

Ottimizzazione dei Sistemi Complessi

G. Liuzzi¹

Venerdì 6 Aprile 2018

¹Istituto di Analisi dei Sistemi ed Informatica IASI - CNR

Condizione Necessaria del II ordine

Supponiamo:

- x^* minimo locale
- f, g, h **due volte** continuamente differenziabili
- x^*, λ^*, μ^* soddisfano KKT

Teorema (C.N. II ordine)

Se $d^\top \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d \geq 0$, per ogni $d \in Y^*$,

$$Y^* = \left\{ d \in \mathbb{R}^n : \begin{aligned} &\nabla h_j(x^*)^\top d = 0, j = 1, \dots, p \\ &\nabla g_i(x^*)^\top d = 0, i \in I_0(x^*), \lambda_i^* > 0, \\ &\nabla g_i(x^*)^\top d \geq 0, i \in I_0(x^*), \lambda_i^* = 0 \end{aligned} \right\}$$

Condizioni di ottimalità

Supponiamo:

- f, g differenziabili e **convesse**
- h **affini**, e.g. $h_j(x) = a^\top x - b = 0$

Teorema (Condizione Sufficiente)

Se $x^* \in \mathcal{F}$ e esistono dei moltiplicatori $\lambda_1^*, \dots, \lambda_m^* \geq 0$, μ_1^*, \dots, μ_p^* tali che:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla h_j(x^*) = 0,$$

$$\lambda_i^* g_i(x^*) = 0, \quad \text{per ogni } i = 1, \dots, m,$$

allora x^* è un minimo locale (globale) di (P_0)

Funzione Lagrangiana

Per il problema (P_0) , definiamo la funzione Lagrangiana

$$\begin{aligned}L(x, \lambda, \mu) &= f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x) \\ &= f(x) + \lambda^\top g(x) + \mu^\top h(x)\end{aligned}$$

Condizioni di ottimalità

Ci sono condizioni sufficienti di ottimo nel caso **non convesso**?

Supponiamo:

- f, g, h **due volte** continuamente differenziabili
- x^*, λ^*, μ^* soddisfano KKT

Teorema (Condizione Sufficiente)

Se $d^\top \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d > 0$, per ogni $d \in Y^*$, $d \neq 0$

$$Y^* = \left\{ d \in \mathbb{R}^n : \begin{aligned} \nabla h_j(x^*)^\top d &= 0, j = 1, \dots, p \\ \nabla g_i(x^*)^\top d &= 0, i \in I_0(x^*), \lambda_i^* > 0, \\ \nabla g_i(x^*)^\top d &\geq 0, i \in I_0(x^*), \lambda_i^* = 0 \end{aligned} \right\}$$

allora x^ è un minimo locale stretto di (P_0)*

Esempio

$$\begin{array}{ll} \min_x & x_1 + x_2 \\ \text{c.v} & x_1^2 + x_2^2 - 2 \leq 0 \end{array}$$

$$\begin{array}{ll} \min_x & -0.1(x_1 - 4)^2 + x_2^2 \\ \text{c.v} & x_1^2 + x_2^2 - 1 \geq 0 \end{array}$$

Problemi vincolati

Problema vincolato nonlineare generale

$$\begin{array}{ll} \min_x & f(x) \\ \text{c.v} & h(x) = 0 \\ & g(x) \leq 0 \end{array} \quad (P_0)$$

con \mathcal{F} insieme ammissibile di (P_0)

Definizione

Teoricamente è possibile definire

$$q_{\infty}(x) = \begin{cases} 0 & \text{se } x \in \mathcal{F} \\ +\infty & \text{se } x \notin \mathcal{F} \end{cases}$$

e quindi

$$P_{\infty}(x) = f(x) + q_{\infty}(x)$$

È evidente che:

- risolvere il problema $\min P_{\infty}(x)$ fornisce soluzioni del problema (P_0) .
- minimizzare $P_{\infty}(x)$ è (o può essere) molto “complicato”

Definizione

Anziché $q_\infty(x)$, definiamo

$$q(x) = \begin{cases} 0 & \text{se } x \in \mathcal{F} \\ > 0 & \text{se } x \notin \mathcal{F} \end{cases}$$

e quindi

$$P_\epsilon(x) = f(x) + \frac{1}{\epsilon}q(x)$$

N.B.:

- se $f(x)$ e $q(x)$ sono cont. differenziabili, allora anche $P_\epsilon(x)$ lo è
- per ogni $x \in \mathbb{R}^n$, $\lim_{\epsilon \rightarrow 0} P_\epsilon(x) = P_\infty(x)$

Espressione di $q(x)$

Il termine $q(x)$ deve “penalizzare con continuità” l’eventuale inammissibilità di x

P.es., con riferimento a (P_0) ,

$$q(x) = \sum_{i=1}^p [h_i(x)]^2 + \sum_{i=1}^m [\max\{0, g_i(x)\}]^2$$

- per nessun valore di $\epsilon > 0$, $\min P_\epsilon(x)$ è equivalente a (P_0)
- per ogni valore di $\epsilon > 0$, una min. non vincolata di $P_\epsilon(x)$ produce (in genere) un punto $x^* \notin \mathcal{F}$, cioè $q(x^*) > 0$

Algoritmo di soluzione

Algoritmo SEQPEN

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

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

Trova una soluzione ($x(\epsilon_k)$) del problema

$$\min_{x \in \mathbb{R}^n} P_{\epsilon_k}(x). \quad (2)$$

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

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

endfor

OUTPUT: una sol. approssimata $x^* = x(\epsilon_k)$

Proprietà di monotonicità

Teorema

Siano soddisfatte le ipotesi seguenti:

- *esiste $x^* \in \mathcal{F}$ soluzione globale del problema vincolato;*
- *per ogni k , la funzione $P_{\epsilon_k}(x)$ ammette minimo globale x^k .*

Allora,

- 1 $P_{\epsilon_k}(x_k) \leq f(x^*);$
- 2 $q(x^k) \geq q(x^{k+1});$
- 3 $f(x^k) \leq f(x^{k+1});$
- 4 $P_{\epsilon_k}(x^k) \leq P_{\epsilon_{k+1}}(x^{k+1}).$

Proprietà di monotonicità – dim.

Dim.

$$P_{\epsilon_k}(x^k) = \min_{x \in \mathbb{R}^n} f(x) + q(x)/\epsilon_k \leq \min_{x \in \mathcal{F}} f(x) + q(x)/\epsilon_k = \min_{x \in \mathcal{F}} f(x) = f(x^*).$$

che prova la (1).

$$\begin{aligned} P_{\epsilon_k}(x^k) &\leq P_{\epsilon_k}(x^{k+1}) \\ P_{\epsilon_{k+1}}(x^{k+1}) &\leq P_{\epsilon_{k+1}}(x^k) \end{aligned}$$

Ovvero

$$\begin{aligned} f(x^k) + q(x^k)/\epsilon_k &\leq f(x^{k+1}) + q(x^{k+1})/\epsilon_k \\ f(x^{k+1}) + q(x^{k+1})/\epsilon_{k+1} &\leq f(x^k) + q(x^k)/\epsilon_{k+1} \end{aligned}$$

$$\begin{aligned} f(x^k) + q(x^k)/\epsilon_k &\leq f(x^{k+1}) + q(x^{k+1})/\epsilon_k \\ -f(x^k) - q(x^k)/\epsilon_{k+1} &\leq -f(x^{k+1}) - q(x^{k+1})/\epsilon_{k+1} \end{aligned}$$

Proprietà di monotonicità – dim. (segue)

da cui segue

$$q(x^k)(1/\epsilon_k - 1/\epsilon_{k+1}) \leq q(x^{k+1})(1/\epsilon_k - 1/\epsilon_{k+1})$$

tenendo conto del fatto che $\epsilon_{k+1} < \epsilon_k$

$$q(x^{k+1}) \leq q(x^k)$$

che prova la (2). Ora, ricordando che

$$f(x^k) - f(x^{k+1}) \leq (q(x^{k+1}) - q(x^k))/\epsilon_k \leq 0$$

che prova la (3). La (4) segue da

$$f(x^k) + q(x^k)/\epsilon_k \leq f(x^{k+1}) + q(x^{k+1})/\epsilon_k \leq f(x^{k+1}) + q(x^{k+1})/\epsilon_{k+1}$$

ricordando che $\epsilon_k > \epsilon_{k+1}$.



Proprietà di convergenza

Teorema

Siano soddisfatte le ipotesi seguenti:

- *esiste $x^* \in \mathcal{F}$ soluzione globale del problema vincolato;*
- *per ogni k , la funzione $P_{\epsilon_k}(x)$ ammette minimo globale x^k ;*
- *tutti i punti x^k rimangono in un insieme compatto D .*

Allora,

- 1 $\lim_{k \rightarrow \infty} q(x^k) = 0;$
- 2 $\lim_{k \rightarrow \infty} f(x^k) = f(x^*);$
- 3 $\lim_{k \rightarrow \infty} P_{\epsilon_k}(x^k) = f(x^*);$
- 4 *ogni punto di accumulazione di $\{x^k\}$ è soluzione globale del problema vincolato;*
- 5 $\lim_{k \rightarrow \infty} q(x^k)\epsilon_k = 0;$

Proprietà di convergenza – dim.

Dim. Sappiamo che

$$f(x^*) \geq P_{\epsilon_k}(x^k) = f(x^k) + q(x^k)/\epsilon_k \geq f(x^0) + q(x^k)/\epsilon_k$$

ovvero

$$f(x^*) - f(x^0) \geq q(x^k)/\epsilon_k.$$

Quando $k \rightarrow \infty$, deve risultare

$$\lim_{k \rightarrow \infty} q(x^k) = 0$$

(altrimenti si avrebbe un assurdo) il che prova la (1).

Sia ora \bar{x} un qualunque punto di accumulazione di $\{x^k\} \subset D$ compatto. Risulta intanto che $q(\bar{x}) = 0$, cioè che $\bar{x} \in \mathcal{F}$.

Proprietà di convergenza – dim. (segue)

$$f(x^*) \geq P_{\epsilon_k}(x^k) \geq f(x^k)$$

Quindi, le successioni $\{P_{\epsilon_k}(x^k)\}$ e $\{f(x^k)\}$ sono monotone non decrescenti e limitate superiormente, quindi ammettono limite.

$$f(x^*) \geq \lim_{k \rightarrow \infty} P_{\epsilon_k}(x^k) = \lim_{k \rightarrow \infty} f(x^k) + q(x^k)/\epsilon_k$$

$$f(x^*) \geq \lim_{k \rightarrow \infty} P_{\epsilon_k}(x^k) = f(\bar{x}) + \lim_{k \rightarrow \infty} q(x^k)/\epsilon_k \geq f(\bar{x}) \geq f(x^*).$$

Da qui segue che valgono (2) e (3). Inoltre, che $f(\bar{x}) = f(x^*)$, cioè che \bar{x} è soluzione globale del problema vincolato e inoltre che vale la (5). □

Esempio 1

Consideriamo il seguente problema vincolato (banale)

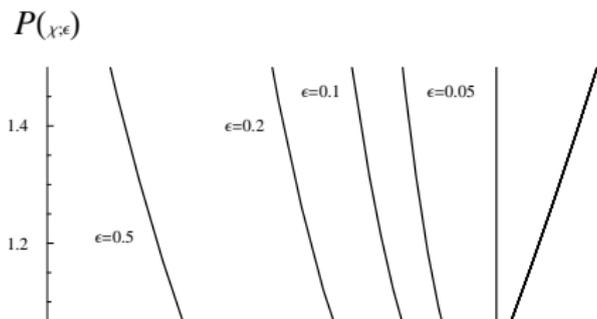
$$\begin{aligned} \min \quad & x^2 \\ \text{s.t.} \quad & x \geq 1 \end{aligned}$$

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

La funzione di penalità esterna è:

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

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



Esempio 2

Consideriamo il seguente problema vincolato (banale)

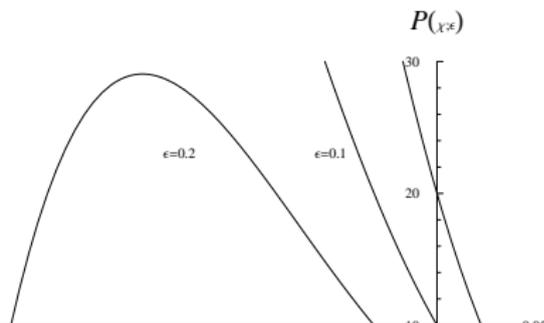
$$\begin{aligned} \min \quad & -x^4 \\ \text{s.t.} \quad & x = 1 \end{aligned}$$

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

La funzione di penalità esterna è:

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

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 $k = 0, 1, \dots, \text{maxit}$

Trova una soluzione globale (x^k) del problema

$$\min_{x \in \mathbb{R}^n} P_{\epsilon_k}(x). \quad (3)$$

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.

$$\|\nabla P_\epsilon(\hat{x})\| \leq \text{tol.}$$

Algoritmo SEQPEN modificato (1)

Algoritmo SEQPEN_{mod}INPUT: maxit, $\rho > 0$, $k = 0$, $\epsilon_0 > 0$ **for** $k = 0, 1, \dots, \text{maxit}$ 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 SEQPEN_{mod}

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

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

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.

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