PROBLEMA DI OTTIMIZZAZIONE

La forma standard di un problema di ottimizzazione è

$$P = \min\{ f(x) : x \in X \}$$

Dove:

1) f: è la funzione obiettivo
2) X: Sono i vincoli da rispettare. $X \subseteq \mathbb{R}^n$
3) Se $X \equiv \mathbb{R}^n$ allora f è senza vincoli

X

I vincoli sono definiti come funzioni vincolo del tipo:

$$X = \{ x \in \mathbb{R}^n : g(x) < 0,\ h(x) = 0 \}$$
$$g(x) = \{ g_1(x), \dots, g_m(x) \} \quad \text{con } g_i : \mathbb{R}^n \to \mathbb{R}$$
$$h(x) = \{ h_1(x), \dots, h_p(x) \} \quad \text{con } h_i : \mathbb{R}^n \to \mathbb{R}$$

Nota: I problemi di minimizzazione li possiamo vedere tutti come:

$$\max(f(x) : x \in X) = -\min(-f(x) : x \in X)$$

VALORE OTTIMO

Dato il problema P, indichiamo la soluzione ottima $v(P)$:

$$v(P) = \inf(f(x) : x \in X)$$

Ottimo Globale

Il valore ottimo globale di P è $x^* \in X : f(x^*) \leq f(x) \quad \forall x \in X$.

Ottimo Locale

Il valore ottimo locale di P è $x^* \in X : f(x^*) \leq f(x)$ per $\forall x \in B(x^*; r)$ con $r > 0$

TEOREMA DI WEIERSTRASS

Se la funzione obiettivo è continua e se la regione X è limitato, allora esiste (Almeno) un ottimo globale.

Dimostrazione

Sia $V(P) = \inf_{x \in X} (f(x))$ l'ottimo.

Definiamo una sequenza $\{x^k\} \in X$ che tende a $V(P)$.

$$f(x^k) \longrightarrow V(P)$$

Per definizione $\{x^k\}$ è limitata, e il teorema di Bolzano-Weierstrass ci garantisce che esiste una sottosequenza $\{x^{k_i}\} \to x^*$. Ma dato che X è chiuso, necessariamente $x^* \in X$. Dato che $f(x)$ è continua $f(x^{k_i}) \to f(x^*)$.

Da cui:

$$f(x^*) = V(P)$$

COROLLARIO 2

Se la funzione obiettivo è continua e la regione ammissibile X è chiusa e esiste $K \in \mathbb{R}$ tale per cui il set sottolivello :

$$S_K(f) = \{x \in X : f(x) \leq K\}$$

è non-vuoto e limitato, ALLORA esiste (almeno) un ottimo globale

Esempio corollario 2

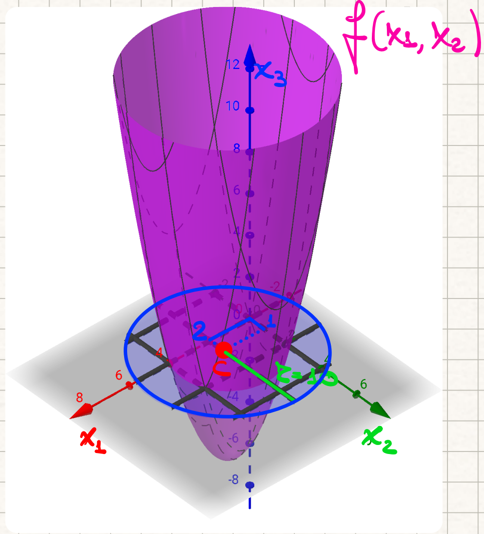
$$f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1 - 2x_2$$

$\exists K \in \mathbb{R}$:

$$x_1^2 + x_2^2 - 4x_1 - 2x_2 \leq K$$

è un cerchio di centro $C = (2, 1)$

e raggio $r = \sqrt{2^2 + 1^2 + K}$



Dimostrazione

Minimizzare f in X equivale a minimizzare f in $S_K(f)$

COROLLARIO 3

Se la funzione obiettivo è continua e coercive, ovvero

$$\lim_{\substack{\|x\| \to \infty \\ x \in X}} f(x) = +\infty$$

e se la regione ammissibile $X \neq \emptyset$ è chiusa, ALLORA esiste (almeno) un ottimo globale.

