ISSN: 2281-1346

Department of Economics and Management

DEM Working Paper Series

A Tutorial on Sensitivity and Stability in Nonlinear Programming and Variational Inequalities under Differentiability Assumptions

Giorgio Giorgi
(Università di Pavia)

Cesare Zuccottti
(Università di Pavia)

 # 159 (06-18)

Via San Felice, 5
I-27100 Pavia
economiaweb.unipv.it

June 2018

A Tutorial on Sensitivity and Stability in Nonlinear Programming and Variational Inequalities under Differentiability Assumptions

Giorgio Giorgi (*) and Cesare Zuccotti (**)

Abstract

In this paper basic results on sensitivity and stability analysis in differentiable nonlinear programming problems are surveyed. We follow mainly the approach of A. V. Fiacco and his co-authors. We give also an overview on sensitivity for right-hand side perturbations of the parameters, on sensitivity for parametric linear programming problems and on sensitivity for parametric variational inequality problems.

Key words

Sensitivity analysis, stability analysis, nonlinear programming, linear programming, variational inequalities.

AMS Subject Classification (2010): Primary, 90C30; Secondary, 49A29.

1. Introduction

The aim of this paper is to offer a tutorial in the main stability and sensitivity aspects of parametric nonlinear programming (NLP) problems and variational inequality (VI) problems, under suitable differentiability assumptions on the functions involved in the said problems. We prefer to follow the basic approach of Fiacco and his collaborators (Fiacco (1976), Fiacco (1980), Fiacco (1983a,b), Fiacco and McCormick (1968), Fiacco and Hutzler (1982a,b), Fiacco and Ishizuka (1990), Fiacco and Kyparisis (1985), Fiacco and Liu(1993)) rather than the more abstract and sophisticated approach of Bonnans and Shapiro (1998, 2000), Bonnans and Cominetti (1996) and other authors.

We point out also the book of Bank and others (1982) for another classical overview on parametric nonlinear programming problems.

The paper will be concerned with the stability and sensitivity aspects of a parametric nonlinear programming (NLP) problem of the form

$$(NLP(\varepsilon)): \quad \begin{cases} \min f(x,\epsilon) \\ \text{s. t. } x \in K(\epsilon) = \{x : g(x,\varepsilon) \leqq 0,\ h(x,\varepsilon) = 0\}, \end{cases}$$

where $f : \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}^m$, $h : \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}^p$, and $\varepsilon \in \mathbb{R}^r$ is a vector of perturbation parameters.

(*) Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, (Italy). E-mail: giorgio.giorgi@unipv.it
(**) Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, (Italy). E-mail: cesare.zuccotti@unipv.it

We shall also be concerned with a parametric variational inequality (VI) problem of the form

$$(VI(\varepsilon)) : \quad \begin{cases} \text{Find } x^* \in K(\epsilon) \text{ such that} \\ F(x^*, \varepsilon)^\top (x - x^*) \geqq 0, \text{ for every } x \in K(\epsilon), \end{cases}$$

where F is a function from $\mathbb{R}^n \times \mathbb{R}^r$ to $\mathbb{R}^n$.

We shall consider that $\varepsilon^0 \in \mathbb{R}^r$ is a fixed parameters vector: if $\varepsilon^0 = 0$ we have the (non perturbed) classical nonlinear programming problem (NLP) and the classical variational inequality problem (VI).

Nonlinear programming problems and variational inequality problems are closely related to each other. It is well-known that the first order optimality conditions (Karush-Kuhn-Tucker conditions) of a nonlinear programming problem are a structured variational inequality problem, thus general results concerning the properties of solutions of variational inequalities apply to stationary solutions of nonlinear programs. On the other hand, a variational inequality can be written as the Karush-Kuhn-Tucker conditions of a nonlinear programming problem, if the Jacobian matrix $\nabla F(x)$ of the vector function F is symmetric for all x and some constraint qualification holds at all x^*. In this case, F is actually the gradient vector function of the objective function of the corresponding nonlinear program.

The paper is organized as follows.

Section 2 intends to give a review of the basic optimality and regularity conditions commonly employed in the study of stability and sensitivity analysis of nonlinear programming problems.

Section 3 is concerned with basic sensitivity results under second order differentiability for $(NLP(\varepsilon))$.

Section 4 contains some notes on the directional differentiability of the perturbed local minimum vector $x(\varepsilon)$ of $(NLP(\varepsilon))$.

Section 5 is concerned with differential stability of the "optimal value function" of $(NLP(\varepsilon))$.

Section 6 treats some basic sensitivity results for right-hand side perturbations, i. e. for a nonlinear programming problem where the perturbation parameters appear only on the right-hand side of the constraints g_i, $i = 1, ..., m$, and h_j, $j = 1, ..., p$.

Section 7 contains some notes on sensitivity results for linear programming.

