ISSN: 2281-1346

Department of Economics and Management

DEM Working Paper Series

Ruolo e uso dei vincoli di uguaglianza in problemi di ottimo con vincoli di disuguaglianza

Giorgio Giorgi
(Università di Pavia)

Angelo Guerraggio
(Università dell'Insubria e Università Bocconi)

 # 6 (10-12)

Via San Felice, 5
I-27100 Pavia
http://epmq.unipv.eu/site/home.html

October 2012

Ruolo e uso dei vincoli di uguaglianza in problemi di ottimo con vincoli di disuguaglianza

G. Giorgi[1] e A. Guerraggio[2]

Sommario

Si prendono in considerazione i legami esistenti tra i problemi "classici" di programmazione matematica, con vincoli espressi da uguaglianze, ed i problemi "moderni" di ottimo vincolato, ossia i problemi di programmazione matematica con vincoli espressi da disuguaglianze.

[1] Facoltà di Economia, Università di Pavia (Italy);
ggiorgi@eco.unipv.it

[2] Facoltà di Economia, Università dell'Insubria di Varese (Italy) e Università Bocconi di Milano (Italy);
angelo.guerraggio@uninsubria.it

1 Introduzione

La presente Nota, di intenti prevalentemente didattici, trae lo spunto da molti lavori (si vedano i riferimenti bibliografici, posti alla fine della Nota), anche di carattere manualistico, e dalla lettura dello storico articolo di Fritz John (1948).

Ad esempio, a pag. 179 della dispensa "Ottimizzazione Statica e Teoria dell'Utilità", Castagnoli e Cigola (2006) scrivono testualmente, allorché introducono i moderni problemi di programmazione matematica: "Il buon vecchio Giuseppe Luigi Lagrange aveva già visto (e detto) tutto. Se $\hat{x}$ è punto di massimo [di un problema con vincoli espressi da disuguaglianze], gli unici vincoli effettivi sono gli attivi ($s \in I(\hat{x}) = \{i \mid g_i(\hat{x}) = 0\}$; gli altri è come se non ci fossero in $\hat{x}$) e perciò il problema di ottimizzazione [con vincoli espressi da disuguaglianze] si può scrivere con i soli vincoli

$$b_s = g_s(\hat{x}), \quad s \in I(\hat{x}).$$

Detto chiaramente, la presenza di vincoli di disuguaglianza non è più generale della presenza di soli vincoli di uguaglianza! L'unico particolare è che i vincoli effettivi (attivi) cambiano da punto a punto (ovvero cambiano con $\hat{x}$), ma in ogni punto ci si può sempre ridurre a vincoli di uguaglianza".

Le precedenti affermazioni, un po' ingenerose e volutamente provocatorie per suscitare maggiore interesse da parte degli studenti, meritano qualche commento, onde evitare che i fondamentali lavori di W. Karush (1939), F. John (1948), H. W. Kuhn e A. W. Tucker (1951) si riducano a banali conseguenze dell'opera (indiscutibilmente geniale) del buon vecchio Giuseppe Luigi Lagrange.

Leggendo il contributo originale di Fritz John (1948) ci si può stupire che l'omonimo teorema non appaia nella stessa forma in cui è riportato da moltissimi testi ed articoli sulle condizioni di ottimalità in programmazione matematica. Anche qui citiamo testualmente da Fritz John (1948, pag. 543 e seguenti).

"Let R be a set of points x in a space E, and $F(x)$ a real-valued function defined in R. We consider a subset R' of R, which is described by a system of inequalities with parameter y :

$$G(x, y) \geq 0, \tag{1}$$

where G is a function defined for all x in R and all "values" of the parameter y [...]. To gain sufficient generality we assume that the "values" of the parameter y vary over a set of points S in a space H [...]. We are interested in conditions a point x^0 of R' has to satisfy in order that

$$M = F(x^0) = \underset{x \subset R'}{Minimum} F(x).$$

In what follows we restrict ourselves to the case, where the space E containing the set R is the $n-$dimensional euclidean space E_n, and where the set S of parameter values y

is a compact set in a metric space H. We make the further assumptions that $F(x)$ and $G(x,y)$, $F_i(x)$, $G_i(x,y)$ are continuous functions of (x,y) in $R \times S$ [...].

Theorem 1

Let x^0 be an interior point of R, and belong to the set R' of all points x of R, which satisfy (1) for all $y \subset S$. Let $F(x^0) = \underset{x \subset R'}{Minimum} F(x)$. Then there exists a finite set of points $y^1, ..., y^s$ in S and numbers $\lambda_0, \lambda_1, ..., \lambda_s$, which do not all vanish, such that

$$G(x^0, y^r) = 0 \;\; \text{for } r = 1, ..., s$$
$$\lambda_0 \geq 0, \lambda_1 > 0, ..., \lambda_s > 0$$
$$0 \leq s \leq n$$

the function

$$\Theta(x) = \lambda_0 F(x) - \sum_{r=1}^{s} \lambda_r G(x, y^r)$$

has a critical point at x^0, i. e.

$$\Theta_i(x^0) = 0 \quad \text{for } i = 1, ..., n\text{''}.$$

Se "traduciamo" il teorema di Fritz John per il caso di un problema di programmazione matematica con un numero finito di vincoli:

$$Min\, \{f(x) \mid g_j(x) \geq 0, \; j = 1, ..., p, \; x \in \mathbb{R}^n\}$$

risulta che se x^0 è soluzione locale di tale problema, esistono moltiplicatori $\lambda_0, \lambda_1, ..., \lambda_p$, *non negativi e non tutti nulli,* tali che

$$\begin{array}{l} \lambda_0 \nabla f(x^0) - \sum_{j=1}^{p} \lambda_j \nabla g_j(x^0) = 0 \\ \lambda_j g_j(x^0) = 0, \quad j = 1, ..., p. \end{array} \tag{1}$$

Inoltre il numero dei vincoli *fortemente attivi in* x^0 (ossia tali per cui è $g_j(x^0) = 0$ e $\lambda_j > 0$ nella (1)) non può superare n. Quest'ultima proprietà, quasi mai osservata nei testi (anche avanzati) di programmazione matematica, è collegata ad un noto teorema di Caratheodory che asserisce che ogni elemento $x \in co(S) \subset \mathbb{R}^n$ (dove $co(S)$ è l'involucro convesso di S) può essere espresso come combinazione lineare convessa di al più $n+1$ ellementi di S (cfr., ad esempio, Mangasarian (1969)). La precisazione induce anch'essa, al pari della provocatoria affermazione di Castagnoli e Cigola, a qualche spunto e a qualche riflessione sul ruolo e sull'uso dei vincoli di uguaglianza nei problemi di ottimo vincolato con vincoli di disuguaglianza, argomento che costituisce il titolo di questa Nota.

Il presente lavoro è strutturato come segue.

Nel § 2 prenderemo in considerazione i legami tra soluzioni di un usuale problema di programmazione non lineare

$$\text{(P)} \qquad \underset{x \in S}{Max} f(x) \qquad S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i = 1, ..., m\}$$

ed il problema (P_1) ottenuto da (P), considerando i soli vincoli attivi di (P) in $x^0 \in S$.

Nel § 3 prenderemo in rassegna i principali procedimenti adottati per ottenere le usuali condizioni di ottimalità di Karush-Kuhn-Tucker per (P), attraverso l'analisi di un problema con soli vincoli di uguaglianza, evitando quindi il ricorso a teoremi dell'alternativa o ad altri strumenti dell'Analisi Convessa.

