Ricerca Operativa

Massimo Pappalardo
Dipartimento di Informatica
Largo B. Pontecorvo 3, Pisa
massimo.pappalardo@unipi.it

Laurea in Ingegneria Informatica
Universitá di Pisa
A.A. 2025/'26

Informazioni su PNL

Testi da consultare

- A.Carpignani-M-Pappalardo: "Theory and Methods of Optimization", Springer 2025.

Conoscenze richieste

Derivate direzionali, gradiente, matrice hessiana, jacobiana.

Obiettivo:

Fornire conoscenze per problemi di ottimizzazione non lineare.

Argomenti

Teoria e metodi della Ottimizzazione Non Lineare.

Minimi

- Minimi locali e/o globali
- Minimi globali unici e/o stretti (sono equivalenti)
- Minimi locali stretti (non è detto che siano localmente unici)
- Minimi locali localmente unici (sono anche stretti)
- Estremo inferiore
- Minimi nel caso lineare

Non sempre il minimo esiste.

Teorema

Se f è continua e coerciva cioè

$$\lim_{\substack{x \in \mathbb{R}^n \\ \|x\| \to +\infty}} f(x) = +\infty,$$

allora esiste il minimo globale di f.
Se $-f$ è coerciva, allora esiste il massimo globale.

Condizioni di ottimalità

Ricordiamo che una matrice A di ordine $n \times n$, si dice:

semidefinita positiva se $d^T A\, d \geq 0 \qquad \forall\ d \in \mathbb{R}^n$;

definita positiva se $d^T A\, d > 0 \qquad \forall\ d \in \mathbb{R}^n \setminus \{0\}$;

Ricordiamo che:

- A è semidefinita positiva se e solo se gli autovalori della sua parte simmetrica $\frac{A+A^T}{2}$ sono positivi;
- A è definita positiva se e solo se gli autovalori della sua parte simmetrica sono strettamente positivi.

  Se la matrice A è simmetrica allora coincide con la sua parte simmetrica.

Indichiamo con ∇f e Hf il gradiente e la matrice hessiana di f.

Teorema

(Condizione necessaria) Sia f di classe $\mathcal{C}^2$ su $\mathbb{R}^n$.

$$x^* \text{ è un minimo locale } \Longrightarrow \left\{ \begin{array}{l} \nabla f\,(x^*) = 0 \\ Hf\,(x^*) \text{ è semidefinita positiva} \end{array} \right. .$$

Teorema

(Condizione sufficiente) Sia f di classe $\mathcal{C}^2$ su $\mathbb{R}^n$.

$$\left. \begin{array}{l} \nabla f\,(x^*) = 0 \\ H\,f\,(x^*) \text{ è definita positiva} \end{array} \right\} \Longrightarrow x^* \text{ è un minimo locale.}$$

ESERCIZI

Esercizio 1. Trovare, se esistono, massimi e minimi locali della funzione

$$f(x_1, x_2) = x_2^3 - x_2 - x_1^2\,x_2^2 + x_1^2.$$

Esercizio 2. Calcolare la derivata direzionale nel punto $(0, 1)$ e nella direzione $(2, 1)$ della funzione

$$f(x_1, x_2) = 2\,x_1^2\,x_2 - 2\,x_1\,x_2 + \frac{2}{3}\,x_2^3.$$

Esercizio 3. Calcolare il gradiente e la matrice hessiana della funzione

$$f(x_1, x_2) = \log(1 + x_1 + x_2^2)$$

La funzione è limitata superiormente? Inferiormente?

Esercizio 4. Trovare, se esistono, massimi e minimi locali della funzione

$$f(x_1, x_2) = 2\,x_1^2\,x_2 - 2\,x_1\,x_2 + \frac{2}{3}\,x_2^3.$$

Esercizio 1.

$$f(x_1, x_2) = x_2^3 - x_2 - x_1^2 x_2^2 + x_1^2.$$

Punti staz.	Max glob.	Min glob.	Max loc.	Min loc.	Sella
$\left(0,\ \frac{1}{\sqrt{3}}\right)$	NO	NO	NO	SI	NO
$\left(0,\ -\frac{1}{\sqrt{3}}\right)$	NO	NO	NO	NO	SI
$(1,\ 1)$	NO	NO	NO	NO	SI
$(-1,\ 1)$	NO	NO	NO	NO	SI

Esercizio 4.

$$f(x_1, x_2) = 2\,x_1^2\,x_2 - 2\,x_1\,x_2 + \frac{2}{3}\,x_2^3.$$

Punti stazionari	Massimo		Minimo		Sella
	globale	locale	globale	locale	
$(0,\ 0)$	NO	NO	NO	NO	SI
$(1,\ 0)$	NO	NO	NO	NO	SI
$(1/2,\ 1/2)$	NO	NO	NO	SI	NO
$(1/2,\ -1/2)$	NO	SI	NO	NO	NO

Convessità

$x \in \mathbb{R}^n$ tale che $\nabla f(x) = 0$ non è necessariamente un punto di minimo o massimo locale di f.

Nel caso in cui la f sia convessa o concava su $\mathbb{R}^n$ la situazione cambia.

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ si dice convessa su $\mathbb{R}^n$ se:

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y) \qquad \forall\ x, y \in \mathbb{R}^n, \quad \forall\ t \in [0,1].$$

Una funzione f si dice concava se $-f$ è convessa.

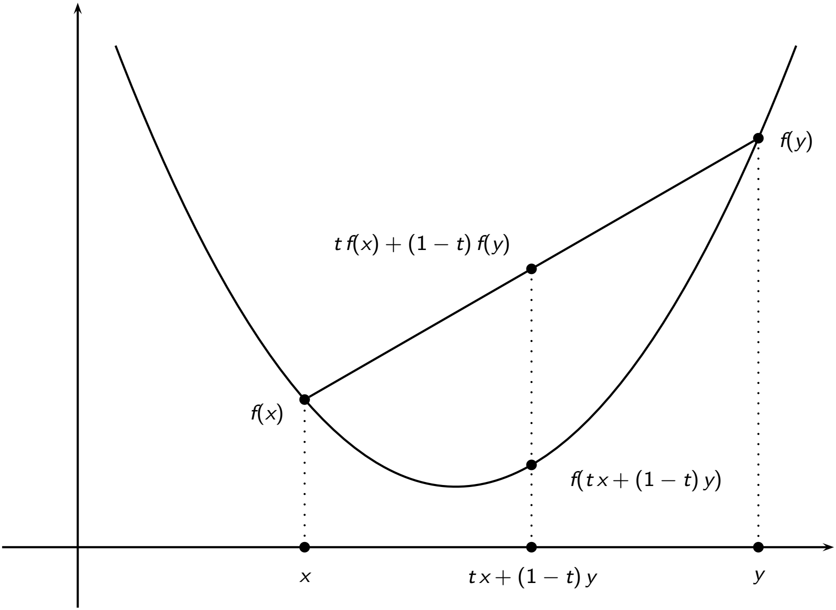


Figure: Funzione convessa.

Teorema
Sia f di classe $\mathcal{C}^2$ su $\mathbb{R}^n$.

$$f \text{ è convessa su } \mathbb{R}^n \quad \Longleftrightarrow \quad Hf(x) \text{ è semidefinita positiva } \forall x \in \mathbb{R}^n.$$

Teorema
Sia f una funzione convessa di classe $\mathcal{C}^1$ su $\mathbb{R}^n$.

$$x^* \text{ è un minimo globale di } f \text{ su } \mathbb{R}^n \quad \Longleftrightarrow \quad \nabla f(x^*) = 0.$$

Convessità di funzioni quadratiche

Sia $f(x) = \frac{1}{2}x^T Qx + c^T x$ una funzione quadratica. f è:

- convessa se e solo se Q è semidefinita positiva

Inoltre

- Se f è convessa e $\alpha > 0$, allora αf è convessa.
- Se f_1 e f_2 sono convesse, allora $f_1 + f_2$ è convessa.
- Se f è convessa, allora $f(Ax + b)$ è convessa.

Esempi

- $$f(x) = -\sum_{i=1}^{m} log(b_i - a_i^T x) \qquad C = \{x \in \mathbb{R}^n : \ b_i - a_i^T x > 0 \quad \forall\, i = 1, \dots, m\}$$
- $f(x) = \|Ax + b\|$

- Se $f_1, \dots, f_m$ sono convesse, allora $f(x) = \max\{f_1(x), \dots, f_m(x)\}$ è convessa.
- Se $\{f_i\}_{i \in I}$ è una famiglia di funzioni convesse, allora $f(x) = \sup_{i \in I} f_i(x)$ è convessa.
- Se $\psi(x, \lambda) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ è convessa in x e concava in λ, allora

$$\begin{aligned} p(x) &= \sup_{\lambda} \psi(x, \lambda) \qquad \text{è convessa} \\ d(\lambda) &= \inf_{x} \psi(x, \lambda) \qquad \text{è concava} \end{aligned}$$

Composizione

$f : \mathbb{R}^n \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$.

Teorema

- Se f è convessa e g è convessa e nondecrescente, allora $g \circ f$ è convessa.
- Se f è concava e g è convessa e noncrescente, allora $g \circ f$ è convessa.
- Se f è concava e g è concava e nondecrescente, allora $g \circ f$ è concava.
- Se f è convessa e g è concava e noncrescente, allora $g \circ f$ è concava.

Per esempio:

- Se g è convessa, allora $e^{g(x)}$ è convessa.
- Se g è concava e positiva, allora $\log g(x)$ è concava.
- Se g è convessa, allora $-\log(-g(x))$ è convessa su $\{x : \; g(x) < 0\}$.
- Se g è concava e positiva, allora $\dfrac{1}{g(x)}$ è convessa.
- Se g è convessa e non negativa, allora $g(x)^p$ è convessa $\forall p \geq 1$.

Algoritmi per f di una variabile

Il metodo di bisezione cerca le soluzioni di $g(x) = 0$ con $g : \mathbb{R} \to \mathbb{R}$.

Per trovare i punti stazionari della funzione f possiamo applicare il metodo di bisezione a $g = f'$.

Se la funzione g è continua su $[a, b]$ e $g(a)\, g(b) < 0$, allora esiste almeno uno zero x^* di f che appartiene all'intervallo (a, b). Per approssimare x^* si divide l'intervallo $[a, b]$ a metà e si stabilisce in quale sottointervallo si trova almeno una soluzione, valutando il segno di $g\left(\frac{a+b}{2}\right)$.

L'ampiezza dell'intervallo contenente la soluzione viene dimezzata ad ogni iterazione.

ALGORITMO DI BISEZIONE

1. Poni $a^0 = a$, $b^0 = b$ e $k = 0$;
2. Poni $x^k = \frac{a^k+b^k}{2}$ e calcola $f'(x^k)$.
3. **Se** un criterio di arresto è soddisfatto
   allora STOP
4. **Se** $f'(a^k)f'(x^k) < 0$
   allora $a^{k+1} = a^k$, $b^{k+1} = x^k$, poni $k = k+1$ e torna al passo 2
   altrimenti $a^{k+1} = x^k$, $b^{k+1} = b^k$, poni $k = k+1$ e torna al passo 2.