Section 8 is concerned with some results on sensitivity for variational inequalities.

The final Section 9 summarizes some results of Armacost and Fiacco (1975, 1978, 1979) and Fiacco (1976) on the use of a class of penalty functions to estimate some sensitivity results for a parametric programming problem $(NLP(\varepsilon))$.

The first-order partial derivatives of f with respect to x and ε are denoted by $\nabla_x f$ and $\nabla_\varepsilon f$, respectively. The second-order partial derivatives of f are denoted by $\nabla_x^2 f$, $\nabla_{\varepsilon x}^2 f$ and $\nabla_\varepsilon^2 f$. Both $\nabla_x f$ and $\nabla_\varepsilon f$ are usually row vectors.

2. Optimality and Regularity Conditions in (NLP)

Consider a nonlinear programming (NLP) problem of the form

$$(NLP): \quad \begin{cases} \min\limits_{x \in \mathbb{R}^n} f(x) & \\ \text{subject to:} & g_i(x) \leqq 0, \ i = 1, ..., m, \\ & h_j(x) = 0, \ j = 1, ..., p, \end{cases}$$

where f, g_i, $h_j : \mathbb{R}^n \longrightarrow \mathbb{R}$; it is assumed in this section that f, every g_i and every h_j are twice continuously differentiable around

$$x^0 \in K = \{x \in \mathbb{R}^n : g_i(x) \leqq 0, \ i = 1, ..., m; \ h_j(x) = 0, \ j = 1, ..., p\}.$$

Note that (NLP) is (NLP(ε)), with $\varepsilon = 0$.

The *Lagrangian function* associated with (NLP) is defined as:

$$\mathcal{L}(x, u, w) = f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{j=1}^{p} w_j h_j(x),$$

where $u = [u_1, ..., u_m]^\top$ and $w = [w_1, ..., w_p]^\top$ are the Lagrange-Kuhn-Tucker multipler vectors associated with inequality and equality constraints g_i and h_j, respectively.

A point $x^0 \in K$ is *feasible* for (NLP). A point $x^0 \in K$ is a local minimum for (NLP) if there exists a neighborhood $N(x^0)$ of x^0 such that $f(x) \geqq f(x^0)$, $\forall x \in N(x^0) \cap K$. The following result is perhaps the most known and important statement of first order necessary conditions for (local) optimality of $x^0 \in K$.

Theorem 1 (Karush (1939), Kuhn and Tucker (1951)). Let $x^0 \in K$ be a local minimum for (NLP) and let an appropriate *constraint qualification* (to be stipulated) hold at x^0. Then, the Karush-Kuhn-Tucker (KKT) conditions hold at x^0, i. e. there exist Lagrange-Kuhn-Tucker multiplier vectors u^0 and w^0 such that:

$$(KKT): \quad \begin{cases} \nabla_x \mathcal{L}(x^0, u^0, w^0) = 0, \\ u_i^0 g_i(x^0) = 0, \ i = 1, ..., m, \\ u_i^0 \geqq 0, \ i = 1, ..., m. \end{cases}$$

Define the *set of active indices* of the inequality constraints at $x^0 \in K$:

$$I(x^0) = \left\{i : 1 \leqq i \leqq m, \ g_i(x^0) = 0\right\};$$

and the *set of strictly active inequality constraints at* $x^0 \in K$:

$$I^+(u^0, w^0) = \left\{ \begin{array}{l} i : i \in I(x^0) \text{ and there is a } (u^0, w^0) \\ \text{satisfying (KKT), with } u_i^0 > 0 \end{array} \right\}.$$

There are many constraint qualifications which assure the validity of the (KKT) conditions. Here we list only the ones which will be used in the present paper. Let $x^0 \in K$.

(a) The *Mangasarian-Fromovitz Constraint Qualification* (MFCQ) holds at x^0 if

(i) The vectors $\nabla h_j(x^0)$, $j = 1, ..., p$, are linearly independent;

(ii) There is z such that

$$\begin{aligned} \nabla g_i(x^0)z &< 0, \ i \in I(x^0), \\ \nabla h_j(x^0)z &= 0, \ j = 1, ..., p. \end{aligned}$$

Applying a theorem of the alternative (see, e. g., Mangasarian (1969)), the equivalent dual form of (MFCQ) states that:

- zero is the unique solution of the relations

$$\sum_{i \in I(x^0)} u_i \nabla g_i(x^0) + \sum_{j=1}^{p} w_j \nabla h_j(x^0) = 0, \ u_i \geqq 0, \quad \forall i \in I(x^0).$$

(b) The *Linear Independence Constraint Qualification* (LICQ) holds at x^0 if the vectors

$$\left\{ \nabla g_i(x^0), \ i \in I(x^0); \ \nabla h_j(x^0), \ j = 1, ..., p \right\}$$

are linearly independent.

