

5 Problemi di ottimizzazione convessi

Tra i problemi di Ottimizzazione sono di particolare interesse i cosiddetti problemi *convessi*.

Ricordiamo che un insieme $\mathcal{C} \subseteq \mathbb{R}^n$ è convesso se, presi comunque due punti $y, z \in \mathcal{C}$, risulta che anche $[y, z] \subseteq \mathcal{C}$, avendo denotato con $[y, z]$ il segmento che congiunge y e z , segmento dato dai punti x ottenuti come:

$$x = (1 - \beta)y + \beta z, \quad \beta \in [0, 1].$$

Si verifica facilmente che l'intersezione di un numero finito di insiemi convessi è un insieme convesso.

Definizione 3 (Funzione (strettamente) convessa) Una funzione $v(x)$ si dice convessa su un insieme convesso \mathcal{C} se, presi comunque due punti $y, z \in \mathcal{C}$ risulta che:

$$v((1 - \beta)y + \beta z) \leq (1 - \beta)v(y) + \beta v(z), \quad \beta \in [0, 1]. \quad (7)$$

La funzione $v(x)$ si dice strettamente convessa se, per $y, z \in \mathcal{C}$, $y \neq z$, risulta

$$v((1 - \beta)y + \beta z) < (1 - \beta)v(y) + \beta v(z), \quad \beta \in (0, 1).$$

□

Una funzione $v(x)$ si dice (strettamente) concava su un insieme convesso \mathcal{C} se la funzione $-v(x)$ è (strettamente) convessa su \mathcal{C} .

Nella (7) $y, z, v(y), v(z)$ sono dati, e β varia tra 0 e 1. Se mettiamo in esplicita evidenza la dipendenza da β , introducendo la funzione

$$\phi(\beta) = v((1 - \beta)y + \beta z)$$

otteniamo per una funzione strettamente convessa che:

$$\phi(\beta) < (1 - \beta)\phi(0) + \beta\phi(1) \quad \beta \in (0, 1)$$

Quest'ultima relazione mette in evidenza che, se si rappresenta graficamente nel piano (β, ϕ) la funzione $\phi(\beta)$, il grafico della funzione nell'intervallo $(0, 1)$ si trova al di sotto del segmento, detto *secante*, che congiunge i punti $(0, \phi(0))$ e $(1, \phi(1))$ e coincide solo negli estremi del segmento. Si può concludere che una funzione strettamente convessa è caratterizzata dalla proprietà di avere il grafico sempre al di sotto di ogni sua secante.

A partire dalla definizione di funzione (strettamente) convessa, è possibile dedurre due importanti proprietà che riguardano le derivate, rispettivamente prime e seconde.

Teorema 4 Una funzione $v(x)$ è convessa su un insieme convesso \mathcal{C} se, e solo se, per ogni $y, z \in \mathcal{C}$ risulta:

$$v(z) \geq v(y) + \nabla v(y)^T(z - y);$$

$v(x)$ è strettamente convessa su \mathcal{C} se, e solo se, per ogni $y, z \in \mathcal{C}$, $y \neq z$, risulta:

$$v(z) > v(y) + \nabla v(y)^T(z - y).$$

In termini di funzione $\phi(\beta)$, la precedente disuguaglianza si riscrive come:

$$\phi(1) > \phi(0) + d\phi(0)/d\beta$$

Pertanto, utilizzando la rappresentazione grafica nel piano (β, ϕ) , si può concludere che una funzione strettamente convessa ha il grafico sempre al di sopra di ogni sua tangente.

Teorema 5 Una funzione $v(x)$ è convessa su un insieme convesso \mathcal{C} se, e solo se, per ogni $x \in \mathcal{C}$ risulta:

$$\frac{1}{2}y^T \nabla^2 v(x)y \geq 0, \quad \text{per ogni } y \in \mathbb{R}^n;$$

inoltre, se risulta

$$\frac{1}{2}y^T \nabla^2 v(x)y > 0, \quad \text{per ogni } y \in \mathbb{R}^n,$$

$v(x)$ è strettamente convessa su \mathcal{C} .

