Nozioni di base di ottimizzazione

A. Brilli[1]

Questi appunti sono la sintesi del materiale utilizzato per il corso di Programmazione Matematica per la triennale in Ingegneria Informatica[2]

[1] Dip. Ingegneria Informatica Automatica e Gestionale, "Sapienza" Univ. di Roma

[2] Acknowledgement: il materiale è stato sviluppato rivisitando gli appunti delle lezioni del Prof. Liuzzi.

Le basi

Sia $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ consideriamo l'insieme

$$Y = \{t \in \mathbb{R} : t = f(x), \forall\ x \in X\} \subseteq \mathbb{R}$$

Ha senso chiedersi quale sia l'estremo inferiore (superiore) di Y e quindi scrivere

$$f^* = \inf Y \equiv \inf\{t \in \mathbb{R} : t = f(x), \forall\ x \in X\} \equiv \inf\{f(x) : x \in X\} \equiv \inf_{x \in X} f(x)$$

N.B. f^* è sempre (ben) definito perchè $f^* \in \mathbb{R} \cup \{-\infty\}$

Può capitare che $f^* = f(x^*)$ per qualche $x^* \in X$. Quando è così, diciamo che f^* è il valore minimo di f su X e scriviamo

$$f^* = \min_{x \in X} f(x)$$

Le prime definizioni

Definizione (Problema di ottimizzazione)

Data $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$, Il problema

$$\begin{aligned} \min \;\; & f(x) \\ & x \in S \subseteq X \end{aligned}$$

è detto **problema di ottimizzazione**

- **non vincolato** *quando* $S = \mathbb{R}^n$
- **vincolato** *quando* $S \subset \mathbb{R}^n$

Problemi inammissibili e illimitati

Il problema di ottimizzazione

$$\begin{aligned} \min \ & f(x) \\ & x \in S \subseteq X \end{aligned}$$

si dice

- **inammissibile** quando $S = \emptyset$
- **illimitato inferiormente** quando per ogni $M > 0$ esiste $x \in S$ tale che $f(x) < -M$

Punti di minimo

Definizione (Minimo globale)

Un punto $x^* \in S$ *è* **minimo globale** *di* f *su* S *quando*

$$f(x^*) \leq f(x) \qquad \forall\ x \in S$$

minimo globale **stretto** *quando* $f(x^*) < f(x) \qquad \forall\ x \in S,\ x \neq x^*$

Definizione (Minimo locale)

Un punto $x^* \in S$ *è* **minimo locale** *di* f *su* S *quando esiste* $\epsilon > 0$ *tale che*

$$f(x^*) \leq f(x) \qquad \forall\ x \in S \cap B(x^*, \epsilon)$$

minimo locale **stretto** *quando* $f(x^*) < f(x) \qquad \forall\ x \in S \cap B(x^*, \epsilon),\ x \neq x^*$

Cosa può succedere? (Esempi)

$$\min \; f(x) \\ x \in S$$

1. Sia $f(x) = x^2$ e $S = \{x \in \mathbb{R} : \; x > 1, x < -1\}$. (P) è inammissibile
2. Sia $f(x) = -x^2$ e $S = [0, +\infty)$. (P) è illimitato inferiormente
3. Sia $f(x) = x^2$ e $S = [1, +\infty)$. (P) ha minimo e soluzione ottima $f^* = 1 = f(1)$, $x^* = 1 \in S$
4. Sia $f(x) = x^2$ e $S = (1, +\infty)$. (P) non ha soluzione ottima ma $f^* = 1 = \inf\{x^2 : \; x \in S\}$

Riassumendo

Dati $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ e $S \subseteq X$ e considerato (ammettendo che abbia senso)

$$\begin{aligned} \min \ & f(x) \\ & x \in S \subseteq X \end{aligned} \tag{P}$$

abbiamo i seguenti casi

1. (P) è **inammissibile** quando $S = \emptyset$
2. (P) è **illimitato inferiormente** quando per ogni $M > 0$ esiste $x \in S$ tale che $f(x) < -M$
3. (P) **ammette soluzione ottima** quando esiste $x^* \in S$ tale che $f(x^*) \leq f(x)$ per ogni $x \in S$
4. (P) ammette estremo inferiore ma **non soluzione ottima**, i.e. $\inf\{f(x) : \ x \in S\} = f^* > -\infty$ ma non esiste $x^* \in S$ tale che $f(x^*) = f^*$

Programmazione matematica

Definizione (Problema di P.M.)