(c) The *Strict Mangasarian-Fromovitz Constraint Qualifications* (SMFCQ) holds at x^0 if:

(i) The gradients

$$\left\{ \nabla g_i(x^0), \ i \in I^+(u^0, w^0); \ \nabla h_j(x^0), \ j = 1, ..., p \right\}$$

are linearly independent.

(ii) There is z such that

$$\begin{aligned} \nabla g_i(x^0)z &< 0, \ i \in I(x^0) \diagdown I^+(u^0, w^0), \\ \nabla g_i(x^0)z &= 0, \ i \in I^+(u^0, w^0), \\ \nabla h_j(x^0)z &= 0, \ j = 1, ..., p. \end{aligned}$$

(d) The *Constant Rank Condition* (CR) holds at x^0 if for any subset $L \subset I(x^0)$ of active constraints, the family

$$\nabla g_i(x), \ i \in L; \ \nabla h_j(x), \ j = 1, ..., p$$

remains of constant rank near the point x^0.

(e) The *Weak Constant Rank Condition* (WCR) holds at x^0 if for any subset $L \subset I^+(u^0, w^0)$ of strictly active constraints, the family

$$\nabla g_i(x), \ i \in L; \ \nabla h_j(x), \ j = 1, ..., p$$

remains of constant rank near the point x^0.

Conditions $(a) - (d)$ are well-known in the literature. In particular, condition (c) was introduced by Kyparisis (1985) who has shown that this condition is both necessary and sufficient to have uniqueness of multiplier vectors in (KKT) conditions. Note however, that the (SMFCQ) condition is not properly a constraint qualification, as it involves the multiplier vectors in its definition. Usually, these multiplier vectors depend also from the objective function f; see, e. g., Wachsmuth (2013). The (CR) condition was introduced by Janin (1984) and the (WCR) condition was introduced by Liu (1995b). The next result summarizes the relationships between these conditions.

Theorem 2. The following implications hold.

- (i) $(LI) \Longrightarrow (SMFCQ) \iff$ uniqueness of (KKT) multipliers $\Longrightarrow (MFCQ) + (WCR)$.
- (ii) $(LI) \Longrightarrow (MFCQ) + (CR) \Longrightarrow (MFCQ) + (WCR)$.
- (iii) $(CR) \Longrightarrow (WCR)$.
- (iv) $(SMFCQ) \nRightarrow (CR)$; $(CR) \nRightarrow (MFCQ)$.

The logical relations $(LI) \Longrightarrow (SMFCQ) \iff$ uniqueness of (KKT) multipliers $\Longrightarrow (MFCQ)$ can be found in Kyparisis (1985).

The logical relations $(LI) \Longrightarrow (MFCQ) + (CR)$ and $(SMFCQ) \nRightarrow (CR)$ were obtained by Kyparisis (1990b).

The relation $(CR) \nRightarrow (MFCQ)$ was proved by Janin (1984).

The implications $(SMFCQ) \Longrightarrow (WCR)$ and $(CR) \Longrightarrow (WCR)$ are evident.

The counterexamples showing $(MFCQ) + (WCR) \nRightarrow (SMFCQ)$ and $(WCR) \nRightarrow (CR)$ were obtained by Liu (1995b).

The next theorem states the classical second order necessary conditions for local optimality in problem (NLP). These conditions are essentially due to McCormick (1967, 1976); see also Fiacco and McCormick (1968).

Theorem 3. Suppose that x^0 is a local solution of (NLP) and that the (LI) conditions hold at x^0. Then, the (KKT) conditions hold at x^0 with associated unique multiplier vectors u^0 and w^0, and the additional Second Order Necessary Conditions (SONC) hold at x^0 :

$$z^\top \nabla_x^2 \mathcal{L}(x^0, u^0, w^0) z \geqq 0$$

for all $z \in Z(x^0)$, where $Z(x^0)$ is the so-called *critical cone or cone of critical directions,* defined as follows:

$$Z(x^0) = \left\{ z \in \mathbb{R}^n : \begin{array}{c} \nabla g_i(x^0) z \leqq 0, \ i \in I(x^0), \\ \nabla g_i(x^0) z = 0, \ i \in I(x^0) \text{ such that } u_i^0 > 0, \\ \nabla h_j(x^0) z = 0, \ j = 1, ..., p \end{array} \right\}.$$

By strengthening (SONC) one obtains the following standard second order sufficient conditions for local "strict" optimality, conditions due to Pennisi (1953), McCormick (1967) and Fiacco and McCormick (1968).

Theorem 4. Suppose that the (KKT) conditions hold at $x^0 \in K$ for (NLP) with multipliers vectors u^0 and w^0, and that the following additional Second Order Sufficient Conditions (SOSC) hold at x^0 :

$$z^\top \nabla_x^2 \mathcal{L}(x^0, u^0, w^0) z > 0$$

for all $z \neq 0$, $z \in Z(x^0)$.