Nel § 4 infine verrà considerato il ruolo dei cosiddetti vincoli "fortemente attivi" per (P) e delle "condizioni di complementarità stretta", nell'ambito dell'analisi di sensitività per (P) e nell'ottenimento di condizioni di ottimalità del secondo ordine.

2 Uso e ruolo dei vincoli attivi

Consideriamo il problema di programmazione non lineare

$$\text{(P)} \qquad \underset{x \in S}{Max} f(x), \quad S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i = 1, ..., m\},$$

ove $f,\ g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ sono differenziabili in un comune insieme aperto $X \subset \mathbb{R}^n$. Se $\bar{x} \in S$, è noto che l'insieme degli indici dei *vincoli attivi (o effettivi o aderenti)* in $\bar{x} \in S$ è dato da

$$I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}.$$

Ovviamente può risultare $I(\bar{x}) = \varnothing$ oppure $I(\bar{x}) = \{1, ..., m\}$. I vincoli tali per cui $g_i(\bar{x}) < 0$ sono detti *vincoli inattivi (o non effettivi o non aderenti)* in $\bar{x} \in S$. É immediato osservare (e l'osservazione è già presente nella storica tesi di W. Karush (1939)) che se $g_i(\bar{x}) < 0$ per qualche $i \in \{1, ..., m\}$, a causa della continuità, risulterà $g_i(x) < 0$ per tutti quei vettori x sufficientemente vicini a $\bar{x}$; di conseguenza è chiaro che i vincoli inattivi in $\bar{x}$ non influiscono *localmente* la "geometria" dell'insieme ammissibile S. Quindi, senza perdere in generalità, si può supporre che, per discorsi inerenti le condizioni necessarie di ottimalità locale per (P), tutti i vincoli siano attivi nel punto di massimo locale $\bar{x} \in S$. É dunque valida l'implicazione

$$\bar{x} \text{ soluzione locale di (P)} \implies \bar{x} \text{ soluzione locale di } (P_0),$$

essendo

$$(P_0) \qquad \underset{x \in S_0}{Max} f(x), \quad S_0 = \{x \in \mathbb{R}^n \mid g_i(x) = 0,\ i \in I(\bar{x})\}.$$

$\big($Ovviamente risulterà pure che $\bar{x}$ è soluzione locale del problema

(P$'_0$) $$\underset{x\in S'_0}{Max} f(x), \quad S'_0 = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i \in I(\bar{x})\}\ \big).$$

Quindi le proprietà di una soluzione *locale* x^0 del problema (P) sono determinate dai vincoli attivi in x^0. Ciò non vuole dire che il discorso valga anche per le soluzioni *globali* di (P) e che quindi in ogni caso possiamo trascurare i vincoli non attivi. Per esempio, se x^0 è soluzione globale di (P), omettendo un vincolo non attivo in x^0, potremmo fare diventare ammissibile un qualche vettore x^1 tale che $f(x^1) > f(x^0)$. Il lavoro di Evans e Gould (1972) fornisce un esempio grafico in $\mathbb{R}^2$ ove si mostra che in genere non è lecito sostituire il problema (P) con il problema (P$_0$), con riferimento a soluzioni globali. Gli stessi autori dimostrano un risultato concernente condizioni che rendono valida tale sostituzione. Va detto che il loro risultato era stato anticipato da Bector (1970) che a sua volta aveva generalizzato un teorema di Künzi, Krelle e Oettli (1966). Premettiamo la seguente definizione (cfr. anche Avriel e altri (1981)).

Definizione 1

La funzione $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, definita sull'insieme convesso $C \subset \mathbb{R}^n$ è detta *semistrettamente quasiconcava* su C se:

$$x^1, x^2 \in C, \quad f(x^2) > f(x^1) \Longrightarrow f(\lambda x^1 + (1-\lambda)x^2) > f(x^1),\ \forall \lambda \in (0,1).$$

É immediato fare vedere che le funzioni semistrettamente quasiconcave possono essere caratterizzate anche dall'implicazione

$$x^1,\ x^2 \in C, \quad f(x^1) \neq f(x^2) \Longrightarrow f(\lambda x^1 + (1-\lambda)x^2) > \min\left\{f(x^1), f(x^2)\right\}, \quad \forall \lambda \in (0,1).$$

Le funzioni semistrettamente quasiconcave sono chiamate "strettamente quasiconcave" da Mangasarian (1969) e Avriel (1976). Mangasarian dimostra che ogni punto di massimo locale di una funzione semistrettamente quasiconcava è anche di massimo globale. É poi noto il risultato di Karamardian (1967) che asserisce che una funzione semistrettamente quasiconcava è quasiconcava se è *superiormente semicontinua.*

Teorema 1 (Evans e Gould (1972))

Sia dato il problema (P), con f funzione semistrettamente quasiconcava su un convesso $C \subset \mathbb{R}^n$, sottoinsieme del dominio di f, e sia ogni $g_i, i \in I(\bar{x})$, funzione quasiconvessa su C $\big($oppure, più generalmente, sia convesso ogni insieme $\{x : g_i(x) \leq 0, i \in I(\bar{x})\}\big)$. Sia $\bar{x}$ soluzione globale di (P). Allora $\bar{x}$ è soluzione globale di (P$_0$).

Dim. Si veda Evans e Gould (1972).

Osservazione 1

1) Va notato che le ipotesi del Teorema 1 non implicano nè l'unicità di $\bar{x}$, nè la convessità dell'insieme ammissibile S. In effetti (P) potrebbe avere soluzioni locali che non sono globali.

2) Senza ulteriori ipotesi di convessità, concavità o concavità generalizzata, se $\bar{x}$ è soluzione globale di (P_0) e $\bar{x}$ è ammissibile per (P), allora $\bar{x}$ è soluzione globale per (P).

3) Abbiamo già puntualizzato che, senza ulteriori ipotesi di convessità, concavità o concavità generalizzata, se $\bar{x}$ è soluzione *locale* di (P), allora $\bar{x}$ è soluzione locale di (P_0). Osserviamo però che il viceversa non vale: se $\bar{x}$ è soluzione locale di (P_0), anche nel caso in cui (P) è problema di programmazione concava, $\bar{x}$ non è necessariamente ammissibile per (P) e, quando anche fosse ammissibile per (P), non è detto che $\bar{x}$ sia soluzione locale di (P).

4) Sulla base dell'implicazione descritta nel precedente punto e tenendo conto che per le funzioni semistrettamente quasiconcave vale il "teorema locale/globale", il risultato di Evans e Gould (di Bector, di Künzi e coautori) non appare proprio sorprendente. In proposito, osserviamo che le ipotesi sulla funzione obiettivo possono essere indebolite, considerando classi di funzioni più generali, per le quali vale comunque il "teorema locale/globale". Ad esempio possono essere considerate le *funzioni connesse per archi* e le loro generalizzazioni, introdotte da Ortega e Rheinboldt (1970) e analizzate sistematicamente da Avriel e Zang (1980).