Data $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$*, Il problema*

$$\begin{aligned} \min \;& f(x) \\ & x \in S \subseteq X \end{aligned}$$

è detto **problema di P.M.** *quando* S *è descritto da un numero finito di equazioni e disequazioni, i.e.*

$$\begin{aligned} S = \{x \in \mathbb{R}^n : & g_i(x) \leq 0, i = 1, \dots, m \\ & h_j(x) = 0, j = 1, \dots, p\} \end{aligned}$$

Introduzione

Un problema di **programmazione matematica**

$$
\begin{aligned}
\min \ & f(x) \\
& g_i(x) \leq 0, \quad i = 1, \ldots, m \\
& h_j(x) = 0, \quad j = 1, \ldots, p
\end{aligned}
$$

tale che:

1. $f(x) = c^\top x$, funzione obiettivo **LINEARE**
2. vincoli di uguaglianza e disuguaglianza **AFFINI**

è detto problema di **programmazione lineare** o problema di PL

$$
\begin{aligned}
\min \ & c^\top x \\
& Ax \leq b
\end{aligned}
$$

Problemi di ottimizzazione

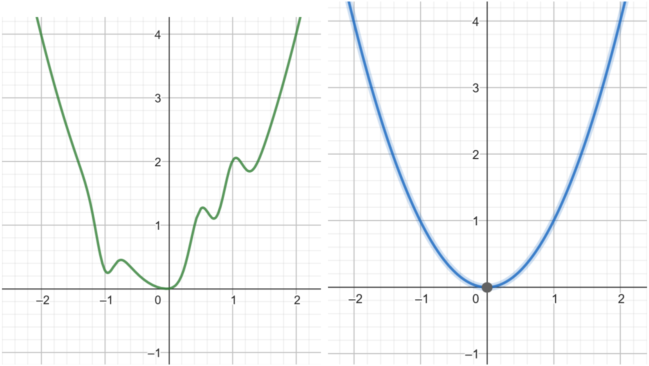
problema di ottimizzazione		problema di programmazione matematica		problema di programmazione lineare
$\min\ f(x)$ $x \in S$ $f : \mathbb{R}^n \to \mathbb{R},\ S \subseteq \mathbb{R}^n$	$\rightsquigarrow$	$\min\ f(x)$ $h(x) = 0$ $g(x) \leq 0$ $h : \mathbb{R}^n \to \mathbb{R}^p,$ $g : \mathbb{R}^n \to \mathbb{R}^m$	$\rightsquigarrow$	$\min\ c^\top x$ $Ax = b$ $x \geq 0$ $c \in \mathbb{R}^n, b \in \mathbb{R}^m,$ $A \in \mathbb{R}^{m \times n}$

Teorema fondamentale della P.L. una ed una sola di queste si verifica

- problema inammissibile
- problema illimitato inferiormente
- esiste soluzione ottima su un vertice

Minimi globali e locali

$\min\{f(x) : \ x \in S\}$ abbiamo visto che oltre al minimo globale possono esistere minimi locali.
Esempio di funzione che presenta minimi locali e funzione che non ne ha



Esistenza del minimo

Dato $X \subseteq \mathbb{R}^n$ e $f : \mathbb{R}^n \to \mathbb{R}$ come possiamo essere sicuri che ha senso scrivere $\min\{f(x), x \in X\}$ cioè che esiste almeno un minimo globale di f su X ?