Then, x^0 is a strict local minimum for (NLP), i. e. $f(x^0) < f(x)$ for all feasible x in some neighborhood of x^0, $x \neq x^0$.

Han and Mangasarian (1979) noted that if the (KKT) conditions are verified at $x^0 \in K$, the restrictions on z are equivalent to: $z \neq 0$, $\nabla f(x^0)z = 0$, $\nabla g_i(x^0)z \leqq 0$, $i \in I(x^0)$, and $\nabla h_j(x^0)z = 0$ for every $j = 1, ..., p$.

Robinson (1982) pointed out that the sufficient conditions of Theorem 4 do not assure that x^0 is an *isolated* (i. e. locally unique) local minimum for (NLP), as indicated by Fiacco and McCormick (1968). See also Fiacco (1983a). Robinson (1982) provides the following counterexample in $\mathbb{R}$.

Example 1. Minimize $f(x) = \frac{1}{2}x^2$, $x \in \mathbb{R}$, subject to $h_1(x) = x^6 \sin\left(\frac{1}{x}\right) = 0$, where $h_1(0) = 0$ by definition.

One can verify that the conditions of Theorem 4 are verified at $x^0 = 0$. Moreover, every point of the set

$$\left\{(n\pi)^{-1},\ n = \pm 1, \pm 2, ...\right\}$$

is an isolated feasible point, and therefore also a local minimum. Thus $x^0 = 0$ is *not* an isolated local minimum.

Other second order sufficient optimality conditions for (NLP), used in the literature, are the following ones,

a) The *Strong Second Order Sufficient Conditions* (SSOSC) hold at $x^0 \in K$ with multipliers (u^0, w^0) if

$$z^\top \nabla_x^2 \mathcal{L}(x^0, u^0, w^0) z > 0$$

for all $z \neq 0$ such that

$$\begin{aligned} \nabla g_i(x^0)z &= 0, \quad \text{if } u_i^0 > 0, \\ \nabla h_j(x^0)z &= 0, \quad j = 1, ..., p. \end{aligned}$$

b) The *General Second Order Suffcient Conditions* (GSOSC) hold at $x^0 \in K$ if

- (SOSC) hold at $x^0 \in K$ with (u^0, w^0), for *every* (u^0, w^0) satisfying (KKT).

c) The *General Strong Second Order Sufficient Conditions* (GSSOSC) hold at $x^0 \in K$ if

- (SSOSC) hold at $x^0 \in K$ with (u^0, w^0), for *every* (u^0, w^0) satisfying (KKT).

The following conditions, due to Robinson (1982), are sufficient for $x^0 \in K$ to be an *isolated* local minimum for problem (NLP).

Theorem 5. Let $x^0 \in K$ and suppose that the (KKT) conditions hold at x^0 with some u^0 and w^0. Suppose that the Mangasarian-Fromovitz constraint qualification holds at x^0; moreover, assume that the General Second Order Sufficient Conditions hold at x^0. Then x^0 is an isolated local minimum of (NLP), i. e. there exists a neighborhood of x^0 such that x^0 is the only local minimum of (NLP) in that neighborhood.

Remark 1. Note that if the (LICQ) or also the (SMFCQ) is substituted for (MFCQ) in Theorem 5, then (GSOSC) coincides with (SOSC), since the multipliers vectors u^0 and w^0 are unique, under the said constraint qualifications.

3. Basic Sensitivity Results under Second Order Differentiability for (NLP(ε))

The parametric nonlinear programming problem we take into consideration is defined as:

$$(NLP(\varepsilon)): \quad \begin{cases} \min f(x,\varepsilon) \\ \text{subject to:} \quad g_i(x,\varepsilon) \leqq 0, \; i = 1, ..., m, \\ \qquad\qquad\quad h_j(x,\varepsilon) = 0, \; j = 1, ..., p, \end{cases}$$

where $\varepsilon \in \mathbb{R}^r$ is the perturbation parameters vector, $f, g_i, h_j : \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}$. For simplicity we assume that the functions f, g_i, h_j are $\mathcal{C}^2$ in (x,ε) in some neighborhood of $(x^0, \varepsilon^0) \in K(\varepsilon)$, where $K(\varepsilon)$ is the feasible set of $(NLP(\varepsilon))$.

The Lagrangian function associated with $(NLP(\varepsilon))$ is defined by:

$$\mathcal{L}(x,u,w,\varepsilon) = f(x,\varepsilon) + u^\top g(x,\varepsilon) + w^\top h(x,\varepsilon),$$

where $g = (g_1, ..., g_m)^\top$ and $h = (h_1, ..., h_p)^\top$.

