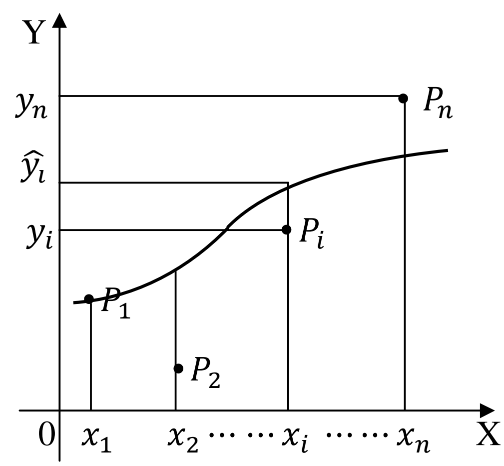
MINIMI QUADRATI

Il Metodo dei "Minimi quadrati" è un procedimento di calcolo approssimato che consente di determinare una funzione matematica rappresentativa di un fenomeno di natura qualsiasi (sociale, economica, fisica, demografica, ecc.).

La funzione prenderà il nome **di funzione interpolante o interpolatrice** se lo scopo è di calcolare tramite essa alcuni dati mancanti nella distribuzione iniziale, mentre si chiamerà **funzione perequatrice** se serve per ricalcolare tutti i valori osservati e, quindi, per sostituire ad essi valori teorici più regolari. La distribuzione iniziale è costituita da un numero finito di coppie di valori (x_i, y_i), osservati per due variabili X e Y (domanda e offerta, altezza e peso, reddito e consumo, ecc.), che rappresentate danno luogo ad una nuvola di punti detta **diagramma a dispersione**.

Sia $y = f(x_i\,; a, b, c, \cdots, k)$ una funzione dall'andamento regolare che approssimi il diagramma a dispersione. Tramite essa si vanno a calcolare i valori y_i imponendo ad ogni punto iniziale P_i l'appartenenza alla curva. Indicheremo con $\hat{y}_i = f(x_i\,; a, b, c, \cdots, k)$ i valori teorici ottenuti. Così facendo, però, si commettono degli errori, per difetto o per eccesso a seconda che il punto si trovi sopra o sotto la curva. Il fine è quello di rendere questi errori quanto più piccoli possibile per ottenere una funzione che più si accosta al diagramma a dispersione.



Per fare ciò, bisogna minimizzare le differenze $d_i = y_i - \hat{y}_i$ tra il valore osservato e quello teorico ma, per evitare che le differenze positive si compensino con quelle negative, **occorre minimizzare i quadrati di queste differenze**. Ciò si attua **determinando dei parametri reali $\boldsymbol{a, b, c, \cdots, k}$ in modo tale che sia minima la funzione:**

$$\boldsymbol{\varphi(a, b, c, \cdots, k) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Condizione necessaria è che le derivate parziali prime siano, contemporaneamente, nulle cioè:

$$\frac{\partial\varphi}{\partial a} = \frac{\partial\varphi}{\partial b} = \frac{\partial\varphi}{\partial c} = \cdots\cdots\cdots = \frac{\partial\varphi}{\partial k} = 0$$

Il che equivale a risolvere il sistema:

$$\begin{cases} 2\sum_{i=1}^{n}[(y_i - f(x_i\,; a, b, c, \cdots, k))] \cdot \frac{\partial f}{\partial a} = 0 \\ 2\sum_{i=1}^{n}[(y_i - f(x_i\,; a, b, c, \cdots, k))] \cdot \frac{\partial f}{\partial b} = 0 \\ \text{---} \\ \text{---} \\ 2\sum_{i=1}^{n}[(y_i - f(x_i\,; a, b, c, \cdots, k))] \cdot \frac{\partial f}{\partial k} = 0 \end{cases}$$

Sia $(a, b, c, \cdots, k)$ una sua soluzione. **Se esiste, essa rende minima la funzione $\boldsymbol{\varphi(a, b, c, \cdots, k)}$ per i seguenti motivi**. **Poiché essa è una somma di quadrati risulta superiormente illimitata e**, pertanto, **non può avere il massimo**. **Del resto**, per questo motivo, **se anche fosse $\boldsymbol{(a, b, c, \cdots, k)}$ una soluzione di massimo, si potrebbero trovare k-ple del tipo $\boldsymbol{(\bar{a}, \bar{b}, \bar{c}, \cdots, \bar{k})}$ tale che sia $\boldsymbol{\varphi(\bar{a}, \bar{b}, \bar{c}, \cdots, \bar{k}) \geq \varphi(a, b, c, \cdots, k)}$. D'altra parte la funzione non può mai essere negativa in quanto somma di quadrati, ma può valere come minimo zero** (ciò accade quando tutti i punti iniziali appartengono alla curva). **Pertanto**, **essendo limitata inferiormente, ha un minimo proprio in $\boldsymbol{(a, b, c, \cdots, k)}$**.