Dimostrazione

Prendiamo $\bar{x} \in X$, e $K = f(\bar{x})$. Costruiamo ora $S_K(f)$

$$S_K(f) = \{x \in X : f(x) \leq K\}$$

Per la definizione di coercività il sottoinsieme $S_K(f)$ è non vuoto e limitato. Applicando il corollario 2 abbiamo la tesi.

TEOREMA CONVESSITÀ e OTTIMO

Assumendo che f sia convessa in un set X a sua volta convesso. ALLORA ogni ottimo locale del problema P è anche globale.

Dimostrazione

Abbiamo x^* ottimo locale di P:

$$f(x^*) \le f(x) \quad \forall x \in B(x^*, \varepsilon)$$

Per CONTRADDIZIONE diciamo che x^* non sia un ottimo globale e che quindi esiste un altro punto $y \in X$:

$$f(y) < f(x^*).$$

Per la definizione di convessità, il segmento che congiunge x^* e y appartiene a sua volta a X:

$$\forall \alpha \in [0,1] \quad \alpha x^* - (1-\alpha) y \in B(x^*, \varepsilon)$$

Abbiamo quindi che

$$\underline{f(x^*)} \le f(\alpha x^* + (1-\alpha) y) \le \alpha f(x^*) + (1-\alpha) f(y) \le$$

$$\le \alpha f(x^*) + (1-\alpha) f(x^*) = \underline{f(x^*)}$$

Che è impossibile!

PROPOSIZIONE 1

Assumiamo f stricly convex nel set convesso X, e che P ammette un ottimo globale x^* ALLORA x^* è unico.

Dimostrazione

Ancora per contraddizione, assumiamo l'esistenza di un altro punto $y \neq x^*$: $f(y) = f(x^*)$.

Dato che $f(\cdot)$ è strictly convex:

$$f\left(\alpha x^* - (1-\alpha) y\right) < \alpha f(x^*) - (1-\alpha) f(y) = f(x^*) \quad \forall \alpha \in (0,1)$$

Il che contraddice l'ipotesi che x^* sia un ottimo globale.

TEOREMA

Se $f: \mathbb{R}^n \to \mathbb{R}$ è strongly convex in $\mathbb{R}^n$ e sia X chiuso ALLORA esiste un ottimo globale.

Dimostrazione

Sappiamo che ogni funzione convessa in $\mathbb{R}^n$ è anche continua. Inoltre ogni funzione strongly convex è definita come la somma di una funzione convessa più un fattore $\tau \|x\|_2^2$ $\forall \tau > 0$:

$$f(x) = \varphi(x) + \tau \|x\|_2^2$$

Dato che abbiamo $\varphi(\cdot)$ convessa può essere delimitata inferiormente da una funzione lineare:

$$f(x) = \varphi(x) + \underbrace{\tau\|x\|_2^2} \geq a^T x + \underbrace{\tau\|x\|_2^2} \geq -\|a\|_2 \cdot \|x\|_2 + \underbrace{\tau\|x\|_2^2}$$

Da qui possiamo affermare che $f(\cdot)$ è coerciva, e quindi vale il corollario 3 che dimostra la tesi per cui esiste un ottimo globale.

COROLLARIO 1

Se f è strongly convex in $\mathbb{R}^n$ e X è chiuso ALLORA esiste un UNICO ottimo globale

CONDIZIONE DI OTTIMALITÀ PER PROBLEMI NON-VINCOLATI

Consideriamo ora X aperto. Definiamo un problema di ottimizzazione non-vincolato come:

$$\min\left\{f(x) : x \in \mathbb{R}^n\right\}$$

TEOREMA 3 - Condizione necessaria di ottimalità

Assumendo X set aperto, e sia $f(\cdot)$ differenziabile in $x^* \in X$. SE x^* è un ottimo locale per P, ALLORA

$$\nabla f(x^*) = \emptyset$$

Dimostrazione

Per contraddizione, assumiamo che $\nabla f(x^*) \neq \emptyset$.
Scegliamo ora una direzione $d = -\nabla f(x^*)$, da cui definiamo $\varphi(t) = f(x^* + td)$. Facciamo un passo nella direzione del gradiente, supposto non nullo.
Ora, per $t = \emptyset$ abbiamo che

$$\varphi'(t) = d^T \cdot \nabla f(x^* + td)$$

$$\varphi'(\emptyset) = d^T \cdot \nabla f(x^*) = -\|\nabla f(x^*)\|^2 < \emptyset$$

Ma se la derivata per $t = \emptyset$ è negativa, vuol dire che $\varphi(t)$ è decrescente in un intorno di $t = \emptyset$. Da qui la contraddizione, perché se x^* è un minimo locale allora $f(x^*)$ deve essere il valore minimo. Ma $\varphi(t) = f(x^* + td)$ è decrescente per cui viene da sé che

$$f(x^* + td) < f(x^*)$$

Il che è impossibile.

