Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Giovedì 4 Aprile 2019

[1] Istituto di Analisi dei Sistemi ed Informatica IASI - CNR

Esempio 1

Consideriamo il seguente problema vincolato (banale)

$$\min \; x^2$$
$$x \geq 1$$

Ammete $x^* = 1$ come **unico** punto di minimo globale

La funzione di penalità esterna è:

$$P_\epsilon(x) = x^2 + \frac{1}{\epsilon} \max\{0, 1 - x\}^2$$

Esempio 1

Consideriamo il seguente problema vincolato (banale)

$$\begin{aligned} \min \ & x^2 \\ & x \geq 1 \end{aligned}$$

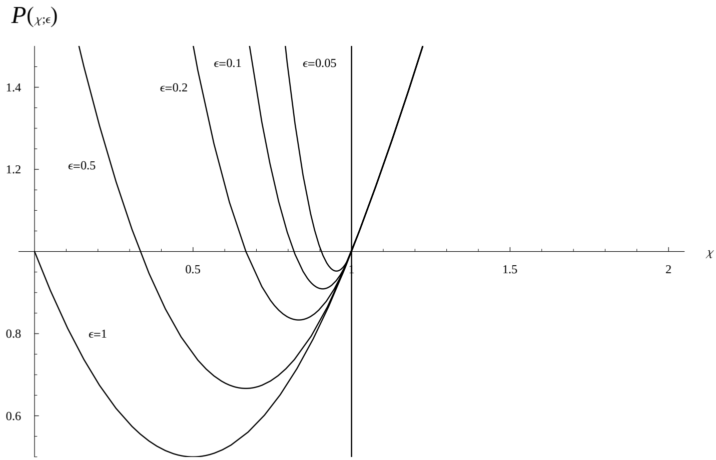
Ammete $x^* = 1$ come **unico** punto di minimo globale

La funzione di penalità esterna è:

$$P_\epsilon(x) = x^2 + \frac{1}{\epsilon}\max\{0, 1 - x\}^2$$

Esempio 1

Grafici di $P_\epsilon(x)$ ($\epsilon = 1, 0.5, 0.2, 0.1, 0.05$)



Esempio 2

Consideriamo il seguente problema vincolato (banale)

$$\min \; -x^4$$
$$x = 1$$

Ammete (ovviamente) $x^* = 1$ come **unico** punto di minimo globale

La funzione di penalità esterna è:

$$P_\epsilon(x) = -x^4 + \frac{1}{\epsilon}(x-1)^2$$

Esempio 2

Consideriamo il seguente problema vincolato (banale)

$$\begin{gathered}\min\ -x^4\\ x = 1\end{gathered}$$

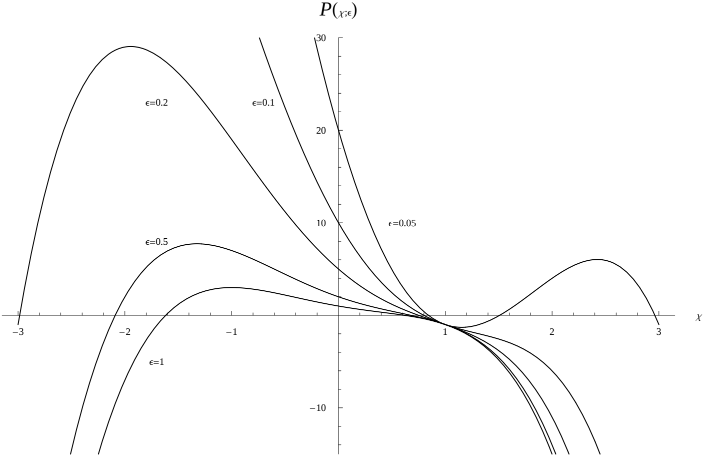
Ammete (ovviamente) $x^* = 1$ come **unico** punto di minimo globale

La funzione di penalità esterna è:

$$P_\epsilon(x) = -x^4 + \frac{1}{\epsilon}(x-1)^2$$

Esempio 2

Grafici di $P_\epsilon(x)$ ($\epsilon = 1, 0.5, 0.2, 0.1, 0.05$)



Algoritmo SEQPEN

Algoritmo `SEQPEN`

INPUT: maxit, $\tau > 0$, $\epsilon_0 > 0$

for $k = 0, 1, \ldots, \texttt{maxit}$

Trova una soluzione globale (x^k) del problema

$$\min_{x \in \mathbb{R}^n} P_{\epsilon_k}(x). \tag{1}$$

if $q(x^k) \leq \tau$ **then** STOP

$\epsilon_{k+1} = \epsilon_k / 2$

endfor

OUTPUT: una sol. approssimata $x^* = x^k$

Algoritmo SEQPEN

L'algoritmo che abbiamo visto:

- ad ogni passo richiede di risolvere all'ottimo **gloable**
$$\min_{x \in \mathbb{R}^n} \; P_\epsilon(x);$$
- è (o può essere) **molto costoso**;
- non sempre è possibile;
- gli algoritmi di min. non vincolata determinano $\hat{x}$ t.c.
$$\|\nabla P_\epsilon(\hat{x})\| \leq \texttt{tol}.$$

Algoritmo SEQPEN modificato (1)

Algoritmo $\texttt{SEQPEN}_{mod}$

INPUT: $\texttt{maxit}$, $\rho > 0$, $k = 0$, $\epsilon_0 > 0$

for $k = 0, 1, \ldots, \texttt{maxit}$

Calcola x^k t.c. $\|\nabla P_{\epsilon_k}(x^k)\| \leq \rho$

if $q(x^k) \leq \rho$ **then** STOP

$\epsilon_{k+1} = \epsilon_k / 2$

endfor

OUTPUT: una sol. approssimata $x^* = x^k$

Algoritmo SEQPEN modificato (1)

Algoritmo $\texttt{SEQPEN}_{mod}$

INPUT: $\texttt{maxit}$, $\{\tau_k\} \to 0$, $\rho > 0$, $k = 0$, $\epsilon_0 > 0$

for $k = 0, 1, \dots, \texttt{maxit}$

Calcola x^k t.c. $\|\nabla P_{\epsilon_k}(x^k)\| \leq \tau_k$

if $\tau_k < \rho$ **and** $q(x^k) \leq \rho$ **then** STOP

$\epsilon_{k+1} = \epsilon_k / 2$

endfor

OUTPUT: una sol. approssimata $x^* = x^k$

Proprietà di convergenza

Teorema

Sia $\{x_k\}$ *una successione t.c.* $\|\nabla P_{\epsilon_k}(x_k)\| \leq \tau_k$ *e* $\{x_k\}_K$ *una sottosucc. convergente ad* x^* *t.c. siano lin. indipendenti*

$$\nabla g_i(x^*),\ i \in I_0(x^*),\ \nabla h_j(x^*),\ \forall\, j.$$

Allora, x^* *è un punto di KKT di* (P_0) *con molt.* λ^* *e* μ^* *t.c.*

$$\left\{\frac{2}{\epsilon_k}\max\{g(x_k), 0\}\right\}_K \to \lambda^*$$

$$\left\{\frac{2}{\epsilon_k} h(x_k)\right\}_K \to \mu^*$$

Un po' di esempi di risoluzione

`esempio3`:

$$\begin{aligned} &\min\ -x_1 - x_2 \\ &s.t.\ x_1^2 + x_2^2 = 1 \end{aligned}$$

$x^0 = (-1,\ -1)^\top$
Ammete $(x^*, y^*) = (1,\ 1)/\sqrt{2}$ come **unico** punto di minimo globale

La funzione di penalità esterna è:

$$P_\epsilon(x) = -x - y + \frac{1}{\epsilon}(1 - x^2 - y^2)^2$$

Un po' di esempi di risoluzione

`esempio3`:

$$\begin{array}{l} \min \; -x_1 - x_2 \\ s.t. \; x_1^2 + x_2^2 = 1 \end{array}$$

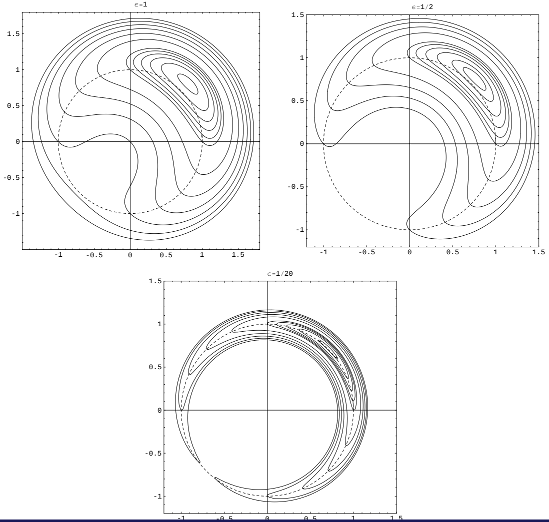
$x^0 = (-1,\ -1)^\top$
Ammete $(x^*, y^*) = (1,\ 1)/\sqrt{2}$ come **unico** punto di minimo globale

La funzione di penalità esterna è:

$$P_\epsilon(x) = -x - y + \frac{1}{\epsilon}(1 - x^2 - y^2)^2$$

Un po' di esempi di risoluzione

Grafici di $P_\epsilon(x)$ ($\epsilon = 1, 0.5, 0.05$)



Un po' di esempi di risoluzione

`esempio4`:

$$\begin{array}{l} \min\ -x_1 - x_2 \\ s.t.\ x_1^2 + x_2^2 \leq 1 \end{array}$$

$x^0 = (-1,\ -1)^\top$

Un po' di esempi di risoluzione

`maratos`:

$$\begin{array}{l} \min \ -x_1 + 2(x_1^2 + x_2^2 - 1) \\ s.t. \ x_1^2 + x_2^2 = 1 \end{array}$$

$x^0 = (-1,\ -1)^\top$

Un po' di esempi di risoluzione

`hs14`:

$$\begin{aligned} \min\ & (x_1-2)^2+(x_2-1)^2 \\ s.t.\ & x_1^2/4+x_2^2 \le 1 \\ & x_1-2x_2=-1 \end{aligned}$$

$x^0 = (2,\ 2)^\top$

Un po' di esempi di risoluzione

`hs24`:

$$
\begin{aligned}
\min\ & x_2^3((x_1-3)^2-9)/(27\sqrt{3}) \\
s.t.\ & x_1/\sqrt{3}-x_2 \geq 0 \\
& x_1+\sqrt{3}x_2 \geq 0 \\
& -x_1-\sqrt{3}x_2+6 \geq 0 \\
& x_1 \geq 0,\ \ x_2 \geq 0
\end{aligned}
$$

$x^0 = (1,\ 1/2)^\top$

Un po' di esempi di risoluzione

`hs32`:

$$\begin{aligned} \min\ & (x_1 + 3 * x_2 + x_3)^2 + 4(x_1 - x_2)^2 \\ s.t.\ & x_1 + x_2 + x_3 = 1 \\ & 6x_2 + 4x_3 - x_1^3 \geq 3 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{aligned}$$

$x^0 = (0.1,\ 0.7,\ 0.2)^\top$

Un po' di esempi di risoluzione

hs41:

$$\begin{aligned} &\min\ 2 - x_1x_2x_3 \\ &s.t.\ x_1 + 2x_2 + 2x_3 - x_4 = 0 \\ &\qquad 0 \leq x_i \leq 1,\ \ i = 1,2,3,4 \end{aligned}$$

$x^0 = (2,\ 2,\ 2,\ 2)^\top$

Un po' di esempi di risoluzione

`hs55`:

$$
\begin{aligned}
\min\ & x_1 + 2x_2 + 4x_5 + e^{x_1 x_4} \\
s.t.\ & x_1 + 2x_2 + 5x_5 - 6 = 0 \\
& x_1 + x_2 + x_3 - 3 = 0 \\
& x_4 + x_5 + x_6 - 2 = 0 \\
& x_1 + x_4 - 1 = 0 \\
& x_2 + x_5 - 2 = 0 \\
& x_3 + x_6 - 2 = 0 \\
& x_i \geq 0,\ \ i = 1, \ldots, 6 \\
& x_1 \leq 1,\ \ x_4 \leq 1
\end{aligned}
$$

$x^0 = (1,\ 2,\ 0,\ 0,\ 0,\ 2)^\top$

Un po' di esempi di risoluzione

`hs60`:

$$
\begin{aligned}
&\min\ (x_1-1)^2+(x_1-x_2)^2+(x_2-x_3)^4 \\
&s.t.\ x_1(1+x_2^2)+x_3^4-4-3\sqrt{2}=0 \\
&\qquad -10 \le x_i \le 10,\ \ i=1,2,3
\end{aligned}
$$

$x^0=(2,\ 2,\ 2)^\top$

Tanto per cominciare

Consideriamo il problema con soli vincoli di uguaglianza

$$\begin{array}{ll} \min & f(x) \\ c.v. & h(x) = 0 \end{array}$$

Funzione Lagrangiana: $L(x,\mu) = f(x) + \mu^\top h(x)$

con gradiente (vi ricordo):

$$\begin{aligned} \nabla_x L(x,\mu) &= \nabla f(x) + \nabla h(x)\mu \\ \nabla_\mu L(x,\mu) &= h(x) \end{aligned}$$

Percui, le C.N. di ottimo si possono riassumere in: $\nabla L(x,\mu) = 0$

Tanto per cominciare

Consideriamo il problema con soli vincoli di uguaglianza

$$\begin{array}{ll} \min & f(x) \\ c.v. & h(x) = 0 \end{array}$$

Funzione Lagrangiana: $L(x, \mu) = f(x) + \mu^\top h(x)$

con gradiente (vi ricordo):

$$\begin{array}{rcl} \nabla_x L(x, \mu) & = & \nabla f(x) + \nabla h(x)\mu \\ \nabla_\mu L(x, \mu) & = & h(x) \end{array}$$

Percui, le C.N. di ottimo si possono riassumere in: $\nabla L(x, \mu) = 0$

Tanto per cominciare

Consideriamo il problema con soli vincoli di uguaglianza

$$\begin{array}{ll} \min & f(x) \\ c.v. & h(x) = 0 \end{array}$$

Funzione Lagrangiana: $L(x, \mu) = f(x) + \mu^\top h(x)$

con gradiente (vi ricordo):

$$\begin{aligned} \nabla_x L(x, \mu) &= \nabla f(x) + \nabla h(x)\mu \\ \nabla_\mu L(x, \mu) &= h(x) \end{aligned}$$

Percui, le C.N. di ottimo si possono riassumere in: $\nabla L(x, \mu) = 0$

Tanto per cominciare

Consideriamo il problema con soli vincoli di uguaglianza

$$\begin{array}{ll} \min & f(x) \\ c.v. & h(x) = 0 \end{array}$$

Funzione Lagrangiana: $L(x,\mu) = f(x) + \mu^\top h(x)$

con gradiente (vi ricordo):

$$\begin{array}{rcl} \nabla_x L(x,\mu) & = & \nabla f(x) + \nabla h(x)\mu \\ \nabla_\mu L(x,\mu) & = & h(x) \end{array}$$

Percui, le C.N. di ottimo si possono riassumere in: $\nabla L(x,\mu) = 0$

Quindi ...

Punti che soddisfano le C.N. di ottimo sono punti staz. della funzione Lagrangiana $L(x, \mu)$!!

Idea: minimizzare $L(x, \mu)$ sullo spazio $\mathbb{R}^{n+p}$ per determinare $(\bar{x}, \bar{\mu})$ tale che $\nabla L(\bar{x}, \bar{\mu}) = \mathbf{0}$

Purtroppo: $L(x, \mu)$ è lineare rispetto alle variabili duali μ!!

Quindi ...

Punti che soddisfano le C.N. di ottimo sono punti staz. della funzione Lagrangiana $L(x, \mu)$!!

Idea: minimizzare $L(x, \mu)$ sullo spazio $\mathbb{R}^{n+p}$ per determinare $(\bar{x}, \bar{\mu})$ tale che $\nabla L(\bar{x}, \bar{\mu}) = \mathbf{0}$

Purtroppo: $L(x, \mu)$ è lineare rispetto alle variabili duali μ!!

Quindi ...

Punti che soddisfano le C.N. di ottimo sono punti staz. della funzione Lagrangiana $L(x, \mu)$!!

Idea: minimizzare $L(x, \mu)$ sullo spazio $\mathbb{R}^{n+p}$ per determinare $(\bar{x}, \bar{\mu})$ tale che $\nabla L(\bar{x}, \bar{\mu}) = \mathbf{0}$

Purtroppo: $L(x, \mu)$ è lineare rispetto alle variabili duali μ!!

Però ...

Siano: x^* un minimo locale (regolare) del problema e μ^* i moltiplicatori di KKT associati.

Supponiamo di conoscere i moltiplicatori μ^* ottimi e di voler determinare x^*.
Cosa succede se tentiamo di minimizzare la funzione $L(x, \mu^*)$ su $\mathbb{R}^n$?

Purtroppo: spesso si ottiene un punto $\bar{x} \neq x^*$ e tale che $h(\bar{x}) \neq \mathbf{0}$!!

Allora? bisogna modificare $L(x, \mu^*)$ "aggiungendo" un termine che consenta di penalizzare la violazione dei vincoli ...

$$\begin{aligned} L_a(x, \mu; \epsilon) &= L(x, \mu) + \frac{1}{\epsilon}\|h(x)\|^2 \\ &= f(x) + \mu^\top h(x) + \frac{1}{\epsilon}\|h(x)\|^2 \end{aligned}$$

Però ...

Siano: x^* un minimo locale (regolare) del problema e μ^* i moltiplicatori di KKT associati.

Supponiamo di conoscere i moltiplicatori μ^* ottimi e di voler determinare x^*.
Cosa succede se tentiamo di minimizzare la funzione $L(x, \mu^*)$ su $\mathbb{R}^n$?

Purtroppo: spesso si ottiene un punto $\bar{x} \neq x^*$ e tale che $h(\bar{x}) \neq \mathbf{0}$!!

Allora? bisogna modificare $L(x, \mu^*)$ "aggiungendo" un termine che consenta di penalizzare la violazione dei vincoli ...

$$\begin{aligned} L_a(x, \mu; \epsilon) &= L(x, \mu) + \frac{1}{\epsilon}\|h(x)\|^2 \\ &= f(x) + \mu^\top h(x) + \frac{1}{\epsilon}\|h(x)\|^2 \end{aligned}$$

Però ...

Siano: x^* un minimo locale (regolare) del problema e μ^* i moltiplicatori di KKT associati.

Supponiamo di conoscere i moltiplicatori μ^* ottimi e di voler determinare x^*.
Cosa succede se tentiamo di minimizzare la funzione $L(x, \mu^*)$ su $\mathbb{R}^n$?

Purtroppo: spesso si ottiene un punto $\bar{x} \neq x^*$ e tale che $h(\bar{x}) \neq \mathbf{0}$!!

Allora? bisogna modificare $L(x, \mu^*)$ "aggiungendo" un termine che consenta di penalizzare la violazione dei vincoli ...

$$\begin{aligned} L_a(x, \mu; \epsilon) &= L(x, \mu) + \frac{1}{\epsilon}\|h(x)\|^2 \\ &= f(x) + \mu^\top h(x) + \frac{1}{\epsilon}\|h(x)\|^2 \end{aligned}$$

Però ...

Siano: x^* un minimo locale (regolare) del problema e μ^* i moltiplicatori di KKT associati.

Supponiamo di conoscere i moltiplicatori μ^* ottimi e di voler determinare x^*.
Cosa succede se tentiamo di minimizzare la funzione $L(x, \mu^*)$ su $\mathbb{R}^n$?

Purtroppo: spesso si ottiene un punto $\bar{x} \neq x^*$ e tale che $h(\bar{x}) \neq \mathbf{0}$!!

Allora? bisogna modificare $L(x, \mu^*)$ "aggiungendo" un termine che consenta di penalizzare la violazione dei vincoli ...

$$\begin{aligned} L_a(x, \mu; \epsilon) &= L(x, \mu) + \frac{1}{\epsilon}\|h(x)\|^2 \\ &= f(x) + \mu^\top h(x) + \frac{1}{\epsilon}\|h(x)\|^2 \end{aligned}$$

Lagrangiano aumentato

Abbiamo "convessificato" $L(x,\mu)$ nell'intorno delle soluzione del problema originario (semplicemente) aggiungendo un termine di penalità, con $\epsilon > 0$ ("sufficientemente" piccolo)

$$L_a(x,\mu;\epsilon) = L(x,\mu) + \frac{1}{\epsilon}\|h(x)\|^2$$

con gradiente

$$\begin{aligned}\nabla_x L_a(x,\mu;\epsilon) &= \nabla_x L(x,\mu) + \frac{2}{\epsilon}\nabla h(x)h(x)\\ \nabla_\mu L_a(x,\mu;\epsilon) &= \nabla_\mu L(x,\mu) = h(x)\end{aligned}$$

Lagrangiano aumentato

Abbiamo "convessificato" $L(x,\mu)$ nell'intorno delle soluzione del problema originario (semplicemente) aggiungendo un termine di penalità, con $\epsilon > 0$ ("sufficientemente" piccolo)

$$L_a(x,\mu;\epsilon) = L(x,\mu) + \frac{1}{\epsilon}\|h(x)\|^2$$

con gradiente

$$\begin{aligned} \nabla_x L_a(x,\mu;\epsilon) &= \nabla_x L(x,\mu) + \frac{2}{\epsilon}\nabla h(x)h(x) \\ \nabla_\mu L_a(x,\mu;\epsilon) &= \nabla_\mu L(x,\mu) = h(x) \end{aligned}$$

Prima proprietà

Proposizione

Per ogni valore $\epsilon > 0$, ogni punto staz. di $L_a(x, \mu; \epsilon)$ è un punto di KKT del problema originario e viceversa

Dim.: Sia $(\bar{x}, \bar{\mu})$ un punto t.c. $\nabla L_a(\bar{x}, \bar{\mu}) = 0$. Allora, dalla espr. del gradiente risulta, banalmente: $h(\bar{x}) = 0$ e quindi $\nabla_x L(\bar{x}, \bar{\mu}) = 0$, ovvero $(\bar{x}, \bar{\mu})$ è un punto di KKT.

Supponiamo ora che $(\bar{x}, \bar{\mu})$ sia un punto di KKT e quindi, in particolare, $h(\bar{x}) = 0$ e $\nabla_x L(\bar{x}, \bar{\mu}) = 0$. Allora, dalla espr. del gradiente di L_a si ricava nuovamente $\nabla L_a(\bar{x}, \bar{\mu}) = 0$. $\square$

Prima proprietà

Proposizione

Per ogni valore $\epsilon > 0$, ogni punto staz. di $L_a(x, \mu; \epsilon)$ è un punto di KKT del problema originario e viceversa

Dim.: Sia $(\bar{x}, \bar{\mu})$ un punto t.c. $\nabla L_a(\bar{x}, \bar{\mu}) = 0$. Allora, dalla espr. del gradiente risulta, banalmente: $h(\bar{x}) = 0$ e quindi $\nabla_x L(\bar{x}, \bar{\mu}) = 0$, ovvero $(\bar{x}, \bar{\mu})$ è un punto di KKT.

Supponiamo ora che $(\bar{x}, \bar{\mu})$ sia un punto di KKT e quindi, in particolare, $h(\bar{x}) = 0$ e $\nabla_x L(\bar{x}, \bar{\mu}) = 0$. Allora, dalla espr. del gradiente di L_a si ricava nuovamente $\nabla L_a(\bar{x}, \bar{\mu}) = 0$. $\square$

Seconda proprietà

Indichiamo con $\mathcal{G}(\mathcal{C})$ l'insieme dei punti di minimo globale del problema

$$\begin{array}{l} \min\ f(x) \\ s.t.\ x \in \mathcal{F} \cap \mathcal{C} \end{array}$$

Supponiamo:

- che esista $\mathcal{C} \subseteq \mathbb{R}^n$ compatto tale che $\mathcal{G}(\mathcal{C}) \subseteq \mathring{\mathcal{C}}$
- $\bar{x} \in \mathcal{G}(\mathcal{C})$ e $\bar{\mu}$ moltiplicatori di KKT associati

Proposizione

Allora **per valori sufficiente piccoli** *di ϵ, $\bar{x}$ è un minimo globale non vincolato di $L_a(x, \bar{\mu}; \epsilon)$ su $\mathcal{C}$*

Dim. Per contraddizione, supponiamo che per ogni $k \in \mathbb{N}$ esiste un $\epsilon_k \leq 1/k$ ed un punto x_k di minimo globale di $L_a(x, \bar{\mu}; \epsilon_k)$ su $\mathcal{C}$ tale che

$$L_a(x_k, \bar{\mu}; \epsilon_k) < L_a(\bar{x}, \bar{\mu}; \epsilon_k) = f(\bar{x})$$

Seconda proprietà – segue

Quindi

$$f(x_k) + \bar{\mu}^\top h(x_k) + \frac{1}{\epsilon_k}\|h(x_k)\| < f(\bar{x}).$$

Ora $\{x_k\}$ è tutta contenuta in un compatto, quindi ammette punti di accumulazione. Esiste quindi certamente una sottosuccessione convergente, che rinominiamo $\{x_k\}$, tale che

$$\lim_{k\to\infty} x_k = \hat{x}.$$

Deve pertanto risultare

$$\lim_{k\to\infty} L_a(x_k, \bar{\mu}; \epsilon_k) \leq f(\bar{x})$$

ovvero $h(\hat{x}) = 0$ e $f(\hat{x}) \leq f(\bar{x})$, ovvero anche il punto $\hat{x}$ è minimo globale del problema su $\mathcal{C}$. Quindi $\hat{x} \in \overset{\circ}{\mathcal{C}}$. Ma allora, per k suff. elevato $x_k \in \overset{\circ}{\mathcal{C}}$. E allora $\nabla_x L_a(x_k, \bar{\mu}; \epsilon) = 0$. Di conseguenza, $(x_k, \bar{\mu})$ è una coppia di KKT per il problema, tale che percui

$$L_a(x_k, \bar{\mu}; \epsilon_k) = f(x_k) < f(\bar{x})$$

che contraddice il fatto che $\bar{x}$ è minimo globale del problema.

Terza proprietà

Sia $\bar{x}$ una sol. del problema vincolato originale tale che:

- $\bar{x}$ è un punto regolare per i vincoli
- la coppia di KKT $(\bar{x}, \bar{\mu})$ soddisfa le C.S. del II ordine

Proposizione

Per **valori sufficientemente piccoli di** $\epsilon > 0$, $\bar{x}$ *è un minimo locale stretto di* $L_a(x, \bar{\mu}; \epsilon)$

N.B. è necessario conoscere $\bar{\mu}$!

Terza proprietà

Sia $\bar{x}$ una sol. del problema vincolato originale tale che:

- $\bar{x}$ è un punto regolare per i vincoli
- la coppia di KKT $(\bar{x}, \bar{\mu})$ soddisfa le C.S. del II ordine

Proposizione

Per **valori sufficientemente piccoli di** $\epsilon > 0$, $\bar{x}$ *è un minimo locale stretto di* $L_a(x, \bar{\mu}; \epsilon)$

N.B. è necessario conoscere $\bar{\mu}$!

Terza proprietà

Sia $\bar{x}$ una sol. del problema vincolato originale tale che:

- $\bar{x}$ è un punto regolare per i vincoli
- la coppia di KKT $(\bar{x}, \bar{\mu})$ soddisfa le C.S. del II ordine

Proposizione

Per **valori sufficientemente piccoli di** $\epsilon > 0$, $\bar{x}$ *è un minimo locale stretto di* $L_a(x, \bar{\mu}; \epsilon)$

N.B. è necessario conoscere $\bar{\mu}$!

Esempio

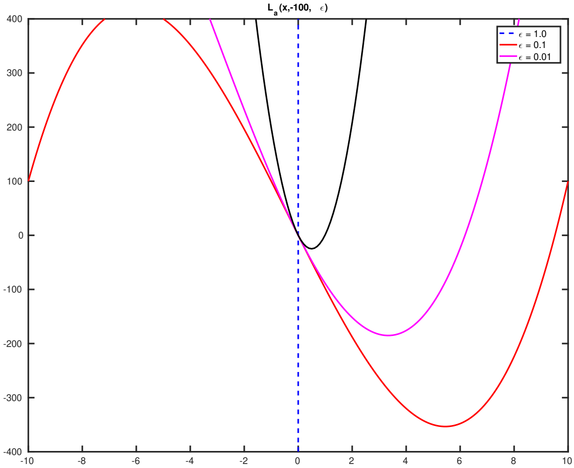
Consideriamo il problema

$$\begin{aligned} &\min \ x^3 \\ &s.t. \ x = 0 \end{aligned}$$

per cui risulta (banalmente) $x^* = 0$ e $\mu^* = 0$

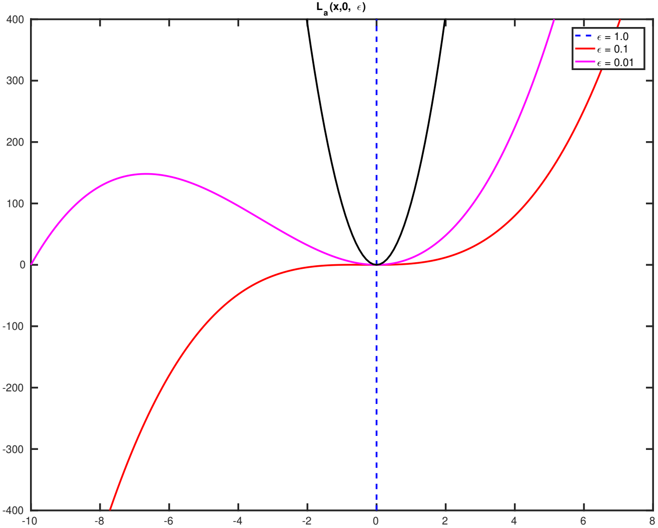
Esempio

Andamenti della funzione $L_a(x, -100; \epsilon)$ per $\epsilon = 1, 0.1, 0.01$



Esempio

Andamenti della funzione $L_a(x, 0; \epsilon)$ per $\epsilon = 1, 0.1, 0.01$



Esempio

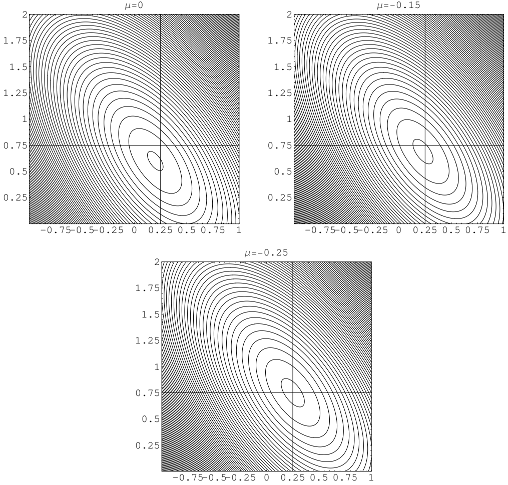
Consideriamo il problema

$$\begin{aligned} \min\ & \frac{1}{2}x^2 + \frac{1}{6}y^2 \\ s.t.\ & x + y = 1 \end{aligned}$$

Calcolare la soluzione del problema (x^*, y^*, μ^*)

Esempio

Curve di livello di $L_a(x, y, \bar{\mu}; \epsilon)$ per $\epsilon = 1$ a $\bar{\mu} = 0, -0.15, -0.25$



Proprietà

Proposizione

- *Per ogni valore* $\epsilon > 0$*, ogni punto staz. di* $L_a(x, \mu; \epsilon)$ *è un punto di KKT del problema originario e viceversa*
- *Sia* $(\bar{x}, \bar{\mu})$ *una sol. del problema originale. Per* **valori sufficientemente piccoli di** $\epsilon > 0$, $\bar{x}$ *è un minimo locale di* $L_a(x, \bar{\mu}; \epsilon)$

Metodo di soluzione

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\mu_\circ$, `maxit`

for $k = 0, 1, \dots,$`maxit`

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k; \epsilon_k)\| \le \tau_k$

if $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, STOP

endif

endfor

Return: miglior coppia trovata (x^*, μ^*)

Metodo di soluzione

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\mu_\circ$, `maxit`

for $k = 0, 1, \dots,$`maxit`

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k; \epsilon_k)\| \le \tau_k$

if $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, STOP

endif

$\epsilon_{k+1} = \epsilon_k / \beta$

Calcola μ_{k+1}

endfor

Return: miglior coppia trovata (x^*, μ^*)

Metodo di soluzione

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\mu_\circ$, `maxit`

for $k = 0, 1, \ldots,$`maxit`

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k; \epsilon_k)\| \leq \tau_k$

if $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, STOP

endif

$\epsilon_{k+1} = \epsilon_k / \beta$

Calcola μ_{k+1}

endfor

Return: miglior coppia trovata (x^*, μ^*)

Metodo di soluzione

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\mu_\circ$, `maxit`

for $k = 0, 1, \dots,$`maxit`

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k; \epsilon_k)\| \leq \tau_k$

if $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, STOP

endif

$\epsilon_{k+1} = \epsilon_k / \beta$

Calcola μ_{k+1}

endfor

Return: miglior coppia trovata (x^*, μ^*)

Aggiornamento del moltiplicatore

Come si passa da μ_k a μ_{k+1}?

Supponiamo di aver calcolato x_k e supponiamo anche che x_k sia stazionario per la funzione $L_a(x, \mu_k; \epsilon_k)$ cioè

$$\nabla_x L_a(x_k, \mu_k; \epsilon_k) = 0$$

Ovvero

$$\nabla_x L_a(x_k, \mu_k; \epsilon_k) = \nabla f(x_k) + \nabla h(x_k) \left(\mu_k + \frac{2}{\epsilon_k} h(x_k) \right) = 0$$

Se definiamo $\mu_{k+1} = \mu_k + \frac{2}{\epsilon_k} h(x_k)$ risulterà

$$\nabla f(x_k) + \nabla h(x_k) \mu_{k+1} = 0$$

per cui x_k soddisfa la condizione di stazionarietà a meno della condizione $h(x_k) = 0$

Aggiornamento del moltiplicatore

Come si passa da μ_k a μ_{k+1}?

Supponiamo di aver calcolato x_k e supponiamo anche che x_k sia stazionario per la funzione $L_a(x, \mu_k; \epsilon_k)$ cioè

$$\nabla_x L_a(x_k, \mu_k; \epsilon_k) = 0$$

Ovvero

$$\nabla_x L_a(x_k, \mu_k; \epsilon_k) = \nabla f(x_k) + \nabla h(x_k) \left(\mu_k + \frac{2}{\epsilon_k} h(x_k) \right) = 0$$

Se definiamo $\mu_{k+1} = \mu_k + \frac{2}{\epsilon_k} h(x_k)$ risulterà

$$\nabla f(x_k) + \nabla h(x_k) \mu_{k+1} = 0$$

per cui x_k soddisfa la condizione di stazionarietà a meno della condizione $h(x_k) = 0$

Aggiornamento del moltiplicatore

Come si passa da μ_k a μ_{k+1}?

Supponiamo di aver calcolato x_k e supponiamo anche che x_k sia stazionario per la funzione $L_a(x, \mu_k; \epsilon_k)$ cioè

$$\nabla_x L_a(x_k, \mu_k; \epsilon_k) = 0$$

Ovvero

$$\nabla_x L_a(x_k, \mu_k; \epsilon_k) = \nabla f(x_k) + \nabla h(x_k) \left(\mu_k + \frac{2}{\epsilon_k} h(x_k) \right) = 0$$

Se definiamo $\mu_{k+1} = \mu_k + \frac{2}{\epsilon_k} h(x_k)$ risulterà

$$\nabla f(x_k) + \nabla h(x_k) \mu_{k+1} = 0$$

per cui x_k soddisfa la condizione di stazionarietà a meno della condizione $h(x_k) = 0$

Metodo di soluzione - Metodo dei Moltiplicatori

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\mu_\circ$, `maxit`

for $k = 0, 1, \dots,$`maxit`

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k; \epsilon_k)\| \leq \tau_k$

if $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, STOP

endif

$\epsilon_{k+1} = \epsilon_k / \beta$

$$\boxed{\mu_{k+1} = \mu_k + \frac{2h(x_k)}{\epsilon_k}}$$

endfor

Return: miglior coppia trovata (x^*, μ^*)