Per giudicare la bontà della scelta si calcolano due indici, detti indici di scostamento I_1 lineare, I_2 quadratico. Se I_1, I_2 sono minori di 0,1 vuol dire che la funzione scelta è la migliore tra quelle dello stesso tipo, ma potrebbe esserci una funzione ancora migliore di diverso tipo (ad esempio una parabola in luogo di una retta).

$$\boldsymbol{I_1 = \frac{\sum_{i=1}^{n}|y_i - \widehat{y_i}|}{\sum_{i=1}^{n} y_i}\,; \quad I_2 = \frac{\sqrt{\frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{n}}}{\frac{\sum_{i=1}^{n} y_i}{n}}}$$

<u>Metodo dei minimi quadrati applicato alla retta</u>: $y = a + bx$

$$\varphi(a,b) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(\mathrm{y_i} - a - bx_i)^2$$

$$\begin{cases} 2\sum_{i=1}^{n}(y_i - a - bx_i)\cdot(-1) = 0 \Rightarrow \\ 2\sum_{i=1}^{n}(y_i - a - bx_i)\cdot(-x_i) = 0 \Rightarrow \end{cases} \quad \begin{cases} na + b\sum x_i = \sum y_i \\ a\sum x_i + b\sum x_i^2 = \sum x_i y_i \end{cases}$$

Dividendo la prima equazione per n e ponendo $\frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$, $\frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$ (medie aritmetiche degli x_i ed y_i) si ha che: $a = \bar{y} - b\bar{x}$, pertanto è sufficiente ricavare b dal sistema. Proviamo che (a, b) rende minima $\varphi(a, b)$. Intanto sappiamo che :

$$\frac{\partial^2\varphi}{\partial a\partial b} = \frac{\partial^2\varphi}{\partial b\partial a} = 2\sum x_i \ ; \ \frac{\partial^2\varphi}{\partial a^2} = 2n \ ; \ \frac{\partial^2\varphi}{\partial b^2} = 2\sum(x_i)^2 \ ;$$

L'Hessiano di ordine 2, costruito con le derivate seconde è:

$$H(a,b) = \begin{vmatrix} 2n & 2\sum x_i \\ 2\sum x_i & 2\sum(x_i)^2 \end{vmatrix} = 4(n\sum(x_i)^2 - (\sum x_i)^2) > 0$$

Per la disuguaglianza di Cauhy-Schwarz[1], eliminando il 4 ed essendo $2n > 0$, ne segue che $(\boldsymbol{a}, \boldsymbol{b})$ è una soluzione di minimo. Se le x_i sono uguali il problema perde di significato.

---

[1] **Disuguaglianza di Cauchy-Schwarz**. Se $(a_1,\ a_2, \cdots a_n)$ e $(b_1,\ b_2, \cdots b_n)$ sono due n-ple di numeri reali di segno qualunque, allora il prodotto delle sommatorie dei quadrati di a_i e b_i è maggiore o uguale del quadrato della sommatoria dei prodotti $a_i b_i$, cioè:

$$\left(\sum_{i=1}^{n} a_i^2\right)\cdot\left(\sum_{i=1}^{n} b_i^2\right) \geq \left(\sum_{i=1}^{n} a_i b_i\right)^2.$$

Vale il segno di uguale se tra a_i e b_i vi è proporzionalità diretta. Caso particolare: se $b_1 = b_2 = \cdots = b_n = 1$ allora la disuguaglianza diventa:

$$n\cdot\left(\sum_{i=1}^{n} a_i^2\right) \geq \left(\sum_{i=1}^{n} a_i\right)^2 \Rightarrow n\cdot\left(\sum_{i=1}^{n} a_i^2\right) - \left(\sum_{i=1}^{n} a_i\right)^2 \geq 0.$$

Il segno di uguale vale se e solo se gli a_i sono tutti uguali.
Esempio: se $n = 3, a_1 = 1, a_2 = 2,\ a_3 = 3$ risulta $3(1 + 4 + 9) > (1 + 2 + 3)^2$ ovvero $42 > 36$.