5) Osserviamo che, a causa della continuità dei vincoli in (P), se è $I(x^0) = \varnothing$, allora $x^0 \in int(S)$. Viceversa, da $x^0 \in int(S)$, non segue che sia $I(x^0) = \varnothing$ (si pensi, ad es., a un vincolo identicamente nullo). Abbiamo poi già sottolineato che nell'analisi delle condizioni per l'ottimalità locale di un punto $\bar{x}$ ammissibile per (P), solo i vincoli di disuguaglianza attivi in $\bar{x}$ necessitano di essere considerati. Ciò, oltre a implicazioni teoriche ha anche implicazioni pratiche. Ricordiamo a tale proposito che la ripartizione dei vincoli (di disuguaglianza) in "attivi" e "non attivi" ha importanza anche in certi algoritmi di calcolo delle soluzioni di un problema del tipo (P). Ci sono infatti dei metodi ("active set methods") che si fondano sulla conoscenza a priori dei vincoli attivi in un problema di programmazione matematica. Le procedure non sono però semplici: si veda Luenberger (1989) e, per sviluppi recenti in tali direzioni, i lavori di Burke e Moré (1988), Burke (1990), Facchinei, Fisher, Kanzow (1999). Anche Rockafellar (1993), in un recente lavoro sui moltiplicatori di Lagrange-Kuhn-Tucker, osserva: "Obviously, in the theoretical study of the local optimality of $\bar{x}$ only the active inequality constraints at $\bar{x}$ need to be considered ... but as a practical matter it may be very hard to know without a lot of computation exactly which of the inequality constraints might turn out to be active".

3 Conversione di vincoli di disuguaglianza in vincoli di uguaglianza

Esporremo in questo paragrafo alcuni artifici, correnti nella letteratura, atti a trasformare un problema del tipo (P) in un problema con vincoli di uguaglianza ed eventuali vincoli di segno sulle variabili. Si possono allora ottenere le condizioni di ottimalità di Karush-Kuhn-Tucker senza utilizzare teoremi dell'alternativa o risultati di Analisi Convessa, facendo sostanzialmente riferimento alla teoria dei classici moltiplicatori di Lagrange. Va comunque detto che in genere tali artifici non permettono di ottenere le condizioni di Karush-Kuhn-Tucker per (P) sotto quelle ipotesi di generalità che, a partire dai lavori di Karush, John, Kuhn e Tucker, sono state di volta in volta adottate negli approcci più moderni.

Le trattazioni che utilizzano gli strumenti tipici dei problemi di ottimo vincolato "classici" (à la Lagrange) possono però avere un'utilità didattica non disprezzabile.

A) Alcuni autori, tra cui, ad es., Intriligator (1972), nell'ottenere le condizioni necessarie di ottimalità per un problema di programmazione non lineare del tipo (P), considerano dapprima il problema

$$(\mathrm{P}_1) \qquad \underset{x\geq 0}{Max} f(x),$$

ossia il problema ove non ci sono disuguaglianze del tipo $g_i(x) \leq 0,\ i = 1, ..., m$, ma solo vincoli di segno sulle variabili di decisione: $x_i \geq 0,\ i = 1, ..., n$. Si assume che f sia di classe C^2 (o anche che sia due volte differenziabile) e si utilizza la formula di Taylor, arrestata al secondo ordine. Se x^* è punto di ottimo locale per (P_1), risulterà

$$f(x^*) \geq f(x^* + tv)$$

con $t > 0$ opportuno e v versore di $\mathbb{R}^n$. Risulta poi

$$f(x^* + tv) = f(x^*) + t\nabla f(x^*)v + \frac{1}{2}t^2 v^\top \nabla_x^2 f(x^* + \theta tv)v, \text{ ove } 0 < \theta < 1.$$

Da tali due espressioni otteniamo la "disuguaglianza fondamentale" (secondo l'espressione di Intriligator)

$$t\nabla f(x^*)v + \frac{1}{2}t^2 v^\top \nabla_x^2 f(x^* + \theta tv)v \leq 0,$$

ossia

$$\nabla f(x^*)v + \frac{1}{2}tv^\top \nabla_x^2 f(x^* + \theta tv)v \leq 0.$$

Considerando il limite per $t \to 0^+$ abbiamo $\nabla f(x^*)v \leq 0$. Se x^* è punto interno all'ortante non negativo di $\mathbb{R}^n$ otteniamo le classiche condizioni di Fermat. Supponiamo allora che sia $x_i^* = 0$ per qualche i. Poniamo $v = e^i$, i-esimo vettore fondamentale di $\mathbb{R}^n$. Allora

$$0 \geq \nabla f(x^*)e^i = \frac{\partial f}{\partial x_i}(x^*).$$

Di conseguenza, indipendentemente dal segno di x_i^* ($x_i^* > 0$ o $x_i^* = 0$), sarà

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

Quindi le condizioni necessarie affinchè x^* sia soluzione locale di (P_1), ove f è, lo ricordiamo, due volte differenziabile in un intorno di x^*, sono:

$$\begin{array}{l} \nabla f(x^*) \leq 0 \\ \nabla f(x^*)x^* = 0 \\ x^* \geq 0. \end{array}$$

A questo punto il problema (P) viene convertito nel problema

$$(P_2) \qquad \underset{x \in S_2}{Max} f(x), \quad S_2 = \{x \in \mathbb{R}^n \mid g(x) + s = 0, s \geq 0\},$$

ove il vettore non negativo $s \in \mathbb{R}^m$ contiene le cosiddette "variabili di scarto" o "variabili ausiliarie" ("slack variables" nella letteratura in lingua inglese). Abbiamo quindi un problema di ottimo vincolato classico, ove però il numero delle variabili è $m + n$, con m variabili non negative. Applicando a (P_2) i risultati trovati per (P_1) si ottengono facilmente le condizioni di Karush-Kuhn-Tucker.Un limite alla generalità di questo approccio è costituito dal fatto che ovviamente occorrerà supporre che in x^* i vincoli $g_i(x)$ siano "regolari", onde potere applicare il classico teorema sui moltiplicatori di Lagrange, e che la funzione obiettivo sia due volte differenziabile.

B) Un approccio interessante, seguito da Luenberger (1989), Peressini, Sullivan e Uhl (1988) e da Simon e Blume (1994), è quello di ricorrere alle proprietà dei vincoli attivi, delineate nel precedente paragrafo. Trasformiamo l'elegante ma stringata dimostrazione di Luenberger in una versione didatticamente più "friendly".

Le ipotesi sono le solite: le funzioni di (P) sono almeno differenziabili e i vettori $\nabla g_i(x^*)$, $i \in I(x^*) = \{i \mid g_i(x^*) = 0\}$ sono linearmente indipendenti. Sia x^* punto di massimo locale per (P): x^* sarà allora anche punto di massimo locale per il problema (P_0), ove i vincoli sono dati da $g_i(x) = 0$, $i \in I(x^*)$. Per (P_0) esisteranno quindi gli usuali moltiplicatori di Lagrange; poichè $\lambda_j = 0$ se è $g_j(x^*) < 0$, otteniamo subito le condizioni di complementarità. Resta da dimostrare che è $\lambda \geq 0$. Ricordiamo il seguente risultato, la cui dimostrazione si trova, ad esempio, in Giannessi (1982).

Siano dati p vincoli di uguaglianza $h_1(x) = 0, \ldots, h_p(x) = 0$, essendo $h : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ di classe C^1 e sia x^0 un punto tale che $h(x^0) = 0$.