Distinguiamo due casi:

- $X \neq \mathbb{R}^n$
- $X = \mathbb{R}^n$

Ottimizzazione vincolata ($X \neq \mathbb{R}^n$)

Teorema (Weierstrass)

Sia $X \subset \mathbb{R}^n$ non vuoto e compatto ed f continua su X. Allora f ammette su X minimo e massimo globale

Ottimizzazione non vincolata ($X = \mathbb{R}^n$)

Sia $f: \mathbb{R}^n \to \mathbb{R}$ e $\alpha \in \mathbb{R}$. Definiamo

$$L_\alpha = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$$

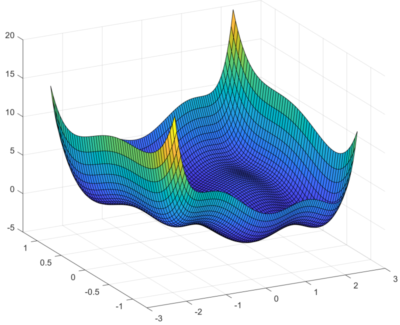
Definizione (Insieme di livello)

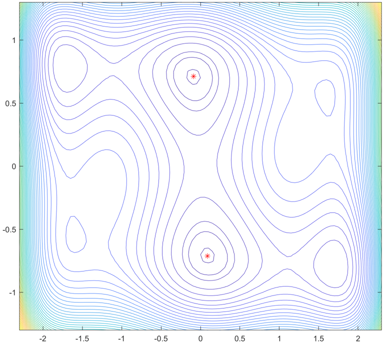
Se L_α per un $\alpha \in \mathbb{R}$ è non vuoto, esso è un insieme di livello inferiore di f

Teorema (Weierstrass)

Sia $f: \mathbb{R}^n \to \mathbb{R}$ continua su $\mathbb{R}^n$. Se esiste un insieme di livello inferiore di f non vuoto e compatto, allora f ammette minimo globale su $\mathbb{R}^n$

Insiemi di livello





Gli strumenti che ci servono

Sia $f\colon \mathbb{R}^n \to \mathbb{R}$ definita in un intorno aperto di x

Definizione (Direzione di discesa)

Una direzione $d \in \mathbb{R}^n, d \neq 0$ è di discesa per f in x quando esiste un $\bar{t} > 0$ tale che

$$f(x + td) < f(x), \quad \forall t \in (0, \bar{t}]$$

1. Se x^* è minimo globale/locale di f su $\mathbb{R}^n$, può esistere una direzione di discesa per f in x^*?
2. Se d è di discesa per f in x, x può essere minimo globale/locale di f?

Direzione di discesa

Sia $f : \mathbb{R}^n \to \mathbb{R}$ e $\bar{x} \in \mathbb{R}^n$. Se esiste una direzione $d \neq 0$ tale che spostandosi da $\bar{x}$ lungo la semiretta $\bar{x} + td$ si ottengono punti strettamente migliori di $\bar{x}$, $\bar{x}$ non è un minimo di f

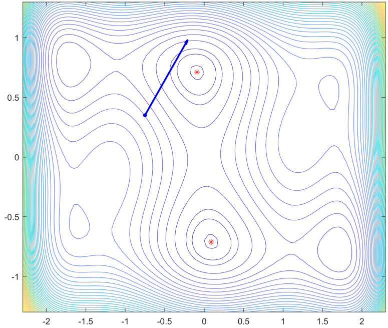
Formalmente

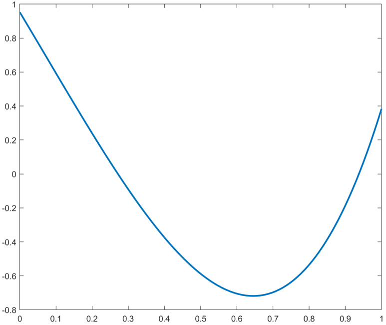
Definizione

Un vettore $d \in \mathbb{R}^n$, $d \neq 0$, è una direzione di discesa per f in $\bar{x}$ se esiste un $\bar{\alpha} > 0$ tale che

$$f(\bar{x} + \alpha d) < f(\bar{x}), \quad \forall\ \alpha \in (0, \bar{\alpha}]$$

Esempio


1
0.5
0
-0.5
-1
-2 -1.5 -1 -0.5 0 0.5 1 1.5 2


1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Cond. di discesa del I ordine

Teorema

Sia $f : \mathbb{R}^n \to \mathbb{R}$ *continuamente differenziabile nell'intorno di un punto* x *e* $d \in \mathbb{R}^n$, $d \neq 0$ *tale che* $\nabla f(x)^\top d < 0$. *Allora* d *è di discesa per* f *in* x.

