OPTIMIZATION BASIC

Dario Comanducci, 11 marzo 2024

Esercizio 1

Describe a (simple) example of an optimization problem: what are the variables? how is the objective function formulated? if present, what are the constraints and their formulation?

Per lavoro mi è capitato di dover modellare la relazione $z = f(y|x)$ tra due grandezze y e z, al variare di un parametro nominale x da cui esse dipendono: con riferimento a Fig. 1, il parametro sull'asse x è stato fissato per una serie di valori nominali $x_1 \ldots x_N$, mentre le grandezze misurate sperimentalmente relative a ciascun x_n fissato sono le coppie di punti (y_j, z_j) per $j = 1 \ldots J$.[1]

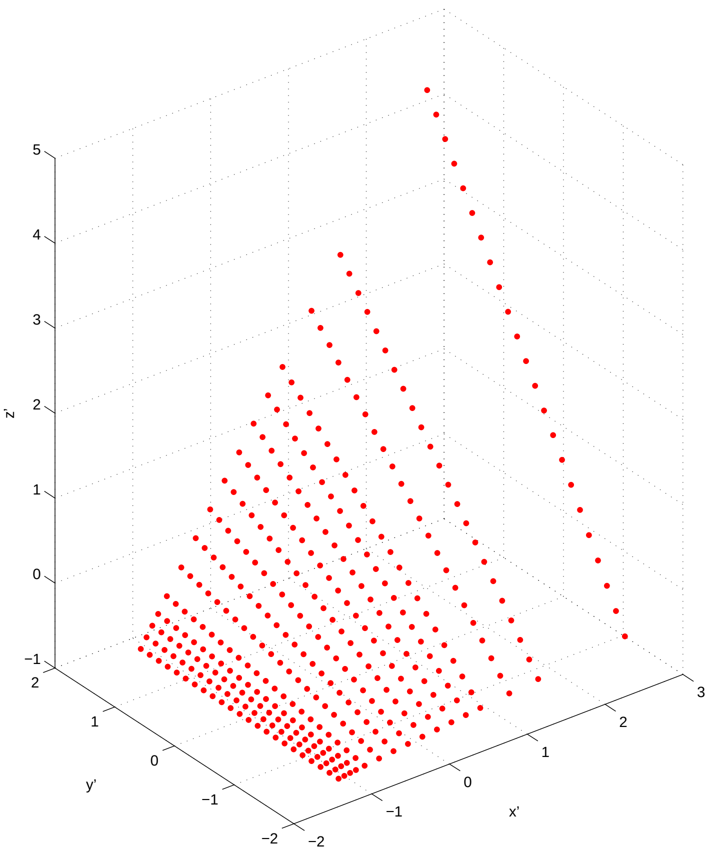

Fig. 1: Dati di input al problema.

[1] Trattandosi di materiale aziendale, preferisco non fornire dettagli sulla natura di x, y e z.

Si può notare da Fig. 1 che i dati raccolti tendono a disporsi in maniera piuttosto regolare nello spazio (x, y, z): seppur il legame $z = f(y|x_n)$ sia grossomodo lineare una volta fissato x_n, per evitare di modellare separatamente N rette (cioè tante quante erano stati i valori nominali scelti a priori), si è preferito interpolare ai minimi quadrati i dati osservati con una superficie, per poi ricavare la dipendenza $z = f(y|x_k)$ intersecando la superficie stimata con il piano $x = x_n$; il modello scelto per la superficie è quello di una quadrica.[2]

Si chiama *quadrica* la superficie luogo di ∞^2 punti le cui coordinate verifichino l'equazione a coefficienti reali e di secondo grado

$$q_{11}x^2+2q_{12}xy+2q_{13}xz+q_{22}y^2+2q_{23}yz+q_{33}z^2+2q_{14}x+2q_{24}y+2q_{34}z+q_{44}=0 \quad (1)$$