If $\varepsilon = \varepsilon^0$ is fixed, then all the definitions given in the previous section apply to problem $(NLP(\varepsilon^0))$. In particular, we denote by $I(x^0, \varepsilon^0)$ the set of active constraints for $(NLP(\varepsilon^0))$ and by $K(\varepsilon^0)$ the feasible set for $(NLP(\varepsilon^0))$.

The following basic results concerning the differentiability of a perturbed solution of a parametric nonlinear programming problem were originally proved by Fiacco and McCormick (1968) for a special case of $(NLP(\varepsilon))$ and later for the general version $(NLP(\varepsilon))$ by Fiacco (1976, 1983a).

Theorem 6. Let be given $(NLP(\varepsilon^0))$. Suppose that $x^0 \in K(\varepsilon^0)$ and that the (KKT) conditions hold for x^0, with associated multiplier vectors u^0, w^0. Moreover, suppose that the (LICQ) holds at x^0, that the *Strict Complementarity Slackness Condition* (SCS) holds at x^0 with respect to u^0, i. e.

$$u_i^0 > 0 \text{ when } g_i(x^0, \varepsilon^0) = 0.$$

Finally, suppose that the (SOSC) hold at x^0, with (u^0, w^0).

Then:

(a) x^0 is an isolated local minimum of $(NLP(\varepsilon^0))$ and the associated multiplier vectors u^0 and w^0 are unique.

(b) For ε in a neighborhood of ε^0 there exists a unique once continuously differentiable vector function

$$y(\varepsilon) = [x(\varepsilon), u(\varepsilon), w(\varepsilon)]^\top$$

satisfying the second order sufficient conditions for a local minimum of $(NLP(\varepsilon))$ such that $y(\varepsilon^0) = (x^0, u^0, w^0)^\top$ and, hence, $x(\varepsilon)$ is a locally unique local minimum of $(NLP(\varepsilon))$ with associated unique multiplier vectors $u(\varepsilon)$ and $w(\varepsilon)$.

(c) The Linear Independence and Strict Complementarity Slackness Conditions hold at $x(\varepsilon)$ for ε near ε^0.

Besides Fiacco (1976, 1983a), Theorem 6 has been considered in an unpublished paper of Duggan and Kalandrakis (2007). See also Robinson (1974) and Bigelow and Shapiro (1974). Diewert (1984) takes into consideration the main results of Theorem 6, referred to a nonlinear parametric programming problem with nonnegative variables and a scalar parameter. Other applications of sensitivity results for $(NLP(\varepsilon))$ in economics are provided, e. g., by Silberberg and Suen (2001) and by Takayama (1977, 1985). Finally, we point out that Fiacco (1983a) and Fiacco and Kyparisis (1985) give Theorem 6 under slightly differentiability assumptions: the functions f, g_i $(i = 1, ..., m)$ and h_j $(j = 1, ..., p)$ and their partial derivatives with respect to x are $\mathcal{C}^1$ in (x, ε) in some neighborhood of (x^0, ε^0). Also Robinson (1974) proved an analogous theorem under weaker differentiability assumptions.

Fiacco (1976) also shows that the derivative of $y(\varepsilon)$ can be calculated by noting that the following system of Karush-Kuhn-Tucker equations will hold at $y(\varepsilon)$ for ε near ε^0 under the assumptions of Theorem 6,

$$\begin{cases} \nabla_x \mathcal{L}(x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon) = 0, \\ u_i(\varepsilon) g_i(x(\varepsilon), \varepsilon) = 0, \ i = 1, ..., m, \\ h_j(x(\varepsilon), \varepsilon) = 0, \ j = 1, ..., p. \end{cases} \tag{1}$$

The assumptions of Theorem 6 imply that the Jacobian $M(\varepsilon)$ of the said system, with respect to (x, u, w) is locally nonsingular; therefore we obtain

$$M(\varepsilon) \nabla_\varepsilon y(\varepsilon) + N(\varepsilon) = 0,$$

where $N(\varepsilon)$ is the Jacobian of system (1) with respect to ε. From this last equality we get

$$\nabla_\varepsilon y(\varepsilon) = \begin{bmatrix} \nabla_\varepsilon x(\varepsilon) \\ \nabla_\varepsilon u(\varepsilon) \\ \nabla_\varepsilon w(\varepsilon) \end{bmatrix} = -(M(\varepsilon))^{-1} N(\varepsilon)).$$