Basta considerare la definizione di derivata direzionale di f in x lungo d

$$\lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t} = \nabla f(x)^\top d < 0$$

allora per t piccoli si deve avere $f(x + td) < f(x)$

Condizione necessaria del I ordine

Proposizione (Condizione necessaria del I ordine)

Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione differenziabile in un intorno di un punto $x^* \in \mathbb{R}^n$. Se x^* è minimo locale di f, allora

$$\nabla f(x^*) = 0.$$

$$f(x, y) = x^4 + y^4 - 36xy$$

$(0, 0)$ soddisfa la condizione necessaria del I ordine? cioè è **stazionario**?

Geometria in $\mathbb{R}^n$

Definizione (Segmento tra due punti)

Siano $x, y \in \mathbb{R}^n$. Il segmento che congiunge x e y è l'insieme

$$[x, y] = \{z \in \mathbb{R}^n : \; z = ty + (1-t)x, t \in [0, 1]\}$$
$$(x, y) = \{z \in \mathbb{R}^n : \; z = ty + (1-t)x, t \in (0, 1)\}$$

Ma perchè i segmenti sono utili?

Definizione (Insieme convesso)

$C \subseteq \mathbb{R}^n$ è convesso quando, comunque scelti $x, y \in C$ risulta

$$[x, y] \subseteq C$$

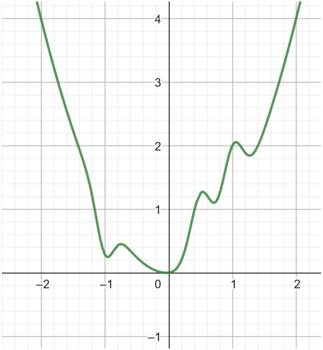
Proprietà

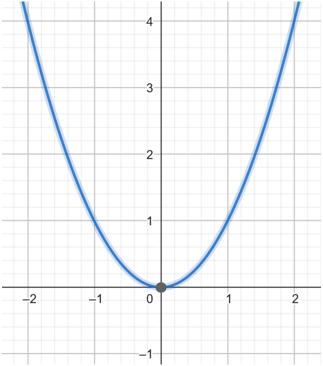
1. $\mathbb{R}^n$ è un insieme convesso (ovvio)
2. $\emptyset$ è un convesso (per mancanza di prove)
3. se $C_1, C_2 \subseteq \mathbb{R}^n$ sono convessi $\Rightarrow$ $C_1 \cap C_2$ è convesso
4. se $C_1, C_2, \ldots, C_p$ sono p insiemi convessi $\Rightarrow$

$$\bigcap_{i=1}^{p} C_i = C \text{ convesso}$$

5. (in generale) l'intersezione di un numero finito o infinito di convessi è un convesso
6. $\{x \in \mathbb{R}^n : \; a^\top x \leq b\}$ (semispazio) è convesso
7. $\{x \in \mathbb{R}^n : \; a^\top x = b\}$ (iperpiano) è convesso (perchè intersezione di due semispazi)
8. un poliedro $P = \{x \in \mathbb{R}^n : \; Ax \leq b\}$ è convesso

Perchè sono così diverse?





- per tanti motivi!
- la curva **BLUE** sta sempre **sotto ad ogni segmento secante**
- la curva **VERDE** no

Funzioni convesse

Intuitivamente: è "convessa" una funzione tale che il suo grafico è sempre al di sotto di ogni segmento secante

Definizione (fun. convessa)

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ *definita su* $S \subseteq \mathbb{R}^n$ *convesso è convessa quando, per ogni* $x, y \in S$

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

Definizione (fun. strettamente convessa)

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ *definita su* $S \subseteq \mathbb{R}^n$ *convesso è strettamente convessa quando, per ogni* $x, y \in S$, $x \neq y$

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y), \quad \forall\, \lambda \in (0,1)$$

Funzioni lineari

Una funzione **lineare** o **affine**

$$f(x) = c^\top x$$

è convessa?