Lemma 1

La matrice Jacobiana $\nabla h(x^0)$ abbia caratteristica massima in x^0 (ossia x^0 è "punto regolare"). Allora per ogni $y \in \mathbb{R}^n$ tale che

$$\nabla h(x^0)y = 0 \tag{2}$$

esiste un arco differenziabile $x = x(t)$, $0 \le t \le \bar{t}$, con $x(0) = x^0$, contenuto nell'insieme ammissibile $S = \{x \mid h(x) = 0\}$, tale che $x'(0) = y$.

Viceversa, se esiste un arco differenziabile $x = x(t)$, $0 \le t \le \bar{t}$, con $x(0) = x^0$, contenuto in S, allora $y = x'(0)$ soddisfa la (2).

Si noti che l'insieme dei vettori y che soddisfano la (2) è, con riferimento a (P_0), il cono che Luenberger chaiama "spazio tangente", ma che usualmente viene chiamato *cono linearizzante* (cfr. Bazaraa e Shetty (1976)), mentre l'insieme dei vettori y tali per cui esiste un arco differenziabile, con le proprietà descritte nel lemma, viene usualmente chiamato *cono raggiungibile o cono delle direzioni raggiungibili o cono di Kuhn-Tucker* (cfr. Bazaraa e Shetty (1976), Giorgi e Guerraggio (1992), Giorgi, Guerraggio e Thierfelder (2004)).

Ritorniamo alla dimostrazione del risultato principale del presente approccio, ossia che il vettore λ dei moltiplicatori (di Kuhn-Tucker) deve essere non negativo. Supponiamo per assurdo che sia $\lambda_j < 0$ per un certo $j \in I(x^*)$. Sia S_0 l'insieme ammissibile di (P_0) e sia $x(t)$ un arco differenziabile in S_0 tale che $x^* = x(0)$. Poichè x^* è punto di massimo locale per (P_0), esisterà $r > 0$ tale che

$$f(x^*) \ge f(x), \ \forall x \in N(x^*, r) \cap S_0$$

e quindi

$$f(x^*) = f(x(0)) \ge f(x(t)) \text{ per ogni } t \in [0, \bar{t}] \, .$$

Ma allora risulta

$$\frac{d}{dt} f(x(t)) \mid_{t=0} = \nabla f(x(0)) x'(0) = \nabla f(x^*) x'(0) \le 0.$$

Consideriamo ora la superficie S_0^j definita da

$$g_i(x) = 0, \ i \in I(x^*), \ i \ne j.$$

Allora $x^* \in S_0^j$ a causa della definizione di $I(x^*)$. Se M_0^j è il relativo cono linearizzante o spazio tangente a S_0^j in x^*, poichè x^* è punto regolare per (P) e quindi per (P_0), esisterà un vettore $y \in S_0^j$ tale che $\nabla g_j(x^*) y < 0$ (si veda anche Simon e Blume (1994) per maggiori dettagli). Il Lemma 1 asserisce l'esistenza di un arco differenziabile $x(t)$ in S_0^j tale che $x(0) = x^*$ e $y = x'(0)$. Ma allora risulta

$$\frac{d}{dt} f(x(t)) \mid_{t=0} = \nabla f(x^*) y = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) y = \lambda_j \nabla g_j(x^*) y > 0,$$

e questo contraddice il fatto che x^* è punto di minimo locale per (P) (e per (P_0)).

Dimostrazioni analoghe si trovano nel già citato volume di Simon e Blume (1994) e in Peressini, Sullivan e Uhl (1988).

C) Alcuni autori (Hadley (1964), Gue e Thomas (1968), Schmidt e Davis (1981)) ricorrono all'artificio di trasformare vincoli di disuguaglianza in vincoli di uguaglianza, aggiungendo o sottraendo il quadrato di variabili di scarto, di modo che non occorre poi imporre esplicitamente la loro non negatività. Dunque, anzichè il problema (P) viene considerato il problema (equivalente, almeno dal punto di vista formale)

$$(\mathrm{P}_3) \qquad \underset{x \in S_3}{Max} f(x), \quad S_3 = \{x \in \mathbb{R}^n \mid g_i(x) + y_i^2 = 0,\ i = 1, ..., m\}\,.$$

La procedura, segnalata anche da Karush (1939) e che risale ad un lavoro di Valentine (1937) sul Calcolo delle Variazioni, ha, come le procedure di cui ai punti A) e B), pregi e difetti. Evita di trattare le disuguaglianze con strumenti "ad hoc", è didatticamente utile, permette di ottenere in modo semplice le condizioni sufficienti di ottimalità del secondo ordine per (P) (si veda, ad esempio, Bertsekas (1999)) ma non è una procedura molto soddisfacente dal punto di vista della sua generalità e delle ipotesi da assumere. Anche Giannessi (1982) è piuttosto critico a questo proposito e conclude che la trasformazione indicata, "pur risultando utile in alcuni casi, non elimina la necessità di trattare il caso delle disequazioni".

Sia x^0 soluzione locale del problema (P). Allora (x^0, y^0) è soluzione locale del problema (P_3), ove $y^0 = (y_1^0, ...y_m^0)$, essendo

$$y_i^0 = \Big(-g_i(x^0)\Big)^{\frac{1}{2}},\ i = 1, ..., m.$$

É facile verificare che x^0 è regolare per il problema (P) se e solo se (x^0, y^0) è regolare per il problema (P_3). Qui un vettore ammissibile x^0 è "regolare" se i gradienti dei vincoli espressi da uguaglianze, valutati in x^0, e/o i gradienti dei vincoli attivi in x^0, valutati sempre in x^0, sono linearmente indipendenti. Applicando allora a (P_3) il classico teorema sui moltiplicatori di Lagrange, otteniamo le condizioni necessarie di ottimalità

$$\begin{aligned}
\frac{\partial L(x^0, \lambda, y^0)}{\partial x_i} &= 0 = \nabla f(x^0) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0), \\
\frac{\partial L(x^0, \lambda, y^0)}{\partial \lambda_i} &= 0 = g_i(x^0) + (y_i^0)^2, \\
\frac{\partial L(x^0, \lambda, y^0)}{\partial y_i} &= 0 = -2\lambda_i y_i^0,\ i = 1, ..., m.
\end{aligned}$$

Poichè $y_i^0 = (-g_i(x^0))^{\frac{1}{2}} > 0$, $\forall i \notin I(x^0)$, l'ultima condizione segnala che è $\lambda_i \geq 0$, $\forall i \notin I(x^0)$. Abbiamo così le condizioni di complementarità. Resta, come al solito, da dimostrare che è $\lambda_i \geq 0$, $\forall i = 1, ..., m$. A questo punto ci si può appoggiare (come fanno gli autori citati all'inizio del presente punto) su proprietà della "funzione del valore ottimale", denotata $f^*(\epsilon)$, per il problema (P(ϵ)) (si veda L'Osservazione 2 nel presente paragrafo e si veda il paragrafo successivo) e in particolare sulla relazione (che vale sotto specifiche ipotesi)