At $\varepsilon = \varepsilon^0$ we have, with $M(\varepsilon^0) = M^0$ and $N(\varepsilon^0) = N^0$,

$$M^0 \nabla_\varepsilon y(\varepsilon^0) = -N^0,$$

where

$$M^0 = \begin{bmatrix} \nabla_x^2\mathcal{L} & (\nabla_x g_1)^\top & \cdots & (\nabla_x g_m)^\top & (\nabla_x h_1)^\top & \cdots & (\nabla_x h_p)^\top \\ u_1\nabla_x g_1 & g_1 & \cdots & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 & \cdots & \cdots & \vdots \\ u_m\nabla_x g_m & 0 & \cdots & g_m & \cdots & \cdots & 0 \\ \nabla h_1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ \nabla h_p & 0 & 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}$$

and

$$N^0 = \left[\nabla^2_{\varepsilon x}\mathcal{L}^\top, u_1(\nabla_\varepsilon g_1)^\top, ..., u_m(\nabla_\varepsilon g_m)^\top, (\nabla_\varepsilon h_1)^\top, ..., (\nabla_\varepsilon h_p)^\top\right]^\top$$

are evaluated at $(x^0, u^0, w^0, \varepsilon^0)$.

The next result, due to McCormick (1976), shows that the conditions imposed in Theorem 6 are nearly minimal for the nonsingularity of the Jacobian of system (1).

Theorem 7. Suppose that the second order necessary conditions for a local minimum of $(NLP(\varepsilon^0))$ hold at x^0 with associated multiplier vectors u^0 and w^0 (i. e. the (KKT) and the (SONC) conditions hold at x^0 with (u^0w^0)). Then the Jacobian matrix of system (1) with respect to (x, u, w) is nonsingular at (x^0, u^0, w^0) *if and only if* the (SOSC), (LICQ) and (SCS) conditions hold at x^0 with (u^0w^0) for $(NLP(\varepsilon^0))$.

We give now some sensitivity results for $(NLP(\varepsilon))$, *without the Strict Complementarity Slackness assumption.* These results are essentially due to Jittorntrum (1984) and, in a more general form, to Robinson (1980).

We recall that, given an arbitrary function $\phi : \mathbb{R}^n \longrightarrow [-\infty, +\infty]$, and a point $(x, d) \in \mathbb{R}^{2n}$, if $|\phi(x)| < +\infty$ and the limit

$$\lim_{\lambda \longrightarrow 0^+} \frac{\phi(x + \lambda d) - \phi(x)}{\lambda}$$

exists, then ϕ is said to be *directionally differentiable at* x in the direction d and the directional derivative of ϕ, $\mathcal{D}\phi(x, d)$, is equal to the value of the above limit. The possibility that $\mathcal{D}\phi(x, d) = \pm\infty$ is not excluded.

Theorem 8 (Robinson (1980)). Suppose that the Karush-Kuhn-Tucker conditions hold at x^0 for $(NLP(\varepsilon^0))$ with some multiplier vectors u^0 and w^0 and that the additional *Strong Second Order Sufficient Conditions* (SSOSC) hold at x^0 with (u^0w^0). Then:

(a) x^0 is an isolated local minimum of $(NLP(\varepsilon^0))$ and the associated multiplier vectors u^0 and w^0 are unique.

(b) For ε in a neighborhood of ε^0 there exists a unique Lipschitz and once directional differentiable vector function $y(\varepsilon) = [x(\varepsilon), u(\varepsilon), w(\varepsilon)]^\top$ satisfying the (KKT) conditions and the (SSOSC) at $x(\varepsilon)$ with $[u(\varepsilon), w(\varepsilon)]$ for $(NLP(\varepsilon))$ such that $y^0(\varepsilon) = (x^0, u^0, w^0)^\top$, and $x(\epsilon)$ is a locally unique local minimum of $(NLP(\varepsilon))$ with associated unique multiplier vectors $u(\varepsilon)$ and $w(\varepsilon)$.

(c) The Linear Independence condition holds at $x(\varepsilon)$ for ε near ε^0.

(d) There exist $t > 0$ and $d > 0$ such that for all ε with $\|\varepsilon - \varepsilon^0\| < d$, it follows that

$$\left\|y(\varepsilon) - y(\varepsilon^0)\right\| \leqq t \left\|\varepsilon - \varepsilon^0\right\|.$$

We now report some sensitivity results on $(NLP(\varepsilon))$ *without the Linear Independence assumption.* The main result of this part is due to Kojima (1980) who proves his results by making extensive use of the *degree theory* of continuous maps. The same results are obtained by Fiacco and Kyparisis (1985) by means of more standard tools of advanced calculus. Let $M(x, \varepsilon)$ denote the set of multiplier vectors (u, w) such that the (KKT) conditions hold for $(NLP(\varepsilon))$ with (u, w).

Theorem 9 (Kojima (1980)). Suppose that the Karush-Kuhn-Tucker conditions hold at x^0 for $(NLP(\varepsilon^0))$ with some multiplier vectors u^0 and w^0, that the additional *General Strong Second Order Sufficient Conditions* (GSSOSC) hold at x^0 and that the (MFCQ) holds at x^0 for $(NLP(\varepsilon^0))$. Then,

(a) x^0 is an isolated local minimum of $(NLP(\varepsilon^0))$ and the set $M(x^0, \varepsilon^0)$ is nonempty, compact and convex.