Minimi locali e globali

Teorema

Sia $f : \mathbb{R}^n \to \mathbb{R}$ convessa su $S \subseteq \mathbb{R}^n$ convesso. Allora, ogni minimo locale di f su S è globale.

Dim. x^* minimo locale di f allora esiste $\epsilon > 0$ tale che

$$f(x^*) \leq f(x), \quad \forall\ x \in B(x^*, \epsilon) \cap S$$

Sia $x \in S \setminus B(x^*, \epsilon)$ consideriamo il segmento $[x^*, x]$. Per convessità deve esistere un $\lambda \in (0, 1)$ tale che $(1 - \lambda)x^* + \lambda x \in B(x^*, \epsilon) \cap S$ quindi

$$f(x^*) \leq f((1 - \lambda)x^* + \lambda x) \leq (1 - \lambda)f(x^*) + \lambda f(x)$$

cioè

$$f(x^*) \leq f(x)$$

Insieme dei minimi globali

Teorema

Sia $f : \mathbb{R}^n \to \mathbb{R}$ *convessa su* $S \subseteq \mathbb{R}^n$ *convesso. L'insieme dei punti di minimo globale è convesso.*

Dim. Sia X^* l'insieme dei minimi globale. Il teorema è banalmente vero se $X^* = \emptyset$ oppure $X^* = \{x^*\}$. Supponiamo quindi che f ammetta su f due minimi globali $\bar{x} \in X^*$ e $\bar{y} \in X^*$. Quindi $f(\bar{x}) = f(\bar{y})$.
Consideriamo un punto $z \in [\bar{x}, \bar{y}]$ cioè $z = (1-\lambda)\bar{x} + \lambda\bar{y}$, $\lambda \in [0, 1]$. Per convessità di f abbiamo

$$f((1-\lambda)\bar{x} + \lambda\bar{y}) \leq (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}) = f(\bar{x})$$

Programmazione lineare

Se $c^\top x$ è convessa e $P = \{x : Ax = b, x \geq 0\}$ è una regione convessa, allora

$$\begin{array}{l} \min \ c^\top x \\ \quad Ax = b, \ x \geq 0 \end{array}$$

1. può avere minimi che non sono globali?
2. Come è definito l'insieme dei minimi globali del problema?

Cond. di convessità del I ordine

$f : \mathbb{R}^n \to \mathbb{R}$ di classe 1 su $S \subseteq \mathbb{R}^n$ convesso e aperto.

Teorema (C.N.S. di convessità)

f è convessa su S **se e solo se** *per ogni $y \in S$ risulta*

$$f(y) + \nabla f(y)^\top (x - y) \leq f(x) \quad \forall\, x \in S$$

Teorema (C.N.S. di convessità)

f è strettamente convessa su S **se e solo se** *per ogni $y \in S$ risulta*

$$f(y) + \nabla f(y)^\top (x - y) < f(x) \quad \forall\, x \in S$$

Convessità di funzioni quadratiche

Sia Q simmetrica e

$$f(x) = \frac{1}{2} x^\top Q x + c^\top x$$

allora

1. f è **convessa** su $\mathbb{R}^n$ **se e solo se** Q è **semi-definita positiva**
2. f è **strettamente convessa** su $\mathbb{R}^n$ **se e solo se** Q è **definita positiva**

Condizione di minimo globale

Proposizione (C.N.S. di minimo globale)

Sia $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$ e convessa. $x^* \in \mathbb{R}^n$ è minimo globale di f se e solo se $\nabla f(x^*) = 0$.

Dim.

$\boxed{\Rightarrow}$ La necessità è una conseguenza del fatto che se x^* è minimo globale, è anche minimo locale di f.

$\boxed{\Leftarrow}$ Se f è convessa su $\mathbb{R}^n$, per ogni $x \in \mathbb{R}^n$, risulta

$$f(x) \geq f(x^*) + \nabla f(x^*)^\top (x - x^*)$$

Quindi, il risultato segue dall'ipotesi $\nabla f(x^*) = 0$. □