AX \text{ para algun } X \in \mathbb{R}^n\}$$

Dado un subespacio podemos construir su ortogonal:

$$S^\perp = \{Y \in \mathbb{R}^m \mid Y \perp Z \text{ para todo } Z \in S\}$$

Pero la relación $Y\perp Z \Longleftrightarrow Y\perp AX \Longleftrightarrow Y^tAX = (A^tY)^tX \Longleftrightarrow A^tY\perp X \Longleftrightarrow A^tY = 0$

Luego $S^\perp = \{Y \in \mathbb{R}^m \mid A^tY = 0\}$

Tambien sabemos que $S\oplus S^\perp$ *luego para cada* $B \in \mathbb{R}^m$ *existe* $B_A \in S$ *y* $B_0 \in S^\perp$ *tal que* $B = B_A + B_0$, *pero como* $B_A \in S$ *existe* $X \in \mathbb{R}^n$ *tal que* $B_A = AX$ *luego* $B = AX + B_0$ *multiplicando por* A^t *tenemos que:*

$$A^tB = A^tAX + A^tB_0 = A^tAX$$

porque $B_0 \in S^\perp$ *por tanto* $A^TB_0 = 0$.

Veamos ahora que los puntos críticos son mínimos de J

Para ello calculemos las derivadas segundas para aplicar el criterio de la matriz Hessiana.

$HT(X) = (D_{li})_{li} = (2\sum_{t=1}^{m} a_{ki}a_{al})_{li} = 2A^tA$ *que es semidefinida positiva por tanto los puntos son mínimos (por ser la función convexa)*

☺

Ejemplo. *Calcular la recta de regresión de los datos* $(0,2),(1,3),(2,1)y(3,4)$

Lo ideal seria encontrar un α *y un* β *tales que* $y = \alpha x + \beta$ *y que verifique:*

$$\begin{cases} 2 = \beta \\ 3 = \alpha + \beta \\ 1 = 2\alpha + \beta \\ 4 = 3\alpha + \beta \end{cases}$$

Pero es claro que este sistema es incompatible, luego la mejor aproximación será la solución a nuestro sistema de mínimos cuadrados.

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 \\ 3 \\ 1 \\ 4 \end{pmatrix}$$

Como $rang(A) = 2$ *es decir es máximo la solución de mínimos cuadrados es* *<u>única</u>. Calculemos pues* A^tA *y* A^tB

$$A^tA = \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 17 \\ 10 \end{pmatrix}$$

y la mejor solución posible o pseudo-solución es $X = (\alpha, \beta) = (\frac{2}{5}, \frac{19}{10})$