(b) There are neighborhoods $N(x^0)$ of x^0 and $N(\varepsilon^0)$ of ε^0 such that for ε in $N(\varepsilon^0)$ there exists a unique continuous vector function $x(\varepsilon)$ in $N(x^0)$ satisfying the (KKT) conditions with some $[u(\varepsilon), w(\varepsilon)]^\top \in M(x(\varepsilon), \varepsilon)$ and the General Strong Second Order Sufficient Conditions such that $x(\varepsilon^0) = x^0$, and hence $x(\varepsilon)$ is the locally unique local minimum of $(NLP(\varepsilon))$ in $N(x^0)$.

(c) The (MFCQ) holds at $x(\varepsilon)$ in $N(\varepsilon^0)$.

Let us now introduce (Fiacco and Kyparisis (1985)) the *General Strict Complementarity Slackness Condition* (GSCS):

• The (GSCS) is said to hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon^0))$ if (SCS) holds at x^0 with respect to (u, w) for *every* $(u, w) \in M(x^0, w^0)$.

Fiacco and Kyparisis (1985) prove the following result.

Theorem 10. The following two sets of assumptions are equivalent:

(a) The (KKT), (SOSC), (SCS) and (LI) conditions hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon^0))$.

(b) The (KKT), (GSOSC), (GSCS) and (MFCQ) conditions hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon^0))$.

In the present section sensitivity analysis for $(NLP(\varepsilon))$ has been carried out under progressive weaker assumptions. However, it should be clear that in the absence of inequality constraints, all those sets of conditions reduce to the (KKT), (SOSC) and (LI) conditions. Additional results were obtained by other authors, e. g., Robinson (1982), under even more general assumptions than those previously adopted. In order to state his main result, we need to introduce the following definitions (see Berge (1963)).

Definition 1. A point-to-set map $\Gamma : D \longrightarrow E$, $D \subset \mathbb{R}^r$, $E \subset \mathbb{R}^s$, is called *lower semicontinuous* at $z^0 \in D$ if for any $v \in \Gamma(z^0)$ and for any open set W with $v \in W$, there exists an open set V with $z^0 \in V$ such that for all $z \in V$, $\Gamma(z) \cap W \neq \varnothing$.

Upper semicontinuity of Γ at a point $z^0 \in D$ means that for any open set $W \subset \Gamma(z^0)$ there is an open set V with $z^0 \in V$ such that for all $z \in V$, $\Gamma(z) \subset W$.

Finally, Γ is called *continuous* at $z^0 \in D$ if it is both upper and lower semicontinuous at z^0.

Theorem 11 (Robinson (1982)). Suppose that the Karush-Kuhn-Tucker conditions hold at $x^0 \in K(\varepsilon^0)$ with some multiplier vectors u^0 and w^0, that the additional General Second Order Sufficient Conditions hold at x^0 for $(NLP(\varepsilon^0))$, and that the (MFCQ) holds at x^0. Then:

(a) x^0 is an isolated local minimum of $(NLP(\varepsilon^0))$ and the set $M(x^0, \varepsilon^0)$ is compact and convex.

(b) There exist neighborhoods $N(x^0)$ of x^0 and $N(\varepsilon^0)$ of ε^0 such that if the point-to-set map $LS : N(\varepsilon^0) \longrightarrow N(x^0)$, defined by

$$LS(\varepsilon) = \left\{x \in N(x^0) : x \text{ is a local minimum of } (NLP(\varepsilon))\right\},$$

then the point-to-set map $SP(\varepsilon) \cap N(x^0)$, where

$$SP(\varepsilon) = x : \{(x, u, w),\ x \in K(\varepsilon),\ \text{is a (KKT) triplet for } (NLP(\varepsilon))\},$$

is continuous at ε^0 and for each $\varepsilon \in N(\varepsilon^0)$ one has $\varnothing \neq LS(\varepsilon) \subset SP(\varepsilon)$.

(c) The (MFCQ) holds at all points in $SP(\varepsilon) \cap N(x^0)$ for ε in $N(\varepsilon^0)$.

Robinson shows by numerical examples that under the assumptions of Theorem 11 the map LS may not be single-valued near $\varepsilon^0 = 0$ and that the inclusion of LS in SP may be strict.