$$\nabla_\epsilon f^*(\epsilon) = \lambda(\varepsilon)$$

che a sua volta implica

$$\nabla_\epsilon f^*(0) = \lambda(0) = \lambda.$$

A causa della relazione $-2\lambda_i y_i^0 = 0$, nella soluzione ottimale x^0 sarà $\lambda_i = 0$ o $y_i^0 = 0$ (o entrambe le relazioni). Assumiamo che sia $y_i^0 \neq 0$ per un indice i. Ciò implica che il vincolo $g_i(x)$ associato a y_i vale come disuguaglianza stretta (nel problema originario). Se consideriamo il vincolo $g_i(x) \leq \epsilon_i$ e permettiamo ad ϵ_i di aumentare con una variazione sufficientemente piccola, la funzione del valore ottimale sarà inalterata e quindi

$$\frac{\partial f^*(\varepsilon)}{\partial \varepsilon_i} = \lambda_i = 0.$$

Supponiamo ora $\lambda_i \neq 0$; ciò implica che $y_i{}^\circ = 0$. Sia, per assurdo $\lambda_i < 0$. Allora sarà

$$\frac{\partial f^*(\varepsilon)}{\partial \varepsilon_i} < 0.$$

Ciò significa che se ϵ_i aumenta, la funzione del valore ottimale diminuisce. Ma ciò è assurdo perchè se ϵ_i aumenta, l'insieme ammissibile diventa più ampio e la funzione del valore ottimale non può certo diminuire. Quindi, in corrispondenza di x^0, sarà $\lambda \geq 0$.

Un'altra procedura atta a dimostrare la non negatività del vettore dei moltiplicatori λ, nel presente approccio, utilizza le condizioni necessarie di ottimalità del secondo ordine per un problema, come (P_3), contenente solo uguaglianze. Occorre quindi assumere che le funzioni siano due volte differenziabili. Questa procedura è adottata, ad esempio, da El-Hodiri (1991); si veda anche Bertsekas (1999) per una trattazione più rigorosa e per i dettagli. Maggiori e più complete considerazioni sulla procedura di cui al presente punto, sono svolte da Taylor (1973) in un lavoro interessante anche per i riferimenti bibliografici. Rimandiamo di nuovo al paragrafo successivo per le condizioni che assicurano la differenziabilità della funzione del valore ottimale nel problema (P(ϵ)).

Osservazione 2

Un approccio ingegnoso all'ottenimento delle condizioni di Karush-Kuhn-Tucker per (P) viene offerto da Sydsaeter e altri (2008), ove si fa riferimento alla sola funzione del valore ottimale $f^*(\epsilon) = f[x^*(\epsilon), \epsilon]$, essendo $x^*(\epsilon)$ soluzione (anche locale) del problema

(P(ϵ)) $$\underset{x}{Max} \{f(x) \mid g_j(x) \leq \epsilon_j, \quad j = 1, ..., m.\}$$

Viene anzitutto fatta l'esplicita ipotesi che $f^*(\epsilon)$ sia differenziabile. Si nota poi che $f^*(\epsilon)$ è non decrescente in ogni variabile $\epsilon_1, ..., \epsilon_m$: come già osservato, se ϵ_j cresce, essendo ogni $\epsilon_i, i \neq j$, tenuta fissa, l'insieme ammissibile diventa più ampio e quindi $f^*(\epsilon)$ non può descrescere. Si fissi ora un vettore $\epsilon = \bar{\epsilon}$ e sia $\bar{x}$ la corrispondente soluzione ottima, ossia soddisfacente l'uguaglianza $f(\bar{x}) = f^*(\bar{\epsilon})$. Per ogni x ammissibile avremo $f(x) \leq f^*(g(x))$ e quindi $f^*(g(x)) \leq f^*(g(x) + \bar{\epsilon} - g(\bar{x}))$, poichè $g(\bar{x}) \leq \bar{\epsilon}$ e f^* è non decrescente. Ne segue che, posto $\varphi(x) = f(x) - f^*(g(x) + \bar{\epsilon} - g(\bar{x}))$, risulta $\varphi(x) \leq 0$ per ogni x. Poichè $\varphi(\bar{x}) = 0$, $\varphi(x)$ ha un massimo in $\bar{x}$, cosicchè

$$0 = \frac{\partial \varphi(\bar{x})}{\partial x_i} = \frac{\partial f(\bar{x})}{\partial x_i} - \sum_{j=1}^{m} \frac{\partial f^*(\bar{\epsilon})}{\partial \epsilon_j} \frac{\partial g_j(\bar{x})}{\partial x_i}, \quad i = 1, ...n. \tag{3}$$

Se definiamo

$$\lambda_j = \frac{\partial f^*(\bar{\epsilon})}{\partial \epsilon_j} \tag{4}$$

le n equazioni (3) forniscono la prima delle condizioni di Karush-Kuhn-Tucker. Poichè $f^*(\epsilon)$ è non decrescente, dalla (4) abbiamo $\lambda_j \geq 0$. Dimostriamo ora che se è $g_j(\bar{x}) < \bar{\epsilon}_j$, risulta $\lambda_j = 0$. Si scelga $\bar{\epsilon}' = (\bar{\epsilon}_1, ..., \bar{\epsilon}_{j-1}, \epsilon_j, \bar{\epsilon}_{j+1}, ..., \bar{\epsilon}_m)$ ove $\epsilon_j \in (g_j(\bar{x}), \bar{\epsilon}_j)$. Allora risulta $g(\bar{x}) \leq \bar{\epsilon}' \leq \bar{\epsilon}$. Essendo f^* non decrescente, abbiamo $f^*(g(\bar{x})) \leq f^*(\bar{\epsilon}') \leq f^*(\bar{\epsilon})$.

Ma l'ottimalità di $\bar{x}$ implica che $f(\bar{x}) = f^*(\bar{\epsilon})$, cosicchè risulta $f^*(\bar{\epsilon}') = f^*(\bar{\epsilon})$, allorchè $\epsilon_j \in (g_j(\bar{x}), \bar{\epsilon}_j)$. Sempre tenendo conto dell'ipotesi di differenziabilità di $f^*(\epsilon)$ in $\bar{\epsilon}$, ne segue che

$$\frac{\partial f^*(\bar{\epsilon})}{\partial \epsilon_j} = 0,$$

cosicchè dalla (4) è $\lambda_j = 0$. Abbiamo così ottenuto anche le condizioni di complementarità.

Osservazione 3

Esistono in letteratura, soprattutto in anni di poco posteriori alla pubblicazione del fondamentale lavoro di Kuhn e Tucker, approcci ad un problema di programmazione non lineare che utilizzano strumenti matematici tipici per problemi con vincoli di uguaglianza. (soprattutto il teorema sulle funzioni implicite). Si veda, ad esempio, Wilde (1962), Bernholtz (1964). Circa il contributo di quest'ultimo autore, va osservato quanto segue (si veda anche Giorgi e Guerraggio (1993)). In Bernholtz (1964) l'autore fa ricorso a due condizioni di qualificazione dei vincoli: la prima è essenzialmente quella utilizzata da Kuhn e Tucker (1951) nel loro famoso lavoro. Come già accennato, Bernholtz usa però, anzichè teoremi dell'alternativa (o di separazione tra insiemi convessi), il classico teorema sulle funzioni implicite, per ottenere le condizioni necessarie di Karush-Kuhn-Tucker. La seconda condizione proposta da Bernholtz è la seguente:

- L'insieme $I(x^0)$ ha cardinalità p, con $0 \leq p \leq m$; è $p \geq n$ e in $x^0 \in S$ almeno uno dei determinanti Jacobiani di ordine n, che possono essere costruiti con le funzioni g_i, $i \in I(x^0)$, è diverso da zero.