Dati K punti $p_k = (x_k,\, y_k,\, z_k)$, $k = 1 \dots K \geqslant 9$, i coefficienti della quadrica sono stimati risolvendo

$$\underbrace{\begin{bmatrix} x_1^2 & 2x_1y_1 & 2x_1z_1 & y_1^2 & 2y_1z_1 & z_1^2 & 2x_1 & 2y_1 & 2z_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_K^2 & 2x_Ky_K & 2x_Kz_K & y_K^2 & 2y_Kz_K & z_K^2 & 2x_K & 2y_K & 2z_K & 1 \end{bmatrix}}_{\mathtt{P}} \mathbf{q} = 0$$

dove $\mathbf{q} = [q_{11}\ q_{12}\ q_{13}\ q_{22}\ q_{23}\ q_{33}\ q_{14}\ q_{24}\ q_{34}\ q_{44}]^\top$. Tuttavia, per $K > 9$ punti p_k soggetti a rumore in genere succederà che

$$\mathtt{P}\mathbf{q} = \varepsilon \neq \mathbf{0}$$

per qualsiasi $\mathbf{q} \neq \mathbf{0}$. Siamo pertanto interessati a determinare $\hat{\mathbf{q}}$ tale da minimizzare la funzione di costo

$$\|\mathtt{P}\mathbf{q}\|^2 = \sum_k \varepsilon_k^2$$

Al fine di evitare la soluzione nulla per $\hat{\mathbf{q}}$ è necessario aggiungere un vincolo: poiché i coefficienti di una quadrica sono definiti a meno di un fattore di scala (l'equazione che modella una quadrica è omogenea, si veda Eq. 1), lo è anche $\mathbf{q}$; per cui possiamo imporre senza perdere di generalità che

$$\|\mathbf{q}\| = 1$$

Riassumendo, dobbiamo trovare

$$\hat{\mathbf{q}} = \arg\min_{\mathbf{q}} \|\mathtt{P}\mathbf{q}\|^2 \ \text{ t.c. } \|\mathbf{q}\| = 1 \qquad (2)$$

La soluzione si può ottenere tramite decomposizione ai valori singolari di $\mathtt{P} = \mathtt{UDV}^\top$, con $\hat{\mathbf{q}}$ data dall'ultima colonna di $\mathtt{V}$ corrispondente al più piccolo valore singolare in $\mathtt{D}$ [1, p. 593] (grazie alle proprietà delle matrici ortogonali $\mathtt{U}$ e $\mathtt{V}$ per le norme, e alla struttura di $\mathtt{D}$). Il risultato stimato sui dati raccolti è mostrato in Fig. 2.

[2] Una quadrica costituisce un buon compromesso tra immunità al rumore, compattezza e ricchezza descrittiva, che può bene modellare la disposizione dei dati osservati; al contrario, il modello di un piano non era in accordo coi valori misurati.

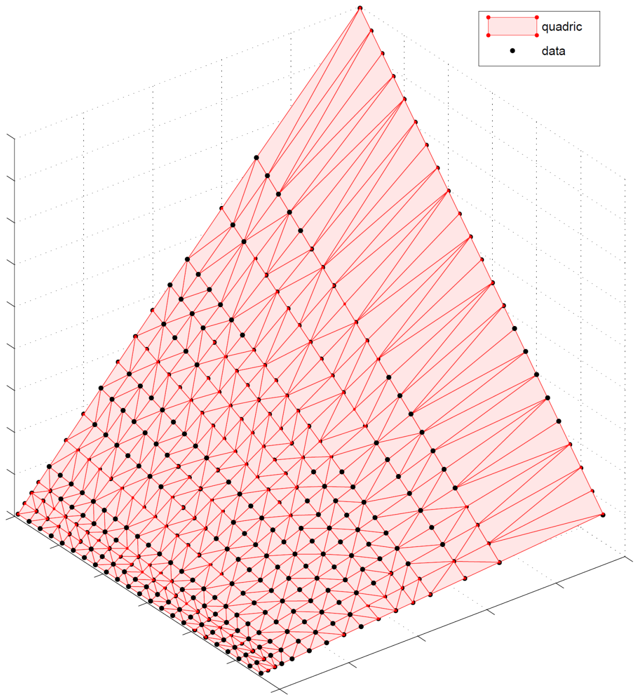

Fig. 2: Quadrica stimata sui punti di Fig. 1.

1 Esercizio 2

Define an optimization problem in one variable with constraints defining a nonempty feasible set S, a local minimum point which is not a global minimum point, multiple global minimum points.