Si noti che l'ultima condizione è solo sufficiente: ad esempio, la funzione $v(x) = x^4$ è strettamente convessa in $\mathcal{C} = \mathbb{R}^n$, ma non soddisfa la condizione, in quanto in $x = 0$ la derivata seconda si annulla. Nel caso di funzioni quadratiche le condizioni del teorema diventano invece necessarie e sufficienti, ovvero si ha

Teorema 6 Una funzione quadratica del tipo $q(x) = \frac{1}{2}x^T Qx + c^T x$ è convessa su un insieme convesso \mathcal{C} se e solo se risulta Q semidefinita positiva. Inoltre, $q(x)$ è strettamente convessa su \mathcal{C} se e solo se risulta Q definita positiva.

Definizione 4 (Problema (strettamente) convesso) Si dice che il Problema (1) è un problema di ottimizzazione convesso se l'insieme ammissibile \mathcal{F} è un insieme convesso e la funzione obiettivo $f(x)$ è una funzione convessa su \mathcal{F} . Se la funzione obiettivo è strettamente convessa su \mathcal{F} , il problema si dice strettamente convesso.

I problemi di ottimizzazione convessi sono di particolare importanza per due motivi. Il primo è che la grande maggioranza dei problemi di ottimizzazione che si incontrano nella pratica sono convessi (vedi la Programmazione Lineare). Il secondo è che la convessità induce alcune proprietà che semplificano l'analisi e la soluzione di un problema convesso.

La prima di tali proprietà (altre ne vedremo nel seguito) consiste nel fatto che:

Teorema 7 Un problema di ottimizzazione convesso o non ha soluzione, o ha solo soluzioni globali; non può avere soluzioni esclusivamente locali.

Dimostrazione. La dimostrazione è per assurdo. Ammettiamo che x^* sia una soluzione locale, ma non globale, di un problema convesso $\min_{x \in \mathcal{C}} f(x)$. Allora esisterà un altro punto $z \in \mathcal{C}$ tale che $f(z) < f(x^*)$. La costruzione utilizzata nella dimostrazione è illustrata schematicamente nella figura 2.

Consideriamo il segmento $[x^*, z]$: per la convessità di f , si ha:

$$f((1 - \beta)x^* + \beta z) \leq (1 - \beta)f(x^*) + \beta f(z) = f(x^*) + \beta(f(z) - f(x^*)), \text{ per ogni } \beta \in [0, 1].$$

Il termine $\beta(f(z) - f(x^*))$ risulta < 0 , per ogni $\beta \in (0, 1]$, e si annulla solo per $\beta = 0$. Quindi, poichè in ogni punto $x \in (x^*, z]$ risulta $f(x) < f(x^*)$, non esiste nessun intorno di raggio $\rho > 0$ in cui x^* può soddisfare la definizione di minimo locale. \square

Una seconda proprietà notevole è espressa dalla seguente proposizione:

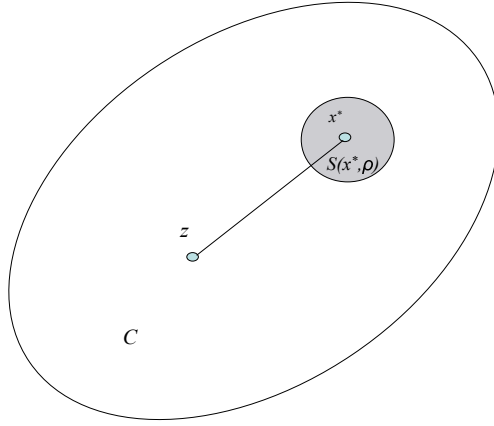


Figure 2: Rappresentazione della costruzione geometrica usata nelle dimostrazione.

Teorema 8 *In un problema di ottimizzazione strettamente convesso la soluzione globale, se esiste, è unica.*

Dimostrazione. Anche questa dimostrazione è per assurdo, e si lascia per esercizio. Sia $f(x)$ strettamente convessa. Si ammetta, per assurdo, che x^* e x^{**} , con $x^* \neq x^{**}$, siano due soluzioni globali di un problema convesso, e se ne traggano le conseguenze. \triangleleft

Riconoscere che un problema di ottimizzazione è convesso fornisce quindi importanti informazioni qualitative sulle sue soluzioni. Per riconoscere che un problema è convesso dobbiamo verificare che \mathcal{F} è un insieme convesso, e che $f(x)$ è convessa su \mathcal{F} , il che non è sempre facile. Ci aiuta la seguente proposizione, di facile utilizzo, e che è verificata molto spesso nella pratica. La proposizione fornisce una condizione sufficiente affinché un problema sia convesso.