Tale condizione non garantisce però la validità delle condizioni di Karush-Kuhn-Tucker. Si consideri, ad esempio, il seguente problema.

$$\begin{array}{l} Max \;\; x_1 \\ (x_1)^2 - x_2 \leq 0 \\ x_2 \leq 0 \\ -x_1 \leq 0. \end{array}$$

Il punto $x^0 = (0, 0)$ che soddisfa la seconda condizione di Bernholtz, è di ottimo ma le corrispondenti condizioni di Karush-Kuhn-Tucker non sono soddisfatte in x^0.

4 Vincoli fortemente attivi

Dato $x^0 \in S$, i vincoli *fortemente attivi* in x^0 sono quei vincoli g_i, $i \in I(x^0)$, tali per cui nelle condizioni di Karush-Kuhn-Tucker per (P) risulta $\lambda_i > 0$. Porremo

$$I_+(x^0, \lambda) = \left\{ i \mid i \in I(x^0), \lambda_i > 0 \right\}$$

(insieme degli indici dei vincoli fortemente attivi in x^0). Si noti che $I_+(x^0, \lambda)$ dipende dal vettore dei moltiplicatori λ nelle condizioni di Karush-Kuhn-Tucker. Se in $x^0 \in S$ risulta $I_+(x^0, \lambda) = I(x^0)$, ossia è $\lambda_i > 0$, $\forall i \in I(x^0)$, allora in x^0 valgono le *condizioni di complementarità stretta*. Sia i vincoli fortemente attivi (in un punto $x^0 \in S$), sia le condizioni di complementarità stretta giocano un ruolo importante nell'analisi del problema (P).

Anzitutto Kyparisis (1985) ha dimostrato che la seguente condizione (di qualificazione dei vincoli) di *Mangasarian-Fromovitz stretta* (MFS):

in $x^0 \in S$ i vettori $\nabla g_i(x^0)$, $i \in I_+(x^0, \lambda)$, sono linearmente indipendenti ed esiste $d \in \mathbb{R}^n$ tale che

$$\begin{array}{rcl} d\nabla g_i(x^0) & < & 0, \forall i \in I(x^0) \setminus I_+(x^0, \lambda) \\ d\nabla g_i(x^0) & = & 0, \forall i \in I_+(x^0, \lambda) \end{array}$$

è *necessaria e sufficiente* per l'unicità dei moltiplicatori λ_i nelle condizioni di ottimalità di Karush-Kuhn-Tucker.

(In Giorgi e Zuccotti (2008) vi è un refuso nella definizione di (MFS): occorre sostituire $I_+(x,u)$ a $I(x)$, nella condizione di indipendenza lineare di $\nabla g_i(x)$). Inoltre si può dimostrare (cfr. Kyparisis (1985), Giorgi e Zuccotti (2008)) che la condizione (MFS) garantisce non solo le usuali condizioni necessarie di ottimalità del primo ordine (di Karush-Kuhn-Tucker), ma anche quelle del secondo ordine. É quindi una condizione di qualificazione dei vincoli, sia del primo che del secondo ordine, ovviamente più generale che non la *condizione di indipendenza lineare*, usata, ad esempio, da Fiacco e McCormick (1968). La *condizione di qualificazione dei vincoli di Mangasarian-Fromovitz* (MF):

dato $x^0 \in S$, esiste $d \in \mathbb{R}^n$ tale che

$$d\nabla g_i(x^0) < 0, \forall i \in I(x^0),$$

è invece una qualificazione dei vincoli soltanto del primo ordine e non anche del secondo ordine.
(Si vedano, al riguardo, i controesempi di Anitescu (2000)e di Baccari (2004)).

Riassumendo, se x^0 è punto ammissibile per (P), denotando con (LI) la condizione di indipendenza lineare, ossia la richiesta che i vettori $\nabla g_i(x^0)$, $\forall i \in I(x^0)$, siano linearmente indipendenti, si ha la seguente catena di implicazioni

$$\text{(LI)} \ [\Longrightarrow (SMF)] \Longrightarrow (MF),$$

ove l'implicazione tra parentesi quadre vale solo se x^0 è soluzione locale di (P), poichè solo in questo caso è possibile fare riferimento ai relativi moltiplicatori di Karush-Kuhn-Tucker, e quindi alla definizione di (SMF).

I vincoli fortemente attivi giocano un ruolo essenziale anche nelle condizioni sufficienti di ottimalità del secondo ordine. Riportiamo il seguente classico risultato, attribuito solitamente a McCormick (1967) ma dovuto sostanzialmente a Pennisi (1953).

Teorema 2

Sia $x^0 \in S$ punto che verifica le condizioni di Karush-Kuhn-Tucker per (P):

$$\nabla f(x^0) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0, \ \lambda_i g_i(x^0) = 0, \ \lambda_i \geq 0, \ i = 1, 2, ..., m,$$

e sia inoltre

$$d\nabla_x^2(f(x^0) - \lambda g(x^0))d < 0, \ \forall d \neq 0, d \in Z(x^0),$$
$$Z(x^0) = \left\{ d \in \mathbb{R}^n \mid d\nabla g_i(x^0) = 0, i \in I_+(x^0, \lambda); d\nabla g_i(x^0) \leq 0, i \in I(x^0) \diagdown I_+(x^0, \lambda) \right\}.$$

Allora x^0 è punto di massimo locale stretto per f su S (ossia soluzione locale stretta per (P)).

Osservazione 4

1) Le ipotesi del teorema 2 permettono in realtà di concludere (Hestenes (1966), (1975)) che x^0 è punto di massimo locale stretto *di ordine due,* ossia che è, con $\alpha > 0$,

$$f(x) \leq f(x^0) - \alpha \parallel x - x^0 \parallel^2, \ \forall x \in S \cap N(x^0),$$

ove $N(x^0)$ è un opportuno intorno di x^0.

2) Se nel punto x^0 valgono le *condizioni di complementarità stretta,* allora il cono $Z(x^0)$ assume la forma

$$Z_1(x^0) = \left\{ d \in \mathbb{R}^n \mid d\nabla g_i(x^0) = 0, \ i \in I(x^0) \right\}$$

e sono quindi applicabili i classici risultati sul segno di una forma quadratica soggetta ad un sistema di vincoli lineari omogenei. Si veda, ad esempio, Debreu (1952).

3) Va notato che alcuni autori (El-Hodiri (1971), (1991), Takayama (1985), Castagnoli e Peccati (1979), Simon e Blume (1994)) sostituiscono, nel precedente teorema, il cono $Z_1(x^0)$ al cono $Z(x^0)$, *senza assumere le condizioni di complementarità stretta.* In tale caso il teorema *non* fornisce condizioni sufficienti di ottimalità (del secondo ordine). Si veda Giorgi (2003) per un controesempio. Con riferimento al pregevole manuale di Simon e Blume, va segnalato che la versione italiana del medesimo, a cura di A. Zaffaroni, è stata emendata da tale errore.