Svolgimento

Svolgendo momentaneamente l'analisi del problema su tutto $\mathbb{R}$, il polinomio mostrato in Fig. 3

$$p(x) = x^8 - 4x^6 - 2x^4 + 12x^2 + 1$$

presenta dei punti stazionari per

$$p'(x) = 8x^7 - 24x^5 - 8x^3 + 24x = 8x(x^6 - 3x^4 - x^2 + 3) = 0$$

ossia per $x = 0$ e per x radice di $x^6 - 3x^4 - x^2 + 3$; posto $z = x^2$, la soluzione per $z^3 - 3z^2 - z + 3 = 0$ è data da:

$$z^3 - 3z^2 - z + 3 = z^2(z-3) - (z-3) = (z-3)(z^2-1) = (z-3)(z-1)(z+1) = 0$$

Tralasciando la soluzione non reale per $x^2 = -1$, in definitiva le radici di $p'(x)$ sono

$$x = 0$$
$$x = \pm 1$$
$$x = \pm\sqrt{3}$$

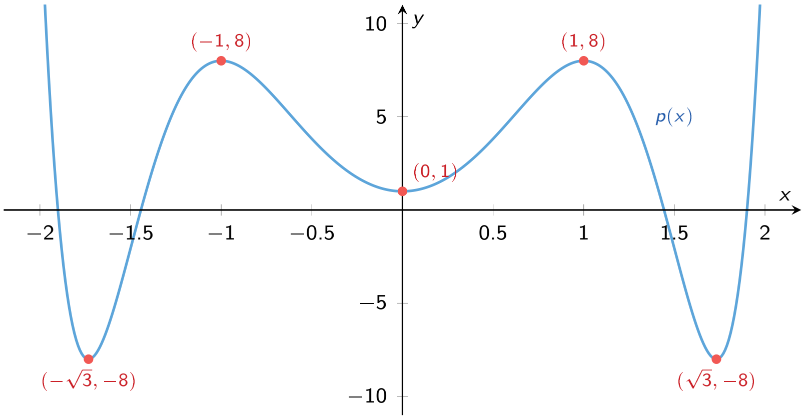

Fig. 3: Polinomio $p(x)$ per l'esercizio 2.

Valutiamo in tali punti la derivata seconda $p''(x) = 8(7x^6 - 15x^4 - 3x^2 + 3)$:

$$p''(0) = 24 > 0 \Rightarrow x = 0 \text{ è un punto di minimo}$$
$$p''(\pm 1) = -64 < 0 \Rightarrow x = \pm 1 \text{ non è un punto di minimo}$$
$$p''(\pm\sqrt{3}) = 384 > 0 \Rightarrow x = \pm\sqrt{3} \text{ è un punto di minimo}$$

Poiché $p(0) = 1 > p(\pm\sqrt{3}) = -8$, abbiamo inoltre che $x = 0$ rappresenta un punto di minimo locale, mentre in $x = \pm\sqrt{3}$ si hanno due minimi globali.[3]

In definitiva, un qualsiasi intervallo $S = (l, u)$, con $l < -\sqrt{3}$ e $u > \sqrt{3}$ va bene come insieme ammissibile al cui interno sono soddisfatte le condizioni richieste dall'esercizio minimizzando la funzione $p(x)$.

Esercizio 3

Prove that any norm is a convex function.

Svolgimento

Dato un insieme convesso S, una funzione $f : S \to \mathbb{R}$ si dice convessa quando

$$f((1-\lambda)x + \lambda y) \leqslant (1-\lambda)f(x) + \lambda f(y) \quad \forall\, x, y \in S, \;\; \forall\, \lambda \in [0,1]$$

Pertanto, valutiamo

$$\|(1-\lambda)x + \lambda y\| \leqslant \|(1-\lambda)x\| + \|\lambda y\| = |1-\lambda|\,\|x\| + |\lambda|\,\|y\|$$

Infine,

$$\lambda \in [0,1] \Rightarrow 1-\lambda \in [0,\, 1] \Rightarrow |1-\lambda|\,\|x\| + |\lambda|\,\|y\| = (1-\lambda)\|x\| + \lambda\|y\|$$

Riferimenti bibliografici

[1] R. I. Hartley, A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004.

[3] Inoltre $\lim_{x \to \pm\infty} p(x) = +\infty$.