Teorema 9 *Si assuma che nel Problema (3) la funzione obiettivo $f(x)$ sia una funzione convessa in \mathbb{R}^n , che i vincoli di disuguaglianza siano dati da funzioni $g_i(x)$ convesse in \mathbb{R}^n , e che i vincoli di uguaglianza siano dati da funzioni affini del tipo $a_j^T x - b_j$. Allora il Problema (3) è convesso.*

Dimostrazione. Facciamo dapprima vedere che, nelle ipotesi poste, l'insieme ammissibile risulta convesso. Osserviamo che l'insieme ammissibile è esprimibile come:

$$\bigcap_{i=1}^p \{x \in \mathbb{R}^n : g_i(x) \leq 0\} \cap \bigcap_{j=1}^m \{x \in \mathbb{R}^n : a_j^T x - b_j = 0\}.$$

Si tratta dell'intersezione di un numero finito di insiemi. Quindi è sufficiente dimostrare che ciascuno di questi insiemi è convesso.

Consideriamo il generico insieme $\{x \in \mathbb{R}^n : g_i(x) \leq 0\}$. Presi due punti y e z che soddisfano il vincolo di disuguaglianza i -esimo, si ha:

$$g_i((1 - \beta)y + \beta z) \leq (1 - \beta)g_i(y) + \beta g_i(z) \leq 0,$$

ove la prima disuguaglianza segue dalla convessità, la seconda segue dal fatto che y e z sono ammissibili. Quindi, poichè in tutti i punti w del segmento $[y, z]$ risulta $g_i(w) \leq 0$, il segmento $[y, z] \subseteq \{x \in \mathbb{R}^n : g_i(x) \leq 0\}$, cioè l'insieme dei punti che soddisfano il vincolo è convesso.

Per il vincolo di uguaglianza $a_j^T x - b_j = 0$, presi due punti y e z che lo soddisfano, si ha

$$a_j^T((1 - \beta)y + \beta z) - b_j = (1 - \beta)(a_j^T y - b_j) + \beta(a_j^T z - b_j) = 0.$$

Quindi, poichè in tutti i punti x del segmento $[y, z]$ risulta $a_j^T x - b_j = 0$, l'insieme dei punti che soddisfano il vincolo è convesso.

Poichè l'intersezione di un numero finito di insiemi convessi è un insieme convesso, possiamo concludere che, nelle ipotesi poste, l'insieme ammissibile del Problema (3) risulta convesso.

Per concludere la dimostrazione, basta osservare che una funzione $f(x)$ convessa in \mathbb{R}^n è convessa anche su ogni sottoinsieme di \mathbb{R}^n . \square

Esempio 5 *Dato il problema*

$$\begin{aligned} \max \quad & x_2 \\ & x_2 - x_1^3 \geq 0 \\ & x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Dire se è convesso. Utilizzare la condizione sufficiente espressa dal Teorema 9 e anche la rappresentazione grafica.

Domande

- 1) Sia $f(x)$ convessa in \mathbb{R}^n . I suoi insiemi di livello sono insiemi convessi?
- 2) Siano $v_1(x), v_2(x)$ funzioni convesse su un insieme convesso \mathcal{C} , e siano $c_1 \geq 0, c_2 \geq 0$ due coefficienti. La funzione $v(x) = c_1 v_1(x) + c_2 v_2(x)$ è convessa su \mathcal{C} ? E se uno dei due coefficienti è negativo?
- 3) Un problema di Programmazione Lineare è convesso?
- 4) Un problema di Programmazione Lineare è strettamente convesso?
- 5) Sia $h_j(x)$ una funzione convessa. Si può affermare che i punti che soddisfano il vincolo $h_j(x) = 0$ è un insieme convesso?
- 6) Si fornisca un esempio di problema di ottimizzazione convesso per cui non vale la Proposizione (9).