Per altre osservazioni sulle condizioni di complementarità stretta, rimandiamo a Giorgi (2006). Qui ci occuperemo soltanto del ruolo svolto da tali condizioni nell'analisi di sensitività (o di sensibilità) per il problema (P(ϵ)). Tale analisi trova classiche applicazioni in Economia nello studio della "statica comparata" e del cosiddetto "teorema dell'inviluppo" (per quest'ultimo argomento si veda Giorgi e Zuccotti (2008)). Benchè l'analisi della sensitività per un problema di programmazione matematica con soli vincoli di uguaglianza (analisi che risale sostanzialmente a Samuelson (1947)) sia quasi sempre svolta in modo corretto (si veda, ad esempio, Afriat (1971), Intriligator (1972), Silberberg (1990)), altrettanto non si può dire con riferimento a quelle opere di economia matematica ove si estendono i risultati di sensitività "classici" a problemi con vincoli di disuguaglianza, del tipo (P). É il caso, ad esempio, di Intriligator (1972), Castagnoli e Peccati (1979), Mas-Colell e altri (1995), Simon e Blume (1994). Takayama (1977)è un pò più accurato.

Con riferimento a problemi più generali di (P), esistono in letteratura molti risultati di sensitività e stabilità. Una citazione d'obbligo è il libro di Fiacco (1983), purtroppo quasi sempre ignorato nei testi e negli articoli di analisi economica che trattano tali tematiche. I risultati di Fiacco sono facilmente "specializzati", con riferimento al problema parametrico

$$(\mathrm{P}(\epsilon)) \qquad \underset{x \in S(\epsilon)}{Max} f(x), \quad S(\epsilon) = \{x \in \mathbb{R}^n \mid g_i(x) \leq \epsilon_i, \ i = 1, ..., m\},$$

ottenendo, non solo i risultati che legano i moltiplicatori di Karush-Kuhn-Tucker alla "sensitività" della funzione del valore ottimale, ma anche risultati di sensitività del secondo ordine (del tutto ignorati nella letteratura economica).

Denotiamo con $f^*(\epsilon)$ la funzione del valore ottimale per $\mathrm{P}(\epsilon)$, detta anche "funzione di perturbazione" ("perturbation function"), ossia $f^*(\epsilon) = f[x(\epsilon), \epsilon]$ ove $x(\epsilon)$ è soluzione di $\mathrm{P}(\epsilon)$. La funzione Lagrangiana per $\mathrm{P}(\epsilon)$ è

$$L(x, \lambda, \epsilon) = f(x) - \sum_{i=1}^{m} \lambda_i (g_i(x) - \epsilon_i).$$

I risultati generali di Fiacco (1983) vengono allora specializzati nel seguente teorema.

Teorema 3

Se

(1) le funzioni implicate in $\mathrm{P}(\epsilon)$ sono di classe C^2 rispetto a x in un intorno di x^* per ϵ in un intorno dell'origine di $\mathbb{R}^m$;

(2) le condizioni sufficienti di ottimalità del secondo ordine per $\mathrm{P}(0)$ valgono in x^*, con associato vettore dei moltiplicatori di Karush-Kuhn-Tucker λ^*;

(3) i gradienti $\nabla_x g_i(x^*)$, $i \in I(x^*)$, sono linearmente indipendenti;

(4) $\lambda_i^* > 0$ per ogni $i \in I(x^*)$, ossia valgono le *condizioni di complementarità stretta.*

Allora:

(a) x^* è punto di minimo locale *isolato* per $\mathrm{P}(0)$ e l'associato vettore dei moltiplicatori λ^* è unico.

(b) Per ϵ in un intorno dell'origine (di $\mathbb{R}^m$), esiste un'unica funzione di classe C^1

$$y(\epsilon) = [x(\epsilon), \lambda(\epsilon)]^T$$

che soddisfa le condizioni sufficienti di ottimalità del secondo ordine per $\mathrm{P}(\epsilon)$ e tale che $y(0) = [x^*, \lambda^*]$ e quindi $x(\epsilon)$ è soluzione locale (unica) di $\mathrm{P}(\epsilon)$, con associato vettore dei moltiplicatori $\lambda(\epsilon)$.

(c) Per ϵ in un intorno dell'origine si ha

$$f^*(\epsilon) = L^*(\epsilon)$$

ove $L^*(\epsilon) = L[x(\epsilon), \lambda(\epsilon), \epsilon]$ può essere definita "funzione Lagrangiana del valore ottimale".

(d) La complementarità stretta e l'indipandenza lineare dei "gradienti attivi" valgono in $x(\epsilon)$ per ϵ in un intorno dell'origine di $\mathbb{R}^m$.

(e) Per ϵ in un intorno dell'origine risulta

$$\nabla_\epsilon f^*(\epsilon) = \lambda(\epsilon).$$

(f) Per ϵ in un intorno dell'origine risulta

$$\nabla_\epsilon^2 f^*(\epsilon) = \nabla_\epsilon \lambda(\epsilon).$$

Ovviamente il risultato (e) è la relazione che caratteriza i cosiddetti "prezzi ombra" nell'interpretazione economica dei moltiplicatori di Lagrange-Kuhn-Tucker. Ribadiamo che senza l'ipotesi (4) di complementarità stretta, di solito ignorata nelle trattazioni economiche, la funzione $f^*(\epsilon)$ può non essere differenziabile. Ciò che allora si può ottenere, come dimostrato da Jittorntrum (1981), è l'esistenza di derivate direzionali per $f^*(\epsilon)$. Si vedano, a tale proposito, anche i lavori di Fiacco e Kyparisis (1985) e Ralph e Dempe (1995). A commento del risultato sub (a), osserviamo che Robinson (1982) ha fornito un esempio di punto di ottimo locale stretto, ma non isolato, con riferimento a un problema di ottimo vincolato.

Bibliografia

S. N. AFRIAT (1971), Theory of maxima and the method of Lagrange, SIAM Journal on Appl. Mathematics, 20, 343-357.

M. ANITESCU (2000), Degenerate nonlinear programming with a quadratic growth condition, SIAM Journal on Optimization, 10, 1116-1135.

M. AVRIEL (1976), Nonlinear Programming. Analysis and Methods, Prentice Hall, Englewood Cliffs, New Jersey.

M. AVRIEL, I. ZANG (1980), Generalized arcwise-connected functions and characterizations of local-global minimum properties, J. Optim. Theory Appl., 32, 407-425.

M. AVRIEL, W. E. DIEWERT, S. SCHAIBLE, W. T. ZIEMBA (1981), Introduction to concave and generalized concave functions; in S. Schaible, W. T. Ziemba (Eds.), Generalized Concavity in Optimization and Economics, Academic Press, New York, 21-50.

A. BACCARI (2004), On the classical necessary second-order optimality conditions, J. Optim. Theory Appl., 123, 213-221.

M. S. BAZARAA, C. M. SHETTY (1976), Foundations of Optimization, Springer Verlag, Berlin.

C. R. BECTOR (1970), Some aspects of quasi-convex programming, Z. Angew. Math. Mech., 50, 495-502.

B. BERNHOLTZ (1964), A new derivation of the Kuhn-Tucker conditions, Operations Research, 12, 295-299.

D. P. BERTSEKAS (1999), Nonlinear Programming - Second Edition, Athena Scientific, Belmont, Mass.

J. V. BURKE, J. J. MORÉ (1988), On the identification of active constraints, SIAM J. Nemer. Anal., 25, 1197-1211.

J. BURKE (1990), On the identification of active constraints. II. The nonconvex case, SIAM J. Numer. Anal., 27, 1081-1103.

E. CASTAGNOLI, L. PECCATI (1979), Matematica per l'Analisi Economica. Vol. 2: Ottimizzazione Statica e Dinamica, ETAS LIBRI, Milano.