4. Directional Differentiability of the Perturbed Local Minimum

This section extends the results of Theorem 8 by weakening the (LICQ) assumption. It is well-known that in absence of (LICQ) we could obtain a set of Kuhn-Tucker multipliers which is not a singleton. It is well-known (see Gauvin (1977)) that if (MFCQ) holds at x^0 for $(NLP(\varepsilon^0))$, then for each (x, ε) near (x, ε^0), the set of Kuhn-Tucker multiplier vectors $M(x, \varepsilon)$ is a *nonempty bounded polyhedral set;* more precisely, (MFCQ) is both necessary and sufficient for the said property to hold.

Let $E(x, \varepsilon)$ be the set of *extreme points* of $M(x, \varepsilon)$. For any (u, w) satisfying the (KKT) conditions, i. e. $(u, w) \in M(x, \varepsilon)$, define

$$I^+(u, w) = \{i : i \in 1 \leqq i \leqq m,\ u_i > 0\}.$$

The first results on directional differentiability of the perturbed local minimum $x(\varepsilon)$ for $(NLP(\varepsilon))$ in absence of strict complementarity slackness (SCS) assumptions were obtained by Jittorntrum (1984). Shapiro (1985) obtains interesting results on that question, weakening the (LICQ) assumption but assuming uniqueness of Karush-Kuhn-Tucker multipliers. Kyparisis

(1985) proved that the uniqueness of Karush-Kuhn-Tucker multipliers is equivalent to (SMFCQ).

Theorem 12 (Shapiro (1985)). Suppose that at x^0, local solution point of $(NLP(\varepsilon^0))$, (KKT), (SMFCQ) and (SSOSC) hold, with associated multipliers vector (u^0, w^0). Then, for ε in some neighborhood of ε^0, there exists a locally unique continuous function $x(\varepsilon)$ which is a local minimum for $(NLP(\varepsilon))$.

Moreover, $x(\varepsilon)$ is directionally differentiable at ε^0 in any direction $s \neq 0$ and its directional derivative $Dx(\varepsilon^0, s) \equiv z$ uniquely solves the following quadratic programming problem:

$$(QP): \quad \begin{cases} \min\limits_z z^\top \nabla_x^2 \mathcal{L} z + 2z^\top \nabla_{\varepsilon x}^2 \mathcal{L} s & \\ \qquad \text{subject to:} & \nabla_x g_i z + \nabla_\varepsilon g_i s = 0, \ i \in I^+(u^0, w^0), \\ & \nabla_x g_i z + \nabla_\varepsilon g_i s \leqq 0, \ i \in I(x^0, \varepsilon^0) \diagdown I^+(u^0, w^0), \\ & \nabla_x h_j z + \nabla_\varepsilon h_j s = 0, \ j = 1, ..., p, \end{cases}$$

where the functions g_i amd h_j are evaluated at (x^0, ε^0) and the function $\mathcal{L}$ at $(x^0, u^0, w^0, \varepsilon^0)$. The set $I(x^0, \varepsilon^0) \diagdown I^+(u^0, w^0)$ is given by $\{i : i \in I(x^0 \varepsilon^0), \ i \notin I^+(u^0, w^0)\}$.

The next result was obtained by Kyparisis (1990b). This author considers the possibility of multiple Karush-Kuhn-Tucker multipliers for $(NLP(\varepsilon^0))$ and employs the *constant rank constraint qualification.*

Theorem 13 (Kyparisis (1990b)). Suppose that conditions (KKT), (MFCQ), (CR) and (GSSOSC) hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon^0))$. Then. for ε in some neighborhood of ε^0, there exists a locally unique continuous local minimum $x(\varepsilon)$ of $(NLP(\varepsilon))$. Moreover, $x(\varepsilon)$ is directionally differentiable at ε^0 in any direction $s \neq 0$ and its directional derivative z uniquely solves the following quadratic program for some $(u, w) \in E(x^0, \varepsilon^0)$, where $E(x^0, \varepsilon^0)$ is the set of the extreme points of $M(x^0, \varepsilon^0)$.

$$(QP(u,w)): \quad \begin{cases} \min\limits_z z^\top \nabla_x^2 \mathcal{L} z + 2z^\top \nabla_{\varepsilon x}^2 \mathcal{L} s & \\ \qquad \text{subject to:} & \nabla_x g_i z + \nabla_\varepsilon g_i s = 0, \ i \in I^+(u, w), \\ & \nabla_x g_i z + \nabla_\varepsilon g_i s \leqq 0, \ i \in I(x^0, \varepsilon^0) \diagdown I^+(u, w), \\ & \nabla_x h_j z + \nabla_\varepsilon h_j s = 0, \ j = 1, ..., p, \end{cases}$$

where the functions g_i and h_j are evaluated at (x^0, ε^0) and the function $\mathcal{L}$ at $(x^0, u^0, w^0, \varepsilon^0)$.

Note that the assumptions (KKT), (MFCQ) and (GSSOSC) were shown by Kojima (1980) to be the weakest ones under which the perturbed solution $x(\varepsilon)$ is locally unique for $(NLