Giogio Giorgi
Cesare Zuccotti

Osservazioni sulle condizioni di ottimalità del secondo ordine in programmazione matematica

Quaderno di ricerca numero X
Anno accademico 2007–2008

ARACNE

www.aracneeditrice.it
info@aracneeditrice.it

00173 Roma
via Raffaele Garofalo, 133 A/B
06 93781065
telefax 06 72678427

ISBN 978-88–548–1785-2

I edizione: maggio 2008

Finito di stampare nel mese di maggio del 2008
dalla tipografia « Braille Gamma S.r.l. » di Santa Rufina di Cittaducale (RI)
per conto della « Aracne editrice S.r.l. » di Roma
Printed in Italy

INDICE

1. Introduzione p. 1

2. Condizioni necessarie di ottimalità del secondo ordine p. 7

3. Condizioni sufficienti di ottimalità del secondo ordine p. 18

4. Condizioni sufficienti del secondo ordine per un minimo locale isolato p. 27

1. Introduzione *

Le condizioni di ottimalità (necessarie e sufficienti) del secondo ordine in problemi di programmazione matematica sono oramai un argomento "classico". Tuttavia, nonostante i numerosi, e a volte molto sofisticati, lavori sull'argomento (alcuni in termini di approssimazioni infinitesimali e coniche del secondo ordine, quali, ad esempio, l'articolo di Cambini, Martein e Vlach (1999) e quelli di Kawasaki (1988), Bonnans, Cominetti e Shapiro (1999), Cominetti (1990), Penot (1998)) permangono ancora questioni che è bene evidenziare e sottolineare, nonchè alcuni errori, anche in libri che sono diventati di riferimento per studi di programmazione matematica, di metodi matematici per l'economia e per discipline analoghe.

Scopo del presente lavoro è di cercare di presentare in modo didatticamente efficace sia le condizioni di ottimalità necessarie del secondo ordine che quelle sufficienti, corredando l'esposizione con opportune osservazioni. Si presenteranno infine delle condizioni del secondo ordine sufficienti per l'ottenimento di punti di minimo isolati per un problema di programmazione non lineare.

Consideriamo dunque un problema standard di programmazione non lineare del tipo

$$(\mathcal{P}) \qquad \min_{x \in \mathbb{R}^n} f(x)$$

ove il vettore x appartiene all'*insieme ammissibile*

$$\mathbb{K} = \left\{ \begin{array}{ccl} x \in \mathbb{R}^n & | & g_i(x) \leq 0, \quad i = 1, 2, ..., m, \\ & & h_j(x) = 0, \quad j = 1, 2, ..., p \end{array} \right\}.$$

* Giorgio Giorgi, Professore Ordinario di Matematica generale e di Matematica per l'economia e la finanza (secs-S/06). Università degli Studi di Pavia, Dipartimento di Ricerche Aziendali "Riccardo Argenziano", Via S. Felice, 7 - 27100 Pavia; tel. 0382986238 – e-mail: ggiorgi@eco.unipv.it

* Cesare Zuccotti, Ricercatore di Teoria delle decisioni e di Applicazioni elettroniche per la gestione e l'ottimizzazione (secs-S/06). Università degli Studi di Pavia, Dipartimento di Ricerche Aziendali "Riccardo Argenziano", Via S. Felice, 7 - 27100 Pavia; tel. 0382986257 – e-mail: czuccotti@eco.unipv.it

Le funzioni $f,\ \{g_i\}\ \{h_j\}$ sono definite e due volte parzialmente derivabili con continuità (cioè di classe C^2) in un comune sottoinsieme aperto $\mathbb{X} \subseteq \mathbb{R}^n$.

Il gradiente e la matrice Hessiana della funzione $\varphi :\ \mathbb{R}^n \to \mathbb{R}$, valutati nel generico punto x, sono denotati, rispettivamente, con $\bigtriangledown\varphi(x)$ e $\bigtriangledown^2\varphi(x)$. Le notazioni $\bigtriangledown_x$ e $\bigtriangledown_x^2$ indicano che le operazioni di derivazione sono state fatte rispetto a x.

Un punto $x^0 \in \mathbb{K}$ è di *minimo locale* per $(\mathcal{P})$ se

$$f\left(x^0\right) \leq f(x), \quad \forall x \in \mathbb{K} \cap \mathbb{N}\left(x^0\right),$$

ove $\mathbb{N}\left(x^0\right)$ è un intorno di x^0.

Un punto $x^0 \in \mathbb{K}$ è di minimo locale *stretto* (o *forte*) se

$$f\left(x^0\right) < f(x), \quad \forall x \in \mathbb{K} \cap \mathbb{N}\left(x^0\right), \quad x \neq x^0.$$

Un punto $x^0 \in \mathbb{K}$ è di minimo locale *stretto di ordine* k, con k intero positivo, se esiste un numero $m > 0$ tale che

$$f(x) \geq f\left(x^0\right) + m(\parallel x - x^0 \parallel)^k, \quad \forall x \in \mathbb{K} \cap \mathbb{N}\left(x^0\right).$$

Infine un punto $x^0 \in \mathbb{K}$ è di minimo locale *isolato* per $(\mathcal{P})$ o punto di minimo *localmente unico* per $(\mathcal{P})$ se è il solo punto di minimo locale per $(\mathcal{P})$ in un intorno $\mathbb{N}\left(x^0\right)$.

Un punto di minimo locale isolato è anche stretto ma il viceversa non vale.

Con riferimento a $(\mathcal{P})$, dato $x^0 \in \mathbb{K}$, chiameremo *insieme degli indici dei vincoli attivi* o *effettivi* in x^0 l'insieme

$$\mathbb{I}\left(x^0\right) = \left\{i \mid g_i\left(x^0\right) = 0\right\}.$$

La funzione *Lagrangiana generalizzata* (o Lagrangiana aumentata o funzione di Fritz John - Lagrange) è definita da

$$\mathbb{L}^g(x,\ u_0,\ u,\ \omega) = u_0\, f(x) + u\, g(x) + \omega\, h(x),$$

ove $u_0,\ u = [u_1,\ u_2,\ ...,\ u_m]\,,\ \ \omega = [\omega_1,\ \omega_2,\ ...,\ \omega_p]$ sono i *moltiplicatori di Fritz John.*

La funzione *Lagrangiana* (o funzione di Lagrange - Kuhn-Tucker) è definita da

$$\mathbb{L}(x,\ u,\ \omega) = f(x) + u\, g(x) + \omega\, h(x),$$

ove $(u,\ \omega)$ sono i *moltiplicatori di Kuhn-Tucker.*

Per ogni $x \in \mathbb{K}$ l'insieme dei moltiplicatori generalizzati di Fritz John - Lagrange è definito da

$$\Lambda^g(x) =$$
$$\left\{ \begin{array}{lll} (u_0, u, \omega) \neq (0,0,0) & \mid & \nabla_x \mathbb{L}^g(x, u_0, u, \omega) = 0,\ u_0 \geq 0, \\ & & \quad u_i \geq 0,\ u_i\, g_i(x) = 0,\ \ i = 1, 2, ..., m \end{array} \right\}.$$

L'insieme dei moltiplicatori di Kuhn-Tucker in $x \in \mathbb{K}$ è definito da

$$\Lambda(x) = \left\{ (u_0, u, \omega) \in \Lambda^g(x) \mid u_0 > 0 \right\}.$$

Di conseguenza, essendo sempre possibile assumere $u_0 = 1$, sarà

$$\Lambda(x) = \left\{ \begin{array}{lll} (1, u, \omega) & \mid & \nabla_x \mathbb{L}(x, u, \omega) = 0, \\ & & u_i \geq 0,\ \ u_i\, g_i(x) = 0,\ \ i = 1, 2, ..., m \end{array} \right\}.$$

E' noto che se $x^0 \in \mathbb{K}$ è punto di minimo locale per $(\mathcal{P})$, allora è $\Lambda^g\left(x^0\right) \neq \varnothing$ (Teorema di Fritz John; si veda, ad esempio, Bazaraa e Shetty (1976), Bazaraa, Sherali e Shetty (1993, 2006), Giorgi, Guerraggio e Thierfelder (2004), Mangasarian (1969)), ma in generale $\Lambda\left(x^0\right) \neq \varnothing$ solo se vale in x^0 una qualche *condizione di qualificazione dei vincoli* (si veda, ad esempio, Bazaraa e Shetty (1976), Giorgi, Guerraggio e Thierfelder (2004), Mangasarian (1969), Peterson (1973)).

Se $\Lambda^g\left(x^0\right) \neq \varnothing$, allora x^0 è *punto stazionario di Fritz John* per $(\mathcal{P})$ e similmente se $\Lambda\left(x^0\right) \neq \varnothing$, allora x^0 è *punto stazionario di Kuhn-Tucker* per $(\mathcal{P})$.

Una condizione di qualificazione dei vincoli molto utilizzata per $(\mathcal{P})$ è la seguente, detta "condizione di Mangasarian - Fromovitz" (Mangasarian e Fromovitz (1967)).

Condizione di Mangasarian - Fromovitz (MF)

In $x \in \mathbb{K}$ sussistono le due proprietà:

1) i vettori $\bigtriangledown h_j(x), \ j = 1, 2, ..., p$, sono linearmente indipendenti;

2) il sistema

$$\begin{aligned} d^T \bigtriangledown g_i(x) < 0, \quad i \in \mathbb{I}(x), \\ d^T \bigtriangledown h_j(x), \quad j = 1, 2, ..., p, \end{aligned}$$

ammette soluzione $d \in \mathbb{R}^n$.

E' possibile esprimere la condizione (MF) nel seguente equivalente modo (si veda, ad esempio, Hestenes (1975) pag. 241):

non esiste nessuna coppia $(\lambda, \ \mu) \neq (0, \ 0)$, tale che

$$\sum_{i \in \mathbb{I}(x)} \lambda_i \bigtriangledown g_i(x) + \sum_{j=1}^{p} \mu_j \bigtriangledown h_j(x) = 0, \quad \lambda_i \geq 0, \quad i \in \mathbb{I}(x).$$

In altre parole i gradienti dei vincoli attivi in x e dei vincoli bilaterali in x sono *positivamente linearmente indipendenti.*

Un'altra condizione di qualificazione dei vincoli che sarà utilizzata in seguito è quella di *indipendenza lineare.*

Condizione di indipendenza lineare (IL)

In $x \in \mathbb{K}$ i vettori

$$\bigtriangledown g_i(x), \quad i \in \mathbb{I}(x), \quad \bigtriangledown h_j(x), \quad j = 1, 2, ..., p,$$

sono linearmente indipendenti.

Vale la seguente importante proposizione (si veda Bonnans e Shapiro (2000)).

Teorema 1.1

Sia x^0 soluzione locale per $(\mathcal{P})$. Allora l'insieme $\Lambda^g(x^0)$ è non vuoto e le seguenti condizioni sono equivalenti:

1) vale in x^0 la condizione (MF);

2) l'insieme $\Lambda^g(x^0)$, ove è $u_0 = 0$, è vuoto;

3) l'insieme $\Lambda(x^0)$ è non vuoto, convesso e compatto (più precisamente, un poliedro convesso e compatto).

E' poi facile dimostrare il seguente risultato.

Teorema 1.2

Sia x^0 soluzione locale di $(\mathcal{P})$ e valga in x^0 la condizione (IL). Allora l'insieme $\Lambda(x^0)$ è un singleton.

La condizione (IL) è dunque condizione sufficiente per ottenere l'*unicità* dei moltiplicatori di Kuhn-Tucker per $(\mathcal{P})$. La condizione necessaria e sufficiente per ottenere la suddetta unicità è stata ottenuta da Kyparisis (1985) ed è detta "condizione di Mangasarian - Fromovitz *stretta*".

Condizione di Mangasarian - Fromovitz stretta (MFS).

In $x \in \mathbb{K}$ i vettori

$$\nabla g_i(x), \quad x \in \mathbb{I}(x), \quad \text{e} \quad \nabla h_j(x), \quad j = 1, 2, ..., p,$$

sono linearmente indipendenti ed esiste $d \in \mathbb{R}^n$ tale che

$$\begin{array}{rcl} d^T \nabla g_i(x) & < & 0, \quad i \in \mathbb{I}(x) \backslash \mathbb{I}_+(x, u); \\ d^T \nabla g_i(x) & = & 0, \quad i \in \mathbb{I}_+(x, u); \\ d^T \nabla h_j(x) & = & 0, \quad j = 1, 2, ..., p, \end{array}$$

ove $\mathbb{I}_+(x,\ u) = \{i \mid i \in \mathbb{I}(x), \ \ u_i > 0\}$ è *l'insieme degli indici fortemente attivi* in $x \in \mathbb{K}$, indice che dipende dal vettore dei moltiplicatori u nelle condizioni di Kuhn-Tucker

$$\nabla_x \mathbb{L}(x,\ u,\ w) = 0, \quad u\, g(x) = 0, \quad u \geq 0.$$

Introduciamo ora il seguente cono:

sia $\overline{x} \in \mathbb{K}$; viene detto *cono critico* o *cono delle direzioni critiche* associato a $\overline{x}$ l'insieme

$$\mathbb{C}(\overline{x}) = \left\{ \begin{array}{lll} d \in \mathbb{R}^n & | & d^T \bigtriangledown f(\overline{x}) \leq 0; \\ & & d^T \bigtriangledown g_i(\overline{x}) \leq 0, \quad i \in \mathbb{I}(\overline{x}); \\ & & d^T \bigtriangledown h_j(\overline{x}) = 0, \quad j = 1, 2, ..., p \end{array} \right\}.$$

Con riferimento a $\mathbb{C}(\overline{x})$, introduciamo anche il seguente sotto-cono del medesimo:

$$\mathbb{C}_1(\overline{x}) = \left\{ \begin{array}{lll} d \in \mathbb{R}^n & | & d^T \bigtriangledown f(\overline{x}) = 0; \\ & & d^T \bigtriangledown g_i(\overline{x}) \leq 0, \quad i \in \mathbb{I}(\overline{x}); \\ & & d^T \bigtriangledown h_j(\overline{x}) = 0, \quad j = 1, 2, ..., p \end{array} \right\}.$$

Dimostriamo ora il seguente risultato che può essere considerato un raffinamento delle condizioni necessarie di ottimalità di Fritz John.

Teorema 1.3

Sia x^0 punto di minimo locale per $(\mathcal{P})$. Allora esistono moltiplicatori $u_0 \geq 0, \quad u_i \geq 0, \quad i \in \mathbb{I}\left(x^0\right), \quad \omega_j \in \mathbb{R}, \quad j = 1, 2, ..., p,$ *non tutti nulli*, tali che

$$u_0 \bigtriangledown f\left(x^0\right) + \sum_{i \in \mathbb{I}(x^0)} u_i \bigtriangledown g_i\left(x^0\right) + \sum_{j=1}^{p} \omega_j \bigtriangledown h_j\left(x^0\right) = 0. \qquad (1.1)$$

Inoltre $u_0 > 0$ se e solo se $\mathbb{C}\left(x^0\right) = \mathbb{C}_1\left(x^0\right).$

Dimostrazione

La prima parte del teorema è notissima e per la relativa dimostrazione rimandiamo, ad esempio, a Mangasarian (1969).
Per la seconda parte, se è $\mathbb{C}\left(x^0\right) \neq \mathbb{C}_1\left(x^0\right),$

$$\begin{array}{lll} \exists \overline{d} \in \mathbb{C}\left(x^0\right) & | & \overline{d^T} \bigtriangledown f\left(x^0\right) < 0; \\ & & \overline{d^T} \bigtriangledown g_i\left(x^0\right) \leq 0, \quad i \in \mathbb{I}\left(x^0\right); \\ & & \overline{d^T} \bigtriangledown h_j\left(x^0\right) = 0, \quad j = 1, 2, ..., p. \end{array}$$

Moltiplicando la (1.1) per $\overline{d}$ otteniamo che deve essere $u_0 = 0$. D'altro canto se è $\mathbb{C}(x^0) = \mathbb{C}_1(x^0)$, concludiamo che il sistema

$$\begin{array}{l} d^T \bigtriangledown f(x^0) < 0; \\ d^T \bigtriangledown g_i(x^0) \leq 0, \quad i \in \mathbb{I}(x^0); \\ d^T \bigtriangledown h_j(x^0) = 0, \quad j = 1, 2, ..., p, \end{array}$$

non ha soluzioni in $d \in \mathbb{R}^n$. Dal noto teorema dell'alternativa di Farkas (cfr. Mangasarian (1969)), deduciamo che esistono moltiplicatori $u_i \geq 0, \quad i \in \mathbb{I}(x^0), \quad \omega_j \in \mathbb{R}, \quad j = 1, 2, ..., p$, tali che

$$\bigtriangledown f\left(x^0\right) + \sum_{i \in \mathbb{I}(x^0)} u_i \bigtriangledown g_i\left(x^0\right) + \sum_{j=1}^{p} \omega_j \bigtriangledown h_j\left(x^0\right) = 0. \qquad \square$$

Dunque, nel caso in cui l'insieme $\Lambda(x^0)$ sia non vuoto, possiamo sostituire in $\mathbb{C}(x^0)$ alla disuguaglianza $d \bigtriangledown f(x^0) \leq 0$, l'uguaglianza $d^T \bigtriangledown f(x^0) = 0$. Faremo in seguito vedere che per ogni $u \in \Lambda(x^0)$ il cono $\mathbb{C}_1(x^0)$ può essere scritto anche in un'altra forma che chiameremo $\mathbb{Z}_1(x^0)$. Non è poi difficile far vedere che, sempre per ogni $u \in \Lambda(x^0)$, il cono $\mathbb{C}_1(x^0)$ può anche essere scritto equivalentemente nella forma

$$\mathbb{C}_1(x^0) = \left\{ \begin{array}{ll} d \in \mathbb{R}^n \mid & d^T \bigtriangledown g_i(x^0) \leq 0, \quad i \in \mathbb{I}(x^0); \\ & d^T \bigtriangledown h_j(x^0) = 0, \quad j = 1, 2, ..., p; \\ & d^T \displaystyle\sum_{i \in \mathbb{I}(x^0)} u_i \bigtriangledown g_i(x^0) = 0 \end{array} \right\}.$$

2. Condizioni necessarie di ottimalità del secondo ordine

Distinguiamo tra condizioni di ottimalità (necessarie) senza qualificazione dei vincoli e con qualificazione dei vincoli. Le prime possono essere considerate generalizzazioni al secondo ordine