Data $\epsilon > 0$, tolleranza fissata, come criterio di arresto usiamo:

$$b^k - a^k = \frac{b-a}{2^k} < \epsilon.$$

La soluzione x^* è approssimata da x^k con un errore minore di ϵ se

$$k > \log_2\left(\frac{b-a}{\epsilon}\right).$$

Esempio

$f(x) = x^4 + 2\,x^2 - 3\,x$. Cerchiamo le soluzioni di:

$$f'(x) = 4\,x^3 + 4\,x - 3 = 0.$$

$f'(0) = -3$ e $f'(1) = 5$, quindi esiste un punto stazionario di f in $[0, 1]$.

La successione di bisezione converge verso il punto $x^* \simeq 0.56736$.

La funzione f è convessa su $\mathbb{R}$ perchè $f''(x) = 12\,x^2 + 4 > 0$. Pertanto, il punto stazionario x^* è anche un minimo globale. $\diamond$

Algoritmo di Newton

Anche l'algoritmo di Newton cerca una soluzione dell'equazione $f(x) = 0$. Partendo da x^k, si fa una linearizzazione della f in un intorno di x^k, cioè

$$f(x) \simeq f(x^k) + f'(x^k)(x - x^k),$$

e si cercano gli zeri di tale linearizzazione:

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}. \tag{1}$$

che è l'intersezione con l'asse delle x della tangente di f in $(x^k, f(x^k))$.

Teorema

Supponiamo che f sia di classe $\mathcal{C}^2$. Sia $f(x^*) = 0$ e $f'(x^*) \neq 0$. Allora esiste un intorno (a, b) di x^* tale che, per ogni x^0 nell'intervallo (a, b), la successione $\{x^k\}$ converge a x^*.

Il metodo di Newton è solo localmente convergente nel senso che la convergenza dipende dalla scelta del punto iniziale.

Esempio

Consideriamo $f(x) = -\frac{1}{4}x^3 + \frac{5}{4}x$. Se scegliamo come punto iniziale $x^0 = 1$, otteniamo $x^1 = -1$, ma di nuovo $x^2 = 1$. La successione oscilla, quindi, tra 1 e -1. $\diamond$

Per trovare i punti stazionari di f, possiamo applicare l'algoritmo di Newton alla derivata f'.

ALGORITMO DI NEWTON IN $\mathbb{R}$

1. Scegli un punto $x^0 \in \mathbb{R}^n$ e poni $k = 0$.
2. Calcola
$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}.$$
3. **Se** un criterio di arresto è soddisfatto
   allora STOP
   altrimenti poni $k = k + 1$ e torna al passo 2.

Gli algoritmi costruiscono successioni ricorsive $\{x^k\}$.

- $\{x^k\}$ converge a x^* che è minimo locale;
- $\{x^k\}$ converge a x^* che è stazionario;
- $\{x^k\}$ ha punti limite che sono stazionari;
- $\{x^k\}$ ha punti limite ed almeno uno di essi è stazionario.

Poiché gli algoritmi non terminano, in generale, in un numero finito di passi, si introducono i criteri di stop; per esempio, fissato $\epsilon > 0$:

- numero dei passi dell'algoritmo;
- $|f(x^{k+1}) - f(x^k)|$ sia $\leq \epsilon$;
- $\|\nabla f(x^k)\|$ sia $\leq \epsilon$.

Algoritmo del gradiente esatto

Per trovare i punti stazionari di f, si costruiscono successioni

$$x^{k+1} = x^k + t^k d^k,$$

dove d^k è una direzione di ricerca discesa mentre t^k è il passo.

Si sceglie come direzione di ricerca quella di massima discesa per la f, ossia quella opposta al gradiente: $d^k = -\nabla f(x^k)$.

Il passo t^k può essere scelto effettuando una "ricerca esatta", ossia, bisogna trovare il minimo di f ristretta alla semiretta uscente da x^k e di direzione $-\nabla f\,(x^k)$:

$$t^k \in \arg\min_{t>0} f\,(x^k - t\,\nabla f\,(x^k)), \tag{2}$$

dove arg min rappresenta l'insieme dei punti di minimo.

ALGORITMO DEL GRADIENTE ESATTO

1. Scegli un punto $x^0 \in \mathbb{R}^n$ e poni $k = 0$.
2. Scegli la direzione $d^k = -\nabla f(x^k)$, calcola

$$t^k \in \arg\min_{t>0} f(x^k + t\, d^k),$$

   e poni $x^{k+1} = x^k + t^k\, d^k$
3. **Se** un criterio di arresto è soddisfatto
   allora STOP
   altrimenti poni $k = k + 1$ e torna al passo 2.

Teorema

1. Due direzioni successive dell'algoritmo sono ortogonali.
2. Se $\nabla f(x^k) \neq 0$ allora $f(x^{k+1}) < f(x^k)$.
3. Se $\lim_{k \to +\infty} x^k = x^*$ allora $\nabla f(x^*) = 0$.

Con qualche ipotesi l'algoritmo del gradiente è globalmente convergente.

Teorema
Se f è coerciva allora, per ogni x_0, $\{x^k\}$ ammette almeno un punto limite ed ogni suo punto limite è un punto stazionario di f.

Vediamo cosa diventa la successione $\{x^k\}$ nel caso in cui f sia quadratica:

$$f(x) = \frac{1}{2} x^T C x + c^T x,$$

con C simmetrica definita positiva. Ponendo $g = C x^k + c$, si ha:

$$f(x^k - t\nabla f(x^k)) = \frac{1}{2} g^T C g\, t^2 - g^T g\, t + f(x^k),$$

$$t^k = \frac{g^T g}{g^T C g},$$

e quindi

$$x^{k+1} = x^k - \frac{g^T g}{g^T C g} g. \tag{3}$$

Il minimo globale di f è $x^* = -C^{-1}c$, allora $x^{k+1} = x^*$ se e solo se

$$Cg = \frac{g^T Cg}{g^T g} g,$$

cioè, se e solo se $\frac{1}{t^k}$ è autovalore di C e g è il suo autovettore. Infatti,

$$-C^{-1}c = x^k - \frac{g^T g}{g^T Cg} g,$$

da cui:

$$-c = Cx^k - \frac{g^T g}{g^T Cg} Cg,$$

ovvero:

$$Cg = \frac{g^T Cg}{g^T g} g.$$

Esempio

Consideriamo $f(x) = x_1^2 + 10\,x_2^2$ che ha minimo globale in $x^* = (0,0)$.

$$Hf(x) = \begin{pmatrix} 2 & 0 \\ 0 & 20 \end{pmatrix} \qquad \forall\; x \in \mathbb{R}^2$$

Applicando l'algoritmo del gradiente esatto a partire da $x^0 = (10,1)$, si può dimostrare che:

$$x^k = \begin{pmatrix} 10\,(\frac{9}{11})^k \\ (-\frac{9}{11})^k \end{pmatrix} \qquad \forall\; k \in \mathbb{N}.$$

Quindi

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = \frac{9}{11} \qquad \forall\; k \in \mathbb{N},$$

La distanza della successione dal minimo globale x^* si riduce ad ogni passo di un fattore $\frac{9}{11}$ segno della lentezza della convergenza.

Algoritmi di discesa: il caso del passo costante

ALGORITMO DEL GRADIENTE con passo costante

1. Scegli un punto $x^0 \in \mathbb{R}^n$ e poni $k = 0$.
2. Scegli la direzione $d^k = -\nabla f(x^k)$ e poni $x^{k+1} = x^k + \eta\, d^k$ dove $0 < \eta < 2/L$, con L costante di Lipschitz di ∇f.
3. **Se** un criterio di arresto è soddisfatto
   allora STOP
   altrimenti poni $k = k + 1$ e torna al passo 2.

Teorema

Supponiamo che f sia coerciva. La successione del metodo del gradiente con passo costante ammette almeno un punto di accumulazione e ogni punto di accumulazione è stazionario.

Algoritmi di discesa con ricerca esatta ed inesatta

Algoritmi di discesa garantiscono $f(x^{k+1}) < f(x^k)$.

Teorema

Se d^k soddisfa

$$\nabla f(x^k)^T d^k < 0. \tag{4}$$

e se t^k è scelto con ricerca esatta, cioè $t^k \in \arg\min_{t>0} f(x^k + t\, d^k)$, allora $f(x^{k+1}) < f(x^k)$.

Sono stati proposti algoritmi chiamati "con ricerca inesatta". Una classe è quella in cui il passo verifica le condizioni di Armijo–Goldstein–Wolfe:

$$\begin{array}{ll} A) & \nabla f(x^{k+1})^T d^k > \beta \nabla f(x^k)^T d^k \\ B) & f(x^{k+1}) < f(x^k) + \alpha\, t^k \nabla f(x^k)^T d^k, \end{array}$$

dove $0 < \alpha < \beta < 1$ sono due parametri prefissati dell'algoritmo.

La condizione A) garantisce che il passo t^k non sia troppo piccolo, in tal caso, infatti, si avrebbe:

$$\nabla f(x^{k+1})^T d^k \simeq \nabla f(x^k)^T d^k,$$

e la condizione A) non potrebbe sussistere.

La condizione B) garantisce che il passo t^k non sia troppo grande; in tal caso, infatti, il membro destro di B) tenderebbe a $-\infty$ e, quindi, B) non potrebbe essere soddisfatta.

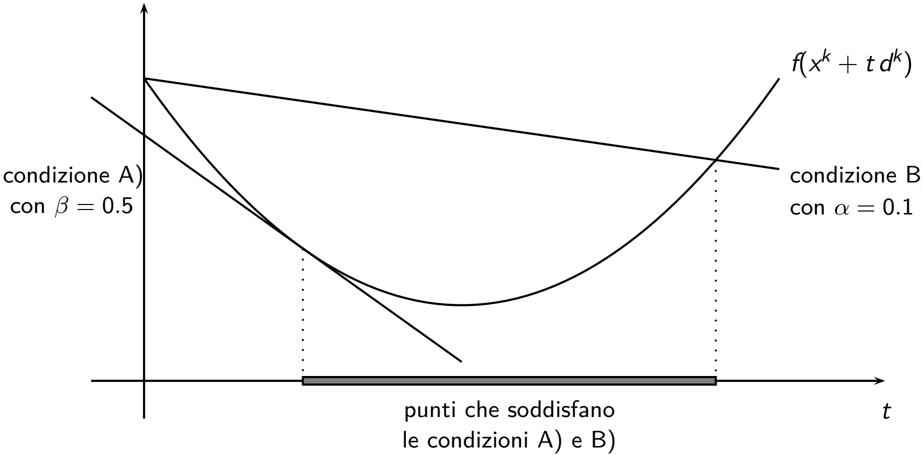


Figure: Condizioni di Armijo–Goldstein–Wolfe.

Esiste un t^k che soddisfi contemporaneamente i criteri A) e B)?

Teorema

Sia f coerciva e sia d^k una direzione di discesa. Allora, esiste un intervallo (a, b), tale che $\forall\ t^k \in (a, b)$, t^k soddisfa le condizioni A) e B).

Per determinare il t^{**}, ovvero il passo, si può usare una tecnica di tipo "backtracking". Allora si sceglie $t^k = 1$ e si procede per ricerche binarie successive, spostandosi a destra o a sinistra, a seconda che la regola violata sia la B) o la A).

METODO "BACKTRACKING" DI RICERCA INESATTA

1. Scegli $\alpha \in (0, \frac{1}{2})$ e $\rho \in (0, 1)$. Poni $t = 1$
2. **Finchè** $f(x^k + t\, d^k) \geq f(x^k) + \alpha\, t\, \nabla f(x^k)^T d^k$,
   poni $t = \rho\, t$.
3. Poni il passo $t^k = t$.

Esempio

Consideriamo $f(x_1, x_2) = x_1^4 + x_1^2 + x_2^2$, il punto $x^k = (1, 1)^T$, la direzione di discesa $d^k = -\nabla f(x^k) = (-6, -2)^T$ ed i parametri $\alpha = 10^{-4}$, $\beta = 0.1$ e $\rho = 0.5$.

Applichiamo il "backtracking". Partiamo con $t = 1$; poichè

$$f(x^k + t\, d^k) = 651 > f(x^k) + \alpha\, t \nabla f(x^k)^T d^k = 2.996,$$

B) non è soddisfatta; riduciamo il passo che diventa $t = 0.5$. Si ha:

$$f(x^k + t\, d^k) = 20 > f(x^k) + \alpha\, t \nabla f(x^k)^T d^k = 2.998,$$

quindi dimezziamo il passo che diventa $t = 0.25$. Infine abbiamo:

$$f(x^k + t\, d^k) = 0.5625 < f(x^k) + \alpha\, t \nabla f(x^k)^T d^k = 2.999$$

perciò il punto $x^{k+1} = x^k + 0.25\, d^k = (-\frac{1}{2}, \frac{1}{2})^T$ soddisfa la condizione B). Osserviamo che esso soddisfa anche la condizione A), infatti:

$$\nabla f(x^{k+1})^T d^k = 7 > \beta \nabla f(x^k)^T d^k = -20.$$

$\diamond$

Vediamo, ora, un teorema di convergenza globale.

Teorema

Supponiamo che f sia $\mathcal{C}^2$ e coerciva. Sia $\{x^k\}$ tale che:

$$x^{k+1} = x^k - t^k \nabla f(x^k),$$

dove t^k verifica le condizioni di Armijo–Goldstein–Wolfe. Allora esiste $\bar{k}$ tale che $\nabla f\,(x^{\bar{k}}) = 0$ oppure $\lim_{k \to +\infty} \|\nabla f\,(x^k)\| = 0$.

Algoritmo di Newton

ALGORITMO DI NEWTON IN $\mathbb{R}^n$

1. Scegli un punto $x^0 \in \mathbb{R}^n$ e poni $k = 0$.
2. Calcola
$$x^{k+1} = x^k - [Hf(x^k)]^{-1} \nabla f(x^k). \tag{5}$$
3. **Se** un criterio di arresto è soddisfatto,
   allora STOP
   altrimenti poni $k = k + 1$ e torna al passo 2.

Teorema

Sia f di classe $\mathcal{C}^3$ e sia x^* punto di minimo locale di f tale che la matrice hessiana $Hf(x^*)$ sia definita positiva. Allora, esiste una costante $\epsilon > 0$ tale che, scelto un punto iniziale x^0, con $\|x^0 - x^*\| < \epsilon$, la successione $\{x^k\}$, definita in (5), converge a x^*.

Osservazione

Nel caso in cui f sia quadratica:

$$f(x) = \frac{1}{2} x^T C x + c^T x,$$

con C simmetrica definita positiva, il metodo di Newton fornisce il minimo globale di f in un solo passo, qualunque sia il punto iniziale x^0. Infatti, se x^* è il minimo di f, cioè $Cx^* + c = 0$, si ha:

$$x^1 = x^0 - C^{-1}(Cx^0 + c) = -C^{-1}c = x^*.$$

Esercizio. Confrontare un'iterazione del metodo del gradiente con una del metodo di Newton a partire dal punto $(-1, 2, -1)$ per:

$$\begin{cases} \max -2\,x_1^4 - 3\,x_3^4 - x_1\,x_2 + 5\,x_1\,x_3 + 4\,x_2 + 6\,x_3 \\ x \in \mathbb{R}^3 \end{cases}$$

Il valore della funzione obiettivo nel punto iniziale è $f(x^0) = 4$.

$$\nabla f\,(x_1, x_2, x_3) = \begin{pmatrix} -8\,x_1^3 - x_2 + 5\,x_3 \\ -x_1 + 4 \\ 5\,x_1 - 12\,x_3^3 + 6 \end{pmatrix}.$$

$$Hf(x_1, x_2, x_3) = \begin{pmatrix} -24\,x_1^2 & -1 & 5 \\ -1 & 0 & 0 \\ 5 & 0 & -36\,x_3^2 \end{pmatrix}.$$

$$d^0 = \nabla f\left(x^0\right) = \begin{pmatrix} 1 \\ 5 \\ 13 \end{pmatrix}.$$

Massimizziamo la f lungo la semiretta di direzione d^0 uscente dal punto x^0. La soluzione ottima è ottenuta per $t \simeq 0.13$, quindi il nuovo punto è $x^1 \simeq \begin{pmatrix} -0.87 \\ 2.65 \\ 0.69 \end{pmatrix}$ e $f(x^1) \simeq 12.22$.

$$Hf(x^0) = \begin{pmatrix} -24 & -1 & 5 \\ -1 & 0 & 0 \\ 5 & 0 & -36 \end{pmatrix}.$$

$$x^1 = x^0 - [Hf\left(x^0\right)]^{-1} \nabla f(x^0).$$

Il nuovo punto è $x^1 \simeq \begin{pmatrix} 4 \\ -111.722 \\ 0.056 \end{pmatrix}$ con $f(x^1) \simeq -510.56$.

Problemi vincolati

Siano $D \subset \mathbb{R}^n$ e $f \colon D \to \mathbb{R}$.

Minimizzare o massimizzare $f(x)$ su D viene indicato così:

$$\begin{cases} \min \; f(x) \\ x \in D \end{cases} \qquad \text{oppure} \qquad \begin{cases} \max \; f(x) \\ x \in D \end{cases}$$

f è detta funzione obiettivo e D regione ammissibile.

Un problema viene detto di PNL se

$$D = \{x \in \mathbb{R}^n : \; g(x) \leq 0, \quad h(x) = 0\},$$

$g(x) = (g_1(x), \ldots, g_m(x))$ e $h(x) = (h_1(x), \ldots, h_p(x))$ sono i vincoli e sono funzioni di classe C^1. Viene anche scritto:

$$\begin{cases} \min \; f(x) \\ g(x) \leq 0 \\ h(x) = 0 \end{cases} \tag{6}$$

Essendo g e h continue D è un dominio chiuso.

Proprietà dei vincoli

Per ottenere condizioni necessarie di ottimalità utili per gli algoritmi bisogna avere conoscenze sui concetti di cono delle direzioni ammissibili, cono tangente, di cono linearizzato (o delle direzioni ammissibili del primo ordine).

Definizione

Dato un punto $x \in D$, un vettore d è detto tangente a D in x se esiste una successione di punti $\{z_k\} \subset D$ con $z_k \to x$ ed una successione di scalari positivi $\{t_k\}$ con $t_k \to 0$ tali che

$$\lim_{k \to \infty} \frac{z_k - x}{t_k} = d.$$

L'insieme di tali d è il cono tangente a D in x, denotato con $T_D(x)$.

PNL

Definizione

Il cono delle direzioni ammissibili a D in x è definito da

$$F_D(x) = \{d \in \mathbb{R}^n : \ \exists \bar{t} \ \ \textit{tale che} \ \ x + td \in D, \ \forall t \in [0, \bar{t}\,]\}.$$

Il cono delle direzioni ammissibili è contenuto nel cono tangente.

Dato un punto $x \in D$, indichiamo con $A(x)$ l'insieme dei vincoli di disuguaglianza che sono attivi in x, cioè $A(x) = \{i : \ g_i(x) = 0\}$.

Definizione

Il cono linearizzato (o delle direzioni ammissibili del primo ordine) a D in x è definito da:

$$L_D^{\leq}(x) = \left\{ d \in \mathbb{R}^n : \quad \begin{array}{ll} d\nabla g_i(x) \leq 0 & \forall \ i \in A(x), \\ d\nabla h_j(x) = 0, & \forall \ j = 1, \ldots, p. \end{array} \right\}$$

Si verifica che $L_D^{\leq}(x)$ è un cono.

Definizione

I vincoli del problema si dicono regolari in $x \in D$ se $T_D(x) = L_D^{\leq}(x)$.

Esempi

Vediamo ora esempi che illustrano questi coni.

Consideriamo la regione $D = \{x \in \mathbb{R}^2 : \; x_1^2 + x_2^2 - 1 = 0\}$, che è la circonferenza di centro $(0, 0)$ e raggio 1.

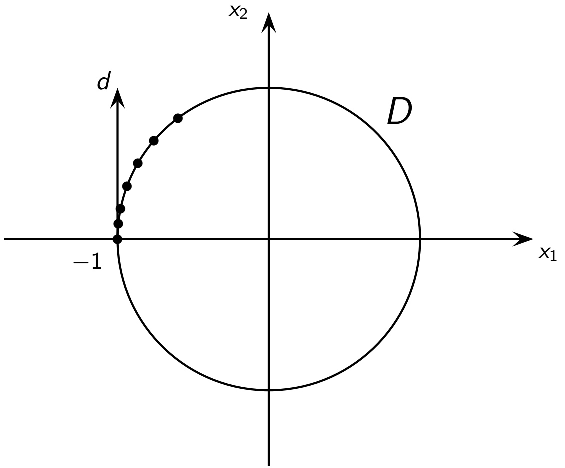
Studiamo la regolarità del vincolo nel punto $x^* = (-1, 0)$.

Calcoliamo prima il cono tangente a D in $(-1, 0)$. Una successione di punti ammissibili che converge a $(-1, 0)$ è:

$$z_k = \left(-\sqrt{1 - \frac{1}{k^2}}, \frac{1}{k} \right).$$

Esempi

Se scegliamo la successione di scalari $t_k = \|z_k - x^*\|$, allora otteniamo il vettore tangente $d = (0, 1)$.



Esempi

In modo analogo, se scegliamo le successioni

$$z_k = \left(-\sqrt{1 - \frac{1}{k^2}}, -\frac{1}{k} \right).$$

e ancora $t_k = \|z_k - x^*\|$, otteniamo il vettore tangente $d = (0, -1)$. Quindi il cono tangente a Din $(-1, 0)$ è

$$T_D(-1, 0) = \{(0, d_2) : \; d_2 \in \mathbb{R}\}.$$

Passiamo al cono delle direzioni ammissibili del primo ordine. Poichè $\nabla h(-1, 0) = (-2, 0)$, dobbiamo trovare i vettori $d \in \mathbb{R}^2$ tali che

$$0 = d\nabla h(x^*) = -2\, d_1,$$

quindi otteniamo che $L_D^{\leq}(-1, 0) = \{d \in \mathbb{R}^2 : \; d_1 = 0\}$ coincide con $T_D(-1, 0)$, quindi il vincolo è regolare nel punto $(-1, 0)$.

Esempi

Consideriamo la regione $D = \{x \in \mathbb{R}^2 : \; x_1^2 + x_2^2 - 1 \leq 0\}$. Anche in questo esempio studiamo la regolarità del vincolo nel punto $x^* = (-1, 0)$.

Calcoliamo prima il cono tangente a D in $(-1, 0)$.

In questo caso, oltre alle due successioni $\{z_k\}$ che abbiamo considerato nell'esempio precedente, dobbiamo tenere conto anche delle successioni di punti appartenenti ad un segmento che convergono verso $(-1, 0)$ dall'interno del cerchio. Queste successioni sono del tipo

$$z_k = (-1, 0) + \frac{1}{k}\, d,$$

dove d è un vettore con la prima componente positiva, cioè $d_1 > 0$. I punti z_k sono ammissibili se $\|z_k\| \leq 1$, cioè se

$$\left(-1 + \frac{d_1}{k}\right)^2 + \left(\frac{d_2}{k}\right)^2 \leq 1,$$

che è vero per $k \geq \frac{\|d\|^2}{2\, d_1}$.

Esempi

Scegliendo $t_k = \frac{\|z_k - x^*\|}{\|d\|}$ otteniamo che d è un vettore tangente.

Inoltre dovremmo considerare anche le successioni che convergono verso $(-1, 0)$ lungo una curva dall'interno del cerchio, ma si troverebbero vettori tangenti già calcolati in precedenza.

Riassumendo, il cono tangente a D in $(-1, 0)$ è

$$T_D(-1, 0) = \{d \in \mathbb{R}^2 : \; d_1 \geq 0\}.$$

Esempi

Passiamo ora al cono linearizzato.

Poichè il vincolo g è attivo in $(-1,0)$ e $\nabla g(-1,0) = (-2,0)$, dobbiamo trovare i vettori $d \in \mathbb{R}^2$ tali che

$$d\nabla g(-1,0) = -2\, d_1 \leq 0,$$

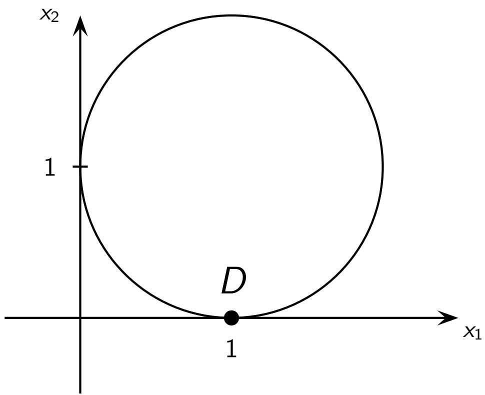
quindi otteniamo che $L_D^{\leq}(-1,0) = \{d \in \mathbb{R}^2 : \ d_1 \geq 0\}$ coincide con $T_D(-1,0)$, quindi il vincolo è regolare nel punto $(-1,0)$.

Consideriamo la regione

$$D = \{x \in \mathbb{R}^2 : \ (x_1-1)^2 + (x_2-1)^2 - 1 \leq 0, \quad x_2 \leq 0\},$$

che coincide con il solo punto $(1,0)$. In questo punto è ovvio che il cono tangente sia $T_D(1,0) = \{(0,0)\}$, poichè l'unica successione ammissibile che converge a $(1,0)$ è la successione costante $z_k = (1,0)$ per ogni $k \in \mathbb{N}$.

Esempi



Inoltre, entrambi i vincoli sono attivi in $(1,0)$, abbiamo che $\nabla g_1(1,0) = (0,-2)$ e $\nabla g_2(1,0) = (0,1)$ e quindi il cono linearizzato è $L_D^{\leq}(1,0) = \{d \in \mathbb{R}^2 : \ d_2 = 0\}$ che non coincide con $T_D(1,0)$. Pertanto i vincoli non sono regolari nel punto $(1,0)$.

Esempi

La definizione di cono tangente si basa sulla geometria della regione D e non sulla sua descrizione algebrica. Al contrario, il cono linearizzato dipende dalla definizione algebrica dei vincoli.

Consideriamo ad esempio la regione D di un esempio precedente che può essere definito anche nel modo seguente:

$$D = \{x \in \mathbb{R}^2 : (x_1^2 + x_2^2 - 1)^2 = 0\}.$$

In questo caso il gradiente del vincolo è

$$\nabla h(x) = \begin{pmatrix} 4\,x_1\,(x_1^2 + x_2^2 - 1) \\ 4\,x_1\,(x_1^2 + x_2^2 - 1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \forall\ x \in D,$$

quindi il cono linearizzato è $L_D^{\leq}(x) = \mathbb{R}^2$. Il vincolo non è regolare nel punto $(-1,0)$ perchè $L_D^{\leq}(-1,0)$ non coincide con $T_D(-1,0) = \{(0,d_2) :\ d_2 \in \mathbb{R}\}$.

Regolarità

Teorema

1. (*Vincoli lineari*) Se g_i e h_j sono funzioni lineari per ogni $i = 1, \dots, m$ e per ogni $j = 1, \dots, p$, allora i vincoli sono regolari in ogni punto $x \in D$.
2. (*Condizione di Slater*) Se le funzioni g_i sono convesse per ogni $i = 1, \dots, m$, le funzioni h_j sono lineari per ogni $j = 1, \dots, p$ ed esiste $\bar{x}$ tale che $g(\bar{x}) < 0$ e $h(\bar{x}) = 0$, allora i vincoli sono regolari in ogni punto punto $x \in D$.
3. (*Condizione di Mangasarian*) Se $\forall\ \bar{x} \in D$ i vettori

$$\begin{cases} \nabla g_i(\bar{x}) & \text{per } i \in A(\bar{x}), \\ \nabla h_j(\bar{x}) & \text{per } j = 1, \dots, p \end{cases}$$

sono linearmente indipendenti, allora i vincoli sono regolari in ogni punto $\bar{x} \in D$.

Un problema è detto convesso se f è convessa e D è convesso.

E' detto concavo se la f è concava e D è convesso.

Se i vincoli g_i sono funzioni convesse, per ogni $i = 1, \ldots, m$ ed i vincoli h_j sono funzioni lineari per ogni $j = 1, \ldots, p$ allora D è convesso.

D'ora in poi supporremo che le funzioni siano di classe $\mathcal{C}^2$.

Analizzeremo condizioni necessarie e/o sufficienti di ottimalità.

Stabilire una condizione necessaria di ottimalità significa trovare una condizione che sia verificata da un punto di minimo.

Stabilire una condizione sufficiente di ottimalità significa invece trovare una condizione che garantisce che il punto trovato sia un minimo. E' possibile che tale condizione non sia soddisfatta da tutti i punti di minimo.

Teorema

Se x^* è un minimo locale di f su D e i vincoli sono regolari in x^*, allora esistono $\lambda^* \in \mathbb{R}^m$, con $\lambda^* \geq 0$, e $\mu^* \in \mathbb{R}^p$ tali che (x^*, λ^*, μ^*) soddisfa il sistema di Lagrange-Karush-Kuhn-Tucker (LKKT):

$$\left\{ \begin{array}{l} \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^p \mu_j^* \nabla h_j(x^*) = 0 \\ \lambda_i^* \, g_i(x^*) = 0 \qquad \forall \ i = 1, \ldots, m \\ h_j(x^*) = 0 \qquad \forall \ i = 1, \ldots, p. \end{array} \right. \tag{7}$$

Se definiamo Lagrangiana, la funzione

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x) + \sum_{j=1}^{p} \mu_j \, h_j(x),$$

allora la prima condizione si può scrivere

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0$$

Se x^* fosse massimo locale il teorema vale ancora cambiando di segno ai moltiplicatori λ. Infatti se fosse massimo locale, x^* sarebbe un minimo locale di $-f$ su D, quindi per il teorema esisterebbero due vettori $\alpha \in \mathbb{R}^m$, con $\alpha \geq 0$, e $\beta \in \mathbb{R}^p$ tali che

$$\begin{cases} -\nabla f(x^*) + \sum_{i=1}^{m} \alpha_i \nabla g_i(x^*) + \sum_{j=1}^{p} \beta_j \nabla h_j(x^*) = 0 \\ \alpha_i^* g_i(x^*) = 0 \qquad \forall\ i = 1, \ldots, m \\ g(x^*) \leq 0 \\ h(x^*) = 0 \end{cases}$$

Definendo $\lambda^* = -\alpha$ e $\mu^* = -\beta$ si ha che (x^*, λ^*, μ^*) risolve il sistema LKKT. c.v.d.

Il Teorema LKKT fornisce solo condizioni necessarie per minimi locali su domini regolari. Consideriamo infatti la funzione $f(x) = x_1 + x_2$ e la regione

$$D = \{x \in \mathbb{R}^2 : \ -x_1^2 - x_2^2 + 2 \leq 0\}.$$

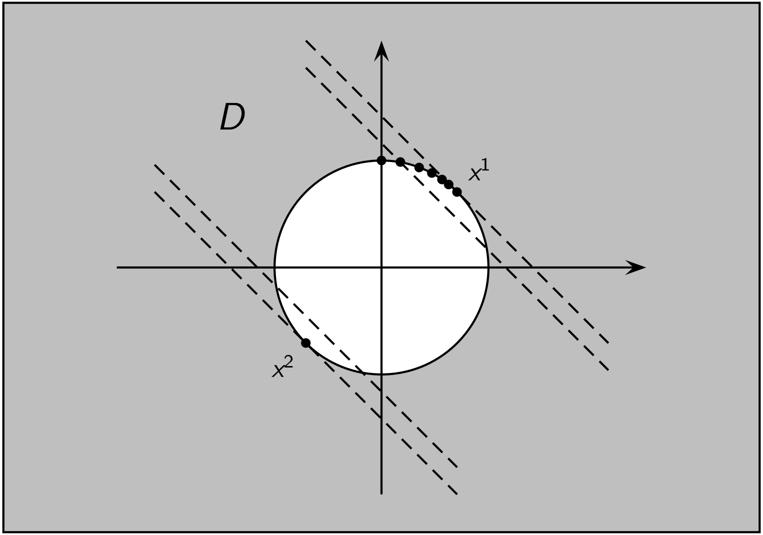
Le soluzioni del sistema LKKT sono:

$$\begin{cases} x^1 = (1,1), \quad \lambda^1 = \frac{1}{2}, \\ x^2 = (-1,-1) \quad \lambda^2 = -\frac{1}{2}. \end{cases}$$

Il punto x^1 soddisfa la condizione necessaria per un minimo locale, ma non è un minimo locale di f su D perchè la successione di punti

$$x^k = \left(1 - \frac{1}{k}, \sqrt{2 - \left(1 - \frac{1}{k}\right)^2}\right),$$

appartiene a D, converge a x^1 e $f(x^k) < f(x^1)$ per ogni $k \in \mathbb{N}$.



In modo analogo x^2 non è un massimo locale di f su D.

Definizione

Se (x^*, λ^*, μ^*) è una soluzione del sistema LKKT, allora x^* è chiamato punto stazionario per domini regolari. Un punto stazionario che non è nè minimo locale nè massimo locale di f su D è detto punto di sella.

Se (x^*, λ^*, μ^*) è una soluzione del sistema LKKT, allora:

- se $\lambda^* \geq 0$, x^* è candidato ad essere un minimo locale di f su D;
- se $\lambda^* \leq 0$, x^* è candidato ad essere un massimo locale di f su D;
- se esistono i e j tali che $\lambda_i > 0$ e $\lambda_j < 0$, allora x^* è una sella.

Cerchiamo massimi e minimi della funzione $f(x) = x_1 + x_2$ sulla regione

$$D = \{x \in \mathbb{R}^2 : \ x_1^2 + x_2^2 - 2 \leq 0\}.$$

Esistono il minimo ed il massimo globale. Il vincolo soddisfa la condizione di Slater perchè la funzione $g(x) = x_1^2 + x_2^2 - 1$ è convessa e $g(0,0) < 0$, quindi è regolare. Risolviamo ora il sistema LKKT:

$$\begin{cases} 1 + 2\,\lambda\,x_1 = 0 \\ 1 + 2\,\lambda\,x_2 = 0 \\ \lambda\,(x_1^2 + x_2^2 - 2) = 0 \\ x_1^2 + x_2^2 - 2 \leq 0 \end{cases}$$

Se $\lambda = 0$, allora la prime due equazioni sono impossibili. Se $\lambda \neq 0$, allora dalle prime due equazioni ricaviamo che $x_1 = x_2 = -\frac{1}{2\,\lambda}$. La terza equazione implica che $x_1^2 + x_2^2 = 2$. Quindi otteniamo $x_1 = \pm 1$. Abbiamo così trovato due soluzioni del sistema LKKT:

$$\begin{array}{ll} x' = (1,1) & \lambda' = -\frac{1}{2} \\ x'' = (-1,-1) & \lambda'' = \frac{1}{2} \end{array}$$

Poichè x' è l'unico punto con $\lambda' \leq 0$, si ha che x' è il massimo globale di f su D. Analogamente x'' è il minimo globale di f su D perchè è l'unico candidato minimo.

Osservazione

Senza l'ipotesi di regolarità dei vincoli, LKKT non è più vero in generale.

Consideriamo $f(x) = x_1 + x_2$ e la regione

$$D = \{x \in \mathbb{R}^2 : \ (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \leq 0, \quad x_2 \leq 0\},$$

che coincide con il solo punto $(1, 0)$, che quindi è sia il massimo globale che il minimo globale di f su D. Tuttavia abbiamo che

$$\nabla f(x^*) = (1, 1), \quad \nabla g_1(x^*) = (0, -2), \quad \nabla g_2(x^*) = (0, 1),$$

quindi non può esistere un vettore λ^* tale che (x^*, λ^*) risolve il sistema LKKT.

Teorema

Supponiamo che f e g_i siano convesse, h_j siano lineari e che (x^*, λ^*, μ^*) sia una soluzione del sistema LKKT con $\lambda^* \geq 0$. Allora x^* è un minimo globale di f su D.

Trovare massimi e minimi globali di $f(x) = x_1 - x_2$ sulla regione

$$D = \{x \in \mathbb{R}^2 : \; -x_1 + x_2^2 - 1 \leq 0 \quad x_1 + x_2^2 - 2 \leq 0\}.$$

La funzione f è convessa e i due vincoli sono convessi. Risolviamo ora il sistema LKKT:

$$\begin{cases} 1 - \lambda_1 + \lambda_2 = 0 \\ -1 + 2\,\lambda_1\,x_2 \; + 2\,\lambda_2\,x_2 = 0 \\ \lambda_1\,(-x_1 + x_2^2 - 1) = 0 \\ \lambda_2\,(x_1 + x_2^2 - 2) = 0 \\ -x_1 + x_2^2 - 1 \leq 0 \\ x_1 + x_2^2 - 2 \leq 0 \end{cases}$$

Distinguiamo 4 casi diversi. Se $\lambda_1 = \lambda_2 = 0$, allora dalla prima equazione si ricava $\lambda_1 = 1$ e dalla seconda $x_2 = \frac{1}{2}$. Inoltre dalla terza equazione si trova che $x_1 = -\frac{3}{4}$. Poichè il punto $(-\frac{3}{4}, \frac{1}{2})$ soddisfa i due vincoli, una soluzione del sistema LKKT è

$$x' = \left(-\frac{3}{4}, \frac{1}{2}\right) \qquad \lambda' = (1, 0).$$

Poichè $\lambda' \geq 0$, il punto x' è un minimo globale di f su D.
Se $\lambda_1 = 0$ e $\lambda_2 \neq 0$ dalla prima equazione si ricava $\lambda_2 = -1$ e dalla seconda $x_2 = -\frac{1}{2}$. Dalla quarta equazione si trova che $x_1 = \frac{7}{4}$.

Poichè $(\frac{7}{4}, -\frac{1}{2})$ soddisfa i vincoli, un'altra soluzione del sistema LKKT è

$$x'' = \left(\frac{7}{4}, -\frac{1}{2}\right) \qquad \lambda'' = (0, -1).$$

Poichè $\lambda'' \leq 0$, il punto x'' è un massimo globale di f su D.

Infine se $\lambda_1 \neq 0$ e $\lambda_2 \neq 0$, allora dalla terza e quarta equazione si ricava che $2\ x_2^2 - 3 = 0$, quindi $x_2 = \pm\frac{\sqrt{6}}{2}$. Dalla prima equazione troviamo $\lambda_1 = 1 + \lambda_2$, mentre dalla terza equazione ricaviamo $x_1 = \frac{1}{2}$. Se $x_2 = \frac{\sqrt{6}}{2}$, allora dalla seconda equazione troviamo $\lambda_2 = \frac{1-\sqrt{6}}{2\ \sqrt{6}} < 0$ e $\lambda_1 = \frac{1+\sqrt{6}}{2\sqrt{6}} > 0$. Quindi il punto $(\frac{1}{2}, \frac{\sqrt{6}}{2})$ è una sella perchè i moltiplicatori λ_1 e λ_2 sono di segno discorde.
Se invece $x_2 = -\frac{\sqrt{6}}{2}$, allora $\lambda_2 = -\frac{1+\sqrt{6}}{2\ \sqrt{6}} < 0$ e $\lambda_1 = \frac{-1+\sqrt{6}}{2\sqrt{6}} > 0$. Quindi in modo analogo a prima otteniamo che $(\frac{1}{2}, -\frac{\sqrt{6}}{2})$ è un punto di sella.

Condizioni del secondo ordine

Abbiamo visto che se x^* è un punto stazionario con $\lambda^* \geq 0$, allora x^* è candidato ad essere un minimo locale di f su D.

Abbiamo condizioni di ottimalità basate sulle derivate seconde di f, g e h.

Definizione

Data una soluzione (x^*, λ^*, μ^*) del sistema LKKT, l'insieme

$$C(x^*, \lambda^*, \mu^*) = \left\{ d \in \mathbb{R}^n : \quad \begin{array}{ll} d\nabla g_i(x^*) = 0 & \forall\ i \in A(x^*) \text{ con } \lambda_i^* \neq 0 \\ d\nabla g_i(x^*) \leq 0 & \forall\ i \in A(x^*) \text{ con } \lambda_i^* = 0 \\ d\nabla h_j(x^*) = 0 & \forall\ j = 1, \ldots, p \end{array} \right\}$$

è chiamato cono critico.

Teorema

Supponiamo che (x^*, λ^*, μ^*) sia una soluzione del sistema LKKT e che valga la condizione di Mangasarian. Se x^* è un minimo locale, allora

$$d^T H_{xx} L(x^*, \lambda^*, \mu^*)\, d \geq 0 \qquad \forall\ d \in C(x^*, \lambda^*, \mu^*).$$

Osservazione
Il Teorema fornisce solo condizioni necessarie per minimi locali, ma non sufficienti.

Consideriamo infatti la funzione $f(x) = x_1^3 + x_2$ e la regione $D = \{x \in \mathbb{R}^2 : \ -x_2 \leq 0\}$. L'unica soluzione del sistema LKKT è il punto $x^* = (0, 0)$, $\lambda^* = 1$. Il vincolo è attivo in x^* e $\nabla g(x^*) = (0, -1)$.

Poichè la matrice $H_{xx}L(x^*, \lambda^*)$ è nulla, il punto x^* soddisfa le condizioni necessarie del secondo ordine sia per un minimo locale che per un massimo locale, tuttavia x^* non è nè un minimo locale, perchè $f(t, 0) < f(0, 0)$ per ogni $t < 0$, nè un massimo locale perchè $\lambda^* > 0$.

Se non sono presenti vincoli, la condizione necessaria del secondo ordine diventa $Hf(x^*)$ semidefinita positiva.

Teorema

Supponiamo che (x^*, λ^*, μ^*) sia una soluzione del sistema LKKT.

1. Se $\lambda^* \geq 0$ e
$$d^T H_{xx} L(x^*, \lambda^*, \mu^*) d > 0 \qquad \forall\ d \in C(x^*, \lambda^*, \mu^*), d \neq 0.$$
allora x^* è un minimo locale.
2. Se $\lambda^* \leq 0$ e
$$d^T H_{xx} L(x^*, \lambda^*, \mu^*) d < 0 \qquad \forall\ d \in C(x^*, \lambda^*, \mu^*), d \neq 0.$$
allora x^* è un massimo locale.

ESERCIZI

Cerchiamo massimi e minimi della funzione $f(x) = x_1 x_2$ sulla regione $D = \{x \in \mathbb{R}^2 : \ x_1^2 + 2x_2^2 - 2 \leq 0\}$. La regione Dè limitata, quindi esistono massimi e minimi globali.

Il vincolo soddisfa la condizione di Slater perchè la funzione g è convessa e $g(0,0) < 0$, quindi il vincolo è regolare in tutti i punti di D. Risolviamo ora il sistema LKKT:

$$\begin{cases} x_2 + 2\lambda x_1 = 0 \\ x_1 + 4\lambda x_2 = 0 \\ \lambda (x_1^2 + 2x_2^2 - 2) = 0 \\ x_1^2 + 2x_2^2 - 2 \leq 0 \end{cases}$$

Se $\lambda = 0$, allora dalle prime due equazioni ricaviamo che $x_1 = x_2 = 0$, quindi una soluzione del sistema è $x^1 = (0,0)$, $\lambda^1 = 0$. Otteniamo le seguenti soluzioni del sistema LKKT:

$$x^2 = \left(1, \frac{\sqrt{2}}{2}\right), \lambda^2 = -\frac{\sqrt{2}}{4} \quad x^3 = \left(-1, -\frac{\sqrt{2}}{2}\right), \lambda^3 = -\frac{\sqrt{2}}{4}$$
$$x^4 = \left(1, -\frac{\sqrt{2}}{2}\right), \lambda^4 = \frac{\sqrt{2}}{4} \quad x^5 = \left(-1, \frac{\sqrt{2}}{2}\right), \lambda^5 = \frac{\sqrt{2}}{4}$$

I punti x^1, x^2 e x^3 sono candidati ad essere massimi globali. Poichè $f(x^1) = 0$ e $f(x^2) = f(x^3) = \frac{\sqrt{2}}{2}$, i punti x^2 e x^3 sono i massimi globali di f su D. Analogamente i punti x^1, x^4 e x^5 sono candidati ad essere minimi globali. Poichè $f(x^1) = 0$ e $f(x^4) = f(x^5) = -\frac{\sqrt{2}}{2}$, i minimi globali di f su Dsono x^4 e x^5.

Rimane da esaminare il punto x^1 che non è nè un minimo globale nè un massimo globale. Non possiamo usare le condizioni sufficienti del primo ordine perchè f non è nè convessa nè concava, essendo indefinita la matrice $Hf(x^1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

La matrice hessiana della Lagrangiana è

$$H^2_{xx}L(x, \lambda) = \begin{pmatrix} 2\lambda & 1 \\ 1 & 4\lambda \end{pmatrix}.$$

Poichè il vincolo non è attivo in x^1, il cono critico $C(x^1, \lambda^1) = \mathbb{R}^2$. La matrice $H^2_{xx}L(x^1, \lambda^1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ è indefinita quindi non sono soddisfatte le condizioni necessarie del secondo ordine nè per un massimo locale, nè per un minimo locale. Pertanto x^1 è una sella.

Cerchiamo massimi e minimi di $f(x) = (x_1 + 1)^2 + (x_2 + 1)^2$ su

$$D = \{x \in \mathbb{R}^2 : \ x_1 + x_2 - 2 \leq 0, \quad -x_1 + x_2 - 2 \leq 0\}.$$

che non è limitata. La funzione f è coerciva, quindi esiste un minimo globale di f su D, ma non esistono massimi globali. I vincoli sono lineari e quindi sono regolari. Risolviamo il sistema LKKT:

$$\begin{cases} 2\,(x_1 + 1) + \lambda_1 - \lambda_2 = 0 \\ 2\,(x_2 + 1) + \lambda_1 + \lambda_2 = 0 \\ \lambda_1\,(x_1 + x_2 - 2) = 0 \\ \lambda_2\,(-x_1 + x_2 - 2) = 0 \\ x_1 + x_2 - 2 \leq 0 \\ -x_1 + x_2 - 2 \leq 0 \end{cases}$$

Se $\lambda_1 = \lambda_2 = 0$, allora ricaviamo la soluzione

$$x^1 = (-1, -1), \quad \lambda^1 = (0, 0).$$

Se $\lambda_1 = 0$ $\lambda_2 \neq 0$, un'altra soluzione del sistema è

$$x^2 = (-2, 0), \quad \lambda^2 = (0, -2).$$

Se $\lambda_1 \neq 0$e $\lambda_2 = 0$, ricaviamo la soluzione

$$x^3 = (1, 1), \quad \lambda^3 = (-4, 0).$$

Infine se $\lambda_1 \neq 0$ e $\lambda_2 \neq 0$, otteniamo la soluzione

$$x^4 = (0, 2), \quad \lambda^4 = (-4, -2).$$

Poichè x^1 è l'unico punto candidato ad essere un minimo, si ricava che x^1 è il minimo globale. I punti x^2, x^3 e x^4 sono candidati ad essere massimi locali. Non possiamo usare le condizioni sufficienti del primo ordine perchè f non è concava. Passiamo quindi alle condizioni del secondo ordine. La matrice hessiana della Lagrangiana non dipende dal punto (x, λ) e vale

$$H_{xx}L(x, \lambda) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Nel punto x^2 è attivo solo il secondo vincolo con $\lambda_2^2 \neq 0$, quindi il cono critico $C(x^2, \lambda^2)$ è

$$C(x^2, \lambda^2) = \{d \in \mathbb{R}^2 : \ -d_1 + d_2 = 0\} = \{(t, t), t \in \mathbb{R}\}.$$

La condizione necessaria del secondo ordine per un massimo locale non è soddisfatta in x^2, infatti se $d \in C(x^2, \lambda^2)$ si ha

$$dH_{xx}L(x^2, \lambda^2)\, d = (t, t) \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} t \\ t \end{pmatrix} = 4\, t^2 > 0 \qquad \text{per ogni } t \neq 0,$$

quindi x^2 è una sella.
Nel punto x^3 è attivo solo il primo vincolo con $\lambda_1^3 \neq 0$, quindi

$$C(x^3, \lambda^3) = \{d \in \mathbb{R}^2 : \; d_1 + d_2 = 0\} = \{(t, -t), t \in \mathbb{R}\}.$$

Anche in questo caso non è soddisfatta la condizione necessaria del secondo ordine per un massimo locale, infatti se $d \in C(x^3, \lambda^3)$ si ha

$$dH_{xx}L(x^3, \lambda^3)\, d = (-t, t) \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} t \\ -t \end{pmatrix} = 4\, t^2 > 0 \qquad \text{per ogni } t \neq 0,$$

quindi x^3 è una sella.

Nel punto x^4 entrambi i vincoli sono attivi con $\lambda_1^4 \neq 0$ e $\lambda_2^4 \neq 0$, quindi

$$C(x^4, \lambda^4) = \{d \in \mathbb{R}^2 : \; d_1 + d_2 = 0, \quad -d_1 + d_2 = 0\} = \{(0,0)\}.$$

La condizione sufficiente del secondo ordine per un massimo locale è ovviamente soddisfatta non essendoci direzioni non nulle in $C(x^4, \lambda^4)$, quindi x^4 è un massimo locale.

Cerchiamo massimi e minimi della funzione $f(x) = x_1 - x_2^2$ sulla regione $D = \{x \in \mathbb{R}^2 : \ 4 - x_1^2 - x_2^2 \leq 0\}$ che non è limitata. Notiamo che nè f ne $-f$ sono coercive perchè sono nulle sulla parabola $x_1 = x_2^2$. Osserviamo però che f non è limitata nè superiormente nè inferiormente su Dperchè $\lim_{t\to+\infty} f(t,0) = +\infty$ e $\lim_{t\to-\infty} f(t,0) = -\infty$, quindi non esistono nè massimi globali nè minimi globali di f su D.

Passiamo ora allo studio dei massimi e minimi locali. Il vincolo è regolare in tutti i punti di D perchè è soddisfatta la condizione di Mangasarian. Risolviamo il sistema LKKT:

$$\begin{cases} 1 - 2\,\lambda\,x_1 = 0 \\ -2\,x_2 - 2\,\lambda\,x_2 = 0 \\ \lambda\,(4 - x_1^2 - x_2^2) = 0 \\ 4 - x_1^2 - x_2^2 \leq 0 \end{cases}$$

Se $\lambda = 0$ la prima equazione diventa impossibile. Se $\lambda \neq 0$, allora il sistema diventa

$$\begin{cases} 1 - 2\,\lambda\, x_1 = 0 \\ x_2\,(1+\lambda) = 0 \\ 4 - x_1^2 - x_2^2 = 0 \end{cases}$$

Se $x_2 = 0$ dalla terza equazione si ha $x_1^2 = 4$ e quindi troviamo due soluzioni:

$$x^1 = (2,0), \quad \lambda^1 = \frac{1}{4}, \qquad x^2 = (-2,0), \quad \lambda^2 = -\frac{1}{4}.$$

Se $x_2 \neq 0$ dalla seconda equazione si ha $\lambda = -1$ e troviamo due soluzioni:

$$x^3 = \left(-\frac{1}{2}, \frac{\sqrt{15}}{2}\right), \quad \lambda^3 = -1, \qquad x^4 = \left(-\frac{1}{2}, -\frac{\sqrt{15}}{2}\right), \quad \lambda^4 = -1.$$

Il punto x^1 è candidato ad essere un minimo locale, mentre x^2, x^3 e x^4 ad essere massimi locali. Non possiamo usare le condizioni sufficienti del primo ordine perchè la funzione $g(x) = 4 - x_1^2 - x_2^2$ non è convessa. Passiamo quindi alle condizioni del secondo ordine. La matrice hessiana della Lagrangiana è

$$H_{xx}L(x,\lambda) = \begin{pmatrix} -2\,\lambda & 0 \\ 0 & -2 - 2\,\lambda \end{pmatrix}.$$

In x^1 il vincolo è attivo con $\lambda^1 \neq 0$, quindi

$$C(x^1, \lambda^1) = \{d \in \mathbb{R}^2 : \; -4\, d_1 = 0\} = \{(0, t), t \in \mathbb{R}\}.$$

La condizione necessaria del secondo ordine per un minimo locale non è soddisfatta, infatti se $d \in C(x^1, \lambda^1)$ si ha

$$dH_{xx}L(x^1, \lambda^1)\, d = (0, t) \begin{pmatrix} -1/2 & 0 \\ 0 & -5/2 \end{pmatrix} \begin{pmatrix} 0 \\ t \end{pmatrix} = -\frac{5}{2}\, t^2 < 0 \qquad \forall\, t \neq 0,$$

quindi x^1 è una sella.

Nel punto x^2 il vincolo è attivo con $\lambda^2 \neq 0$, quindi

$$C(x^2, \lambda^2) = \{d \in \mathbb{R}^2 : \; 4\, d_1 = 0\} = \{(0, t), t \in \mathbb{R}\}.$$

In questo caso è soddisfatta la condizione sufficiente del secondo ordine per un massimo locale, infatti se $d \in C(x^2, \lambda^2)$ si ha

$$dH_{xx}L(x^2, \lambda^2)\, d = (0, t) \begin{pmatrix} 1/2 & 0 \\ 0 & -3/2 \end{pmatrix} \begin{pmatrix} 0 \\ t \end{pmatrix} = -\frac{3}{2}\, t^2 < 0 \qquad \forall\, t \neq 0,$$

quindi x^2 è un massimo locale.

Il vincolo è attivo in x^3 con $\lambda^3 \neq 0$, quindi

$$C(x^3, \lambda^3) = \{d \in \mathbb{R}^2 : \; d_1 - \sqrt{15}\, d_2 = 0\} = \{(\sqrt{15}\, t, t), t \in \mathbb{R}\}.$$

La condizione necessaria del secondo ordine per un massimo locale non è soddisfatta, infatti se $d \in C(x^3, \lambda^3)$ si ha

$$dH_{xx}L(x^3, \lambda^3)\, d = (\sqrt{15}\, t, t) \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{15}\, t \\ t \end{pmatrix} = 30\, t^2 > 0 \qquad \forall \; t \neq 0,$$

quindi x^3 è una sella.

Nel punto x^4 il vincolo è attivo con $\lambda^4 \neq 0$, quindi

$$C(x^4, \lambda^4) = \{d \in \mathbb{R}^2 : \; d_1 + \sqrt{15}\, d_2 = 0\} = \{(-\sqrt{15}\, t, t), t \in \mathbb{R}\}.$$

La condizione necessaria del secondo ordine per un massimo locale non è soddisfatta, infatti se $d \in C(x^4, \lambda^4)$ si ha

$$d H_{xx} L(x^4, \lambda^4)\, d = (-\sqrt{15}\, t, t) \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -\sqrt{15}\, t \\ t \end{pmatrix} = 30\, t^2 > 0 \qquad t \neq 0$$

quindi x^4 è una sella.

Cerchiamo massimi e minimi della funzione $f(x) = x_1^3 + x_2^3$ sulla regione

$$D = \{x \in \mathbb{R}^2 : \ x_1 - 1 \leq 0, \quad x_2 - 1 \leq 0\}$$

che non è limitata. Osserviamo che nè f nè $-f$ sono coercive, tuttavia f non è limitata inferiormente su D perchè $\lim_{t\to-\infty} f(t,0) = -\infty$, quindi non esistono minimi globali. Inoltre dai vincoli ricaviamo che $f(x) \leq 2$ per ogni $x \in D$ e che $f(1,1) = 2$, quindi il punto $(1,1)$ è il massimo globale. Passiamo allo studio dei massimi e minimi locali. I vincoli sono lineari e quindi sono regolari in tutti i punti di D. Risolviamo il sistema LKKT:

$$\begin{cases} 3\,x_1^2 + \lambda_1 = 0 \\ 3\,x_2^2 + \lambda_2 = 0 \\ \lambda_1\,(x_1 - 1) = 0 \\ \lambda_2\,(x_2 - 1) = 0 \\ x_1 - 1 \leq 0 \\ x_2 - 1 \leq 0 \end{cases}$$

Distinguendo quattro casi in base al segno di λ_1 e λ_2 otteniamo quattro soluzioni:

$$\begin{array}{ll} x^1 = (0,0), \lambda^1 = (0,0) & x^2 = (0,1), \lambda^2 = (0,-3) \\ x^3 = (1,0), \lambda^3 = (-3,0) & x^4 = (1,1), \lambda^4 = (-3,-3). \end{array}$$

Il punto x^1 è candidato ad essere sia un massimo locale che un minimo locale, x^2 e x^3 sono candidati ad essere massimi locali, mentre sappiamo già che x^4 è il massimo globale. Non possiamo usare le condizioni sufficienti del primo ordine perchè nè f nè $-f$ sono convesse su D. Passiamo alle condizioni del secondo ordine. La matrice hessiana della Lagrangiana è

$$H_{xx}L(x,\lambda) = \begin{pmatrix} 6\,x_1 & 0 \\ 0 & 6\,x_2 \end{pmatrix}.$$

In x^1 nessuno dei due vincoli è attivo quindi $C(x^1,\lambda^1) = \mathbb{R}^2$. Inoltre la matrice $H_{xx}L(x^1,\lambda^1) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

quindi sono soddisfatte le condizioni necessarie del secondo ordine sia per un massimo locale che per un minimo locale, ma non è soddisfatta nè la condizione sufficiente per un massimo locale nè quella per un minimo locale. In questo caso quindi non bastano le condizioni del secondo ordine per analizzare il punto x^1. Notiamo però che $f(x^1) = 0$ e

$$f(x) \left\{ \begin{array}{ll} > 0 & \text{se } x_1, x_2 > 0 \\ < 0 & \text{se } x_1, x_2 < 0 \end{array} \right\},$$

quindi x^1 è una sella.
Nel punto x^2 è attivo il secondo vincolo con $\lambda_2^2 \neq 0$, quindi

$$C(x^2, \lambda^2) = \{d \in \mathbb{R}^2 : \ d_2 = 0\} = \{(t, 0), t \in \mathbb{R}\}.$$

Per ogni $d \in C(x^2, \lambda^2)$ si ha

$$dH_{xx}L(x^2, \lambda^2)\, d = (t, 0) \begin{pmatrix} 0 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} t \\ 0 \end{pmatrix} = 0,$$

quindi è soddisfatta la condizione necessaria del secondo ordine per un massimo locale, ma non quella sufficiente. Anche in questo caso le condizioni del secondo ordine non bastano per studiare il punto x^2. Osserviamo però che $f(x^2) = 1$ e che $f(t, 1) > 1$ se $t \in (0, 1]$, quindi x^2 è una sella.

Nel punto x^3 è attivo il primo vincolo con $\lambda_1^3 \neq 0$, quindi

$$C(x^3, \lambda^1) = \{d \in \mathbb{R}^2 : \; d_1 = 0\} = \{(0, t), t \in \mathbb{R}\}.$$

Per ogni $d \in C(x^3, \lambda^3)$ si ha

$$d H_{xx} L(x^3, \lambda^3)\, d = (0, t) \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ t \end{pmatrix} = 0,$$

quindi anche in questo caso le condizioni del secondo ordine non bastano per studiare il punto x^3 perchè è soddisfatta la condizione necessaria del secondo ordine per un massimo locale, ma non quella sufficiente. Notiamo però che $f(x^3) = 1$ e che $f(1, t) > 1$ se $t \in (0, 1]$, quindi x^3 è una sella.

Esercizio. Sia data $f(x_1, x_2) = 2\,x_1^2 + x_2$ e la regione ammissibile definita dai vincoli $x_1^2 + x_2^2 = 1$ e $x_2 \geq 0$. Trovare i massimi ed i minimi.

Esistono massimi e minimi globali perchè la f è continua e la regione $\mathcal{R}$ è compatta. I vincoli sono regolari perchè in ogni punto i gradienti dei vincoli attivi sono linearmente indipendenti. Il sistema LKKT associato al problema è:

$$x_1\,(2 + \mu) = 0 \tag{8}$$

$$1 + 2\,\mu\,x_2 - \lambda = 0 \tag{9}$$

$$x_1^2 + x_2^2 = 1 \tag{10}$$

$$\lambda\,x_2 = 0 \tag{11}$$

$$x_2 \geq 0 \tag{12}$$

Distinguiamo due casi: $\lambda = 0$ e $\lambda \neq 0$. Se $\boxed{\lambda \neq 0}$, allora dalla (11) si ottiene $x_2 = 0$. Sostituendo in (9) e in (10) si ottiene $\lambda = 1$ e $x_1 = \pm 1$. Inoltre dalla (8) si ha $\mu = -2$. Se $\boxed{\lambda = 0}$, allora il sistema KKT diventa:

$$\begin{aligned} x_1\,(2+\mu) &= 0 &\quad (13)\\ 1 + 2\,\mu\,x_2 &= 0 &\quad (14)\\ x_1^2 + x_2^2 &= 1 &\quad (15)\\ x_2 &\geq 0 &\quad (16) \end{aligned}$$

Distinguiamo due casi: $x_1 = 0$ e $x_1 \neq 0$. Se $x_1 = 0$, allora $x_2 = 1$ e $\mu = -1/2$. Se $x_1 \neq 0$, allora $\mu = -2$. Sostituendo in (14) si ha $x_2 = 1/4$ e sostituendo in (15) si trova $x_1 = \pm\sqrt{15}/4$. Le soluzioni di LKKT sono:

$$\begin{array}{llll} x = (1,0) & \lambda = 1 & \mu = -2 & f(x) = 2 \\ x = (-1,0) & \lambda = 1 & \mu = -2 & f(x) = 2 \\ x = (0,1) & \lambda = 0 & \mu = -1/2 & f(x) = 1 \\ x = \left(\frac{\sqrt{15}}{4}, \frac{1}{4}\right) & \lambda = 0 & \mu = -2 & f(x) = 17/8 \\ x = \left(-\frac{\sqrt{15}}{4}, \frac{1}{4}\right) & \lambda = 0 & \mu = -2 & f(x) = 17/8 \end{array}$$

Il punto di minimo globale è $(0,1)$ mentre i punti di massimo globale sono $\left(\frac{\sqrt{15}}{4}, \frac{1}{4}\right)$ e $\left(-\frac{\sqrt{15}}{4}, \frac{1}{4}\right)$.

Esercizio. Sia data $f(x_1, x_2) = (x_1 + 2)^2 + 4\,x_2^2$ e i vincoli $x_1^2 - x_2^2 = 1$ e $x_1 \leq 0$. Il minimo globale esiste perchè la f è coerciva. Il massimo globale non esiste perchè

$$\lim_{x_1 \to -\infty} f(x) = +\infty.$$

I vincoli sono regolari perchè in ogni punto i gradienti dei vincoli attivi sono linearmente indipendenti. Il sistema LKKT è:

$$2(x_1 + 2) + 2\,\mu\,x_1 + \lambda = 0 \tag{17}$$
$$8\,x_2 - 2\,\mu\,x_2 = 0 \tag{18}$$
$$x_1^2 - x_2^2 = 1 \tag{19}$$
$$\lambda\,x_1 = 0 \tag{20}$$
$$x_1 \leq 0 \tag{21}$$

Distinguiamo due casi: $\lambda = 0$ e $\lambda \neq 0$. Se $\boxed{\lambda \neq 0}$, allora dalla (20) si ottiene $x_1 = 0$, ma sostituendo nella (19) si ottiene $x_2^2 = -1$ che non ha soluzioni reali. Se $\boxed{\lambda = 0}$:

$$
\begin{aligned}
x_1 + 2 + \mu\, x_1 &= 0 && (22)\\
x_2\,(4 - \mu) &= 0 && (23)\\
x_1^2 - x_2^2 &= 1 && (24)\\
x_1 &\leq 0 && (25)
\end{aligned}
$$

Distinguiamo due casi: $x_2 = 0$ e $x_2 \neq 0$. Se $x_2 = 0$, allora $x_1 = \pm 1$. La soluzione con $x_1 = 1$ non è ammissibile. Per $x_1 = -1$ si ha $\mu = 1$. Se $x_2 \neq 0$, allora $\mu = 4$. Sostituendo nella (22) si ha $x_1 = -2/5$, ma sostituendo nella (15) si trova $x_2^2 = -21/25$ che non ha soluzioni reali. Riassumendo, l'unica soluzione del sistema LKKT è $(-1, 0)$ che è anche il punto di minimo globale.

Così come abbiamo visto per i problemi non vincolati, anche per un problema vincolato gli algoritmi forniscono soluzioni stazionarie; in questo caso soluzioni che soddisfano la condizione LKKT.

Consideriamo un problema di PNL vincolato su un poliedro limitato:

$$\begin{cases} \min f(x) \\ Ax \leq b, \end{cases} \tag{26}$$

Al generico passo k, l'algoritmo di Frank-Wolfe, prima trova un vertice ottimo y^k del problema linearizzato nel punto corrente x^k, poi minimizza la funzione f sull'intervallo $[x^k, y^k]$.

ALGORITMO DI FRANK-WOLFE

1. Poni $k = 0$ e scegli un punto x^0 ammissibile.

2. Costruisci il problema linearizzato nel punto x^k:

$$\begin{cases} \min \ \nabla f(x^k)^T x \\ Ax \le b \end{cases} \qquad (\mathrm{PL}(x^k))$$

e calcola un vertice ottimo y^k del problema linearizzato $\mathrm{PL}(x^k)$.

3. Calcola:

$$t^k \in \arg \min_{0 \le t \le 1} f(x^k + t\,(y^k - x^k))$$

e poni $x^{k+1} = x^k + t^k(y^k - x^k)$.

4. **Se** un criterio di arresto è soddisfatto
 allora STOP
 altrimenti poni $k = k + 1$ e torna al passo 2.

Teorema

Per ogni x^0, l'algoritmo genera una successione $\{x^k\}$ i cui punti di accumulazione sono soluzioni del sistema LKKT.

Esempio

$$\begin{cases} \min\ \frac{1}{2}\,x_1^2 + x_2^2 - 2\,x_1\,x_2 + x_1 \\ -3\,x_1 + 2\,x_2 \leq 7 \\ x_1 + 2\,x_2 \leq 4 \\ 0 \leq x_2 \leq 2. \end{cases} \tag{27}$$

Se $x^0 = (0,1)$ si ha che $\nabla f(x^0) = (-1,2)$; il vertice ottimo di:

$$\begin{cases} \min -x_1 + 2\,x_2 \\ -3\,x_1 + 2\,x_2 \leq 7 \\ x_1 + 2\,x_2 \leq 4 \\ 0 \leq x_2 \leq 2 \end{cases}$$

è $y^0 = (4,0)$.

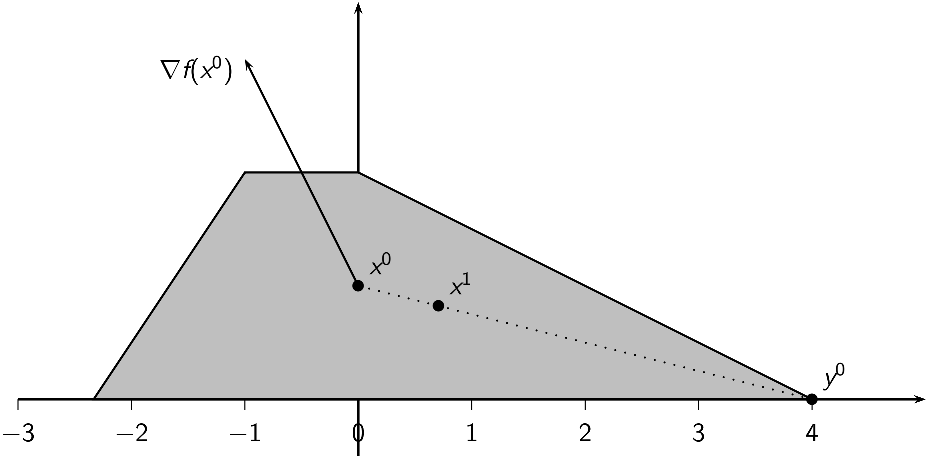
Occorre determinare il minimo della funzione $f(x)$, ristretta al segmento di estremi $(0, 1)$ e $(4, 0)$, le cui equazioni parametriche sono:

$$\begin{cases} x_1(t) = 4\,t \\ x_2(t) = 1 - t. \end{cases}$$

Indicando con φ la restrizione di f al segmento, si ha:

$$\varphi(t) = \frac{1}{2}(4t)^2 + (1-t)^2 - 8\,t\,(1-t) + 4\,t = 17\,t^2 - 6\,t + 1,$$

che assume minimo per $t = \frac{3}{17}$, quindi $x^1 = \left(\frac{12}{17}, \frac{14}{17}\right)$.



$\nabla f(x^0)$
x^0
x^1
y^0
−3
−2
−1
0
1
2
3
4

Algoritmo del gradiente proiettato

Se il punto x^k è interno alla regione ammissibile, viene scelta $-\nabla f(x^k)$ come direzione di discesa; se invece x^k appartiene alla frontiera della regione ammissibile, si sceglie la proiezione di $-\nabla f(x^k)$ sul sottospazio vettoriale definito dall'insieme $\mathcal{I}$ dei vincoli attivi nel punto x^k, cioè $\mathcal{I} = \{i : \ A_i x^k = b_i\}$.

ALGORITMO DEL GRADIENTE PROIETTATO

1. Sia x^0 ammissibile; $\mathcal{I}$ l'insieme dei vincoli attivi in x^k e la matrice M, avente le righe A_i della matrice A, per ogni $i \in \mathcal{I}$.
2. Costruisci la matrice di proiezione:

$$H = \begin{cases} I & \text{se } \mathcal{I} = \emptyset \\ I - M^T(MM^T)^{-1}M & \text{se } \mathcal{I} \neq \emptyset \end{cases}$$

   e calcola la direzione $d^k = -H\nabla f(x^k)$.

3. **Se** $d^k = 0$ **allora** vai al passo 4, **altrimenti** sia $\hat{t}^k$ la soluzione di:

$$\begin{cases} \max\ t \\ A(x^k + t\,d^k) \leq b \end{cases}$$

$$\textit{sia } t^k \in \arg\min_{0 \leq t \leq \hat{t}^k} f(x^k + t\,d^k),$$

   poni $x^{k+1} = x^k + t^k\,d^k$, $k = k+1$ e torna al passo 2.

4. Calcola $u = -(MM^T)^{-1}M\nabla f(x^k)$. **Se** $u \geq 0$ **allora** STOP (x^k è stazionario), **altrimenti** sia $u_j = \min\{u_i :\ i \in \mathcal{I}\}$, togli da M la riga A_j e vai al passo 2.

Lemma
Per ogni direzione $y \in \mathbb{R}^n$, Hy è la proiezione ortogonale di y sul sottospazio vettoriale $S = \{x \in \mathbb{R}^n : \ Mx = 0\}$.

Teorema

1. se $d_k \neq 0$, allora f decresce lungo la direzione d_k;
2. se $d_k = 0$ e $u \geq 0$, allora (x_k, u) soddisfa le condizioni LKKT.

Esempio

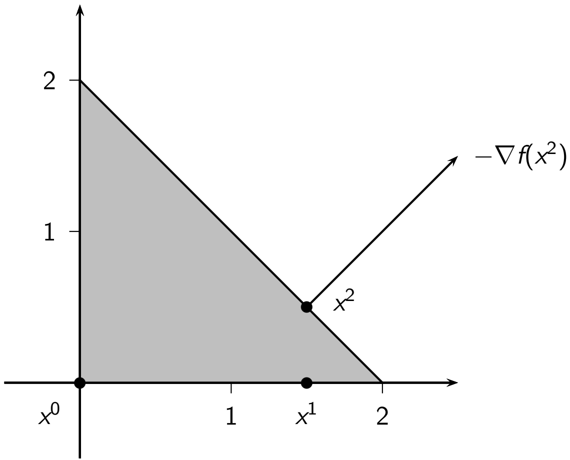
Consideriamo il problema:

$$\begin{cases} \min 2\,x_1^2 + 2\,x_2^2 - 2\,x_1\,x_2 - 6\,x_1 \\ x_1 + x_2 \leq 2 \\ x_1 \geq 0 \\ x_2 \geq 0. \end{cases}$$

Partiamo da $x^0 = (0,0)$, quindi $M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $H = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ e $d^0 = (0,0)$. Si esegue il passo 5: $u = (-6,0)$. Si calcola $M = (0,-1)$ e poi $H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $d^0 = (6,0)$. Si esegue il passo 4, ottenendo $\hat{t}^0 = \frac{1}{3}$ e $t^0 = \frac{1}{4}$, perciò $x^1 = (\frac{3}{2}, 0)$.

Si esegue, ora, il passo 2: $M = (0,-1)$, $H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ e $d^1 = (0,0)$. Si esegue il passo 5: $u = -3$, $M = 0$ e $H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. La direzione diventa $d^1 = (0,3)$ quindi abbiamo $\hat{t}^1 = \frac{1}{6}$ e $t^1 = \frac{1}{6}$, quindi $x^2 = (\frac{3}{2}, \frac{1}{2})$.

Si esegue il passo 2: $M = (1,1)$, $H = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ e $d^2 = (0,0)$. Si esegue il passo 5: $u = 1$, quindi, x^2 soddisfa le condizioni KKT (esso è anche un minimo globale, poichè f è convessa).



Esercizio 1. Confrontare un'iterazione del gradiente proiettato con una di Frank-Wolfe da $(2,0,2)$ per

$$\begin{cases} \min\ x_1^3 + 2\,x_2^3 + 2\,x_3^2 - 10\,x_1 - 4\,x_2 - 7\,x_3 + 50 \\ 2\,x_1 + x_2 + 3\,x_3 \leq 10 \\ x_i \geq 0 \end{cases}$$

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} 3\,x_1^2 - 10 \\ 6\,x_2^2 - 4 \\ 4\,x_3 - 7 \end{pmatrix}.$$

$f(x^0) = 32$ e $\nabla f(x^0) = \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix}.$

I due vincoli attivi in x^0 sono $2\,x_1 + x_2 + 3\,x_3 \leq 10$ e $x_2 \geq 0$, quindi si ha la matrice

$$M = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \end{pmatrix}.$$

La direzione di spostamento è $d^0 = -H\nabla f(x^0)$ con
$H = I - M^T(MM^T)^{-1}M$.

Otteniamo $d^0 = \begin{pmatrix} -\frac{12}{13} \\ 0 \\ \frac{8}{13} \end{pmatrix}$ e la semiretta $x^0 + t\,d^0 = \begin{pmatrix} 2 - \frac{12}{13}\,t \\ 0 \\ 2 + \frac{8}{13}\,t \end{pmatrix}$ e calcoliamo $\hat{t}_0$:

$$\begin{array}{lll} 2\,x_1 + x_2 + 3\,x_3 \leq 10 & \text{è soddisfatto} & \text{per ogni } t \in \mathbb{R} \\ x_1 \geq 0 & \text{è soddisfatto} & \text{per } t \leq \frac{13}{6} \\ x_2 \geq 0 & \text{è soddisfatto} & \text{per ogni } t \in \mathbb{R} \\ x_3 \geq 0 & \text{è soddisfatto} & \text{per } t \geq -\frac{13}{4} \end{array}$$

Il passo si ottiene per $t \simeq 0.107$, pertanto il nuovo punto è
$x^1 \simeq \begin{pmatrix} 1.90 \\ 0 \\ 2.07 \end{pmatrix}$ e $f(x^1) \simeq 31.9$.

Il vertice ottimo del linearizzato è $y^0 = \begin{pmatrix} 0 \\ 10 \\ 0 \end{pmatrix}$ e la direzione è
$d^0 = y^0 - x^0 = \begin{pmatrix} -2 \\ 10 \\ -2 \end{pmatrix}$.
L'ottimo si ottiene per $t \simeq 0.082$, perciò il nuovo punto è
$x^1 \simeq \begin{pmatrix} 1.83 \\ 0.83 \\ 1.83 \end{pmatrix}$ e $f(x^1) \simeq 29.5$.

Esercizio 2.

$$\begin{cases} \min \ \left((x_1-5)^2 + (2-x_2)^2\right) \\ 2 \le x_1 \le 4 \\ 1 \le x_2 \le 3 \end{cases}$$

Fare un'iterazione del metodo di Frank-Wolfe da $x^0 = (3,3)$.

sol. ottima del problema linearizzato in x^0	(4,1)
direzione	$(1,-2)$
passo	$4/5$
x^1	(19/5, 7/5)

Fare un'iterazione del metodo del gradiente proiettato da $x^0 = (3, 3)$.

matrice M	(0, 1)
matrice H	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
direzione	(2/3, 0)
max spostamento possibile lungo la direzione	3/2
passo	3/2
x^1	(4, 3)