E. CASTAGNOLI, M. CIGOLA (2006), Ottimizzazione Statica e Teoria dell'Utilità, Le Dispense del Pellicano, Libreria Egea, Milano.

G. DEBREU (1952), Definite and semidefinite quadratic forms, Econometrica, 20, 295-300.

M. A. EL HODIRI (1971), Constrained Extrema. Introduction to the Differentiable Case withe Economic Applications, Springer Verlag, Lecture Notes in Operations Research and Mathematical Systems, N. 56, Berlin.

M. A. EL HODIRI (1991), Extrema of Smooth Functions. With Examples from Economic Theory, Springer Verlag, Berlin.

J. P. EVANS, F. J. GOULD (1972), on using equality-constraint algorithms for inequality constrained problems, Math. Programming, 2, 324-329.

F. FACCHINEI, A. FISCHER, C. KANZOW (1999), On the accurate identification of active constraints, SIAM J. Optim., 9, 14-32.

A. V. FIACCO (1983), Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, Academic Press, New York.

A. V. FIACCO, G. P. McCORMICK (1968), Nonlinear Programming: Sequential Unconstrained Minimization Techniques, J. Wiley, New York.

A. V. FIACCO, J. KYPARISIS (1985), Sensitivity analysis in nonlinear programming under second order assumptions, Lecture Notes in Control and Information Sciences, 66, Springer Verlag, 74-97.

F. GIANNESSI (1982), Metodi Matematici della Programmazione. Problemi Lineari e Non Lineari, Pitagora Editrice, Bologna.

G. GIORGI (2003), Some remarks on second-order optimality conditions, International Journal of Mathematical Sciences, 2, 97-102.

G. GIORGI (2006), Il ruolo della complementarità stretta in programmazione matematica, Report n. 290, Dipartimento di Statistica e Matematica Applicata all'Economia, Università di Pisa.

G. GIORGI, A. GUERRAGGIO (1992), On the notion of tangent cone in mathematical programming, Optimization, 25, 11-23.

G. GIORGI, A. GUERRAGGIO (1993), Constraint qualifications: the invex case, the set constraint case and other remarks, Atti XVII Convegno A.M.A.S.E.S., Ischia, 8-11 Settembre 1993, Istituto Italiano per gli Studi Filosofici, Napoli, 485-509.

G. GIORGI, A. GUERRAGGIO, J. THIERFELDER (2004), Mathematics of Optimization. Smooth and Nonsmooth Case, Elsevier, Amsterdam.

G. GIORGI, C. ZUCCOTTI (2008), Osservazioni sulle condizioni di ottimalità del secondo ordine in programmazione matematica, Quaderno di Ricerca X, a. a. 2007-2008, Facoltà di Economia, Univ. di Pavia, Dipartimento di Ricerche Aziendali, Aracne, Roma.

R. L. GUE, M. E. THOMAS (1968), Mathematical Methods in Operations Research, Collier-MacMillan, London.

G. HADLEY (1964), Nonlinear and Dynamic Programming, Addison-Wesley, Reading, Mass.

M. R. HESTENES (1966), Calculus of Variations and Optimal Control Theory, J. Wiley, New York.

M. R. HESTENES (1975), Optimization Theory: The Finite-Dimensional Case, J. Wiley, New York.

M. D: INTRILIGATOR (1972), Mathematical Optimization and Economic Theory, Prentice Hall, Englewood Cliffs, N. J.

K. JITTORNTRUM (1981), Solution point differentiability without strict complementarity in nonlinear programming, Math. Programming Study, 21, 127-138.

F. JOHN (1948), Extremum problems with inequalities as subsidiary conditions; in Studies and Essays, Courant Anniversary Volume, Interscience, New York, 187-204.

S. KARAMARDIAN (1967), Strictly quasi-convex (concave) functions and duality in mathematical programming, J. Math. Anal. Appl., 20, 344-358.

W. KARUSH (1939), Minima of Functions of Several Variables with Inequalities as Side Conditions, Master Thesis, Department of Mathematics, University of Chicago.

H. W. KUHN, A. W. TUCKER (1951), Nonlinear programming; in J. Neyman (Ed.), Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, Berkeley, 402-411.

H. P. KÜNZI, W. KRELLE , W. OETTLI (1966), Nonlinear Programming, Blaisdell Publishing Company, London.

J. KYPARISIS (1985), On uniqueness of Kuhn-Tucker multipliers in nonlinear programming, Math. Programming, 32, 242-246.

D. G. LUENBERGER (1989), Linear and Nonlinear Programming, Addison-Wesley Publishing Company, Reading, Mass.

O. L. MANGASARIAN (1969), Nonlinear Programming, McGraw-Hill Book Company, New York.

A. MAS-COLELL, M. D. WHINSTON, J. R. GREEN (1995), Microeconomic Theory, Oxford Univ. Press, Oxford.

G. P. McCORMICK (1967), Second order conditions for constrained minima, SIAM Journal on Applied Mathematics, 15, 641-652.

J. M. ORTEGA, W. RHEINBOLDT (1970), Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York.

L. L. PENNISI (1953), An indirect sufficiency proof for the problem of Lagrange with differential inequalities as side conditions, Trans. Amer. Math. Soc., 74, 177-198.

A. L. PERESSINI, F. E. SULLIVAN, J. J. UHL (1988), The Mathematics of Nonlinear Programming, Springer Verlag, New York.

D. RALPH, S. DEMPE (1995), Directional derivatives of the solution of a parametric nonlinear program, Mathematical Programming, 70, 159-172.

S. M. ROBINSON (1982), Generalized equations and their solutions, Part II (Applications to nonlinear programming), Math. Programming Study, 19, 200-211.

R. T. ROCKAFELLAR (1993), Lagrange multipliers and optimality, SIAM Review, 35, 183-238.

P. A. SAMUELSON (1947), Foundations of Economic Analysis, Harvard Univ. Press, Cambridge, Mass.

J. W. SCHMIDT, R. P. DAVIS (1981), Foundations of Analysis in Operations Research, Academic Press, New York.

E. SILBERBERG (1990), The Structure of Economics: A Mathematical Analysis, 2nd ed., Mc Graw-Hill, New York.

C. P. SIMON, L. BLUME (1994), Mathematics for Economists, Northon, New York. Versione italiana a cura di A. Zaffaroni: Matematica per l'Economia e le Scienze Sociali (2 volumi), Egea, Milano, 2002.

K. SYDASETER, P. HAMMOND, A. SEIERSTAD, A. STROM (2008), Further Mathematics for Economic Analysis, Second edition, Prentice Hall, London, New York.

A. TAKAYAMA (1977), Sensitivity analysis in economic theory, Metroeconomica, 29, 9-37.

A. TAKAYAMA (1985), Mathematical Economics, Cambridge Univ. Press, Cambridge.

J. C. TAYLOR (1973), A squared-variable transformation approach to nonlinear programming optimality conditions, Naval Research Logistics Quart., 20, 25-39.

F. VALENTINE (1937), The problem of Lagrange with differential inequalities as added side conditions; in Contributions to the Calculus of Variations, 1933-1937, The University of Chicago Press, Chicago, Illinois, 403-447.

D. WILDE (1962), Differential calculus in nonlinear programming, Operations Research, 10, 764-773.