TEOREMA 4 - Second order necessary optimality condition
Sia X aperto, sia $x^* \in X$ ottimo locale per P. ALLORA

1) $\nabla f(x^*) = \emptyset$

2) $\nabla^2 f(x^*) \geq 0 \Rightarrow$ Hessiana semi definita positiva

TEOREMA 5 - Second order sufficient optimality condition
Sia X aperto, sia $x^* \in X$ tale per cui valgono:

1) $\nabla f(x^*) = \emptyset$

2) $\nabla^2 f(x^*) \geq 0 \Rightarrow$ Hessiana definita positiva

ALLORA x^* è un ottimo locale per P

TEOREMA 6 - Optimality condition for convex problem
Sia $f(\cdot)$ una funzione convessa differenziabile nel set X convesso e aperto, ALLORA $x^* \in X$ è un ottimo globale per P SSE $\nabla f(x^*) = \emptyset$

Dimostrazione

La condizione necessaria è verificata grazie al teorema 3.
Assumiamo ora che $\nabla f(x^*) = \emptyset$.

Ricordando che, dato che la $f(\cdot)$ è differenziabile per ipotesi, $f(\cdot)$ è anche convessa SSE

$$f(x) - f(y) \geq (x-y)^T \nabla f(y) \quad \forall x, y \in X$$

Ponendo $y = x^*$ otteniamo che

$$f(x) - f(x^*) \geq 0$$

$$\underbrace{f(x) \geq f(x^*)}_{\text{tesi}} \quad \forall x \in X$$

TEOREMA 7

Sia $f(\cdot)$ differenziabile e strictly convex in X, set aperto e convesso, ALLORA $x^* \in X$ è l'unico ottimo globale per $\mathcal{P}$ SSE $\nabla f(x^*) = 0$

UNCONSTRAINED QUADRATIC PROGRAMMING P

Considerando il problema quadratico:

$$\begin{cases} \min f(x) = \frac{1}{2} x^T Q x + C^T x \\ x \in \mathbb{R}^n \end{cases} \Rightarrow P$$

COROLLARIO

Esiste un ottimo globale x^* per P SSE valgono le seguenti condizioni:

1) $Qx + C = \emptyset$

2) $Q \succeq \emptyset$

Nota: Per la condizione $Q \succeq \emptyset$, un problema non vincolato ammette un ottimo globale SSE $f(\cdot)$ è convessa

Esempio

$$f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + x_3^2 + 3x_1x_2 + x_2x_3 - 6x_1 - 4x_2 - 3x_3$$

$f(\cdot)$ ammette un minimo globale in $\mathbb{R}^3$.

Calcoliamo l'Hessiana Q, calcolando le derivate seconde parziali di $f(\cdot)$:

$$\frac{\partial f(\cdot)}{\partial x_1 \partial x_1} = \left(4x_1 + 3x_2 - 6\right)_{x_1} = 4$$

$$\frac{\partial f(x)}{\partial x_1 \partial x_2} = 3$$

$$\frac{\partial f(x)}{\partial x_1 \partial x_3} = \emptyset$$

Facendo tutte le derivate seconde:

$$Q = \begin{bmatrix} 4 & 3 & \emptyset \\ 3 & 4 & 1 \\ \emptyset & 1 & 2 \end{bmatrix}$$

Calcolando gli autovalori:

$$\lambda_1 = 0.61 \qquad \lambda_2 = 2.28 \qquad \lambda_3 = 7.09$$

Da cui segue che f è strongly convex e che il minimo locale è $x^* = -Q^{-1}c$, con $c = [-6, -4, -3]^T$

Risolvendo troviamo

$$x^* = \begin{bmatrix} 2.7 \\ -1.6 \\ 2.3 \end{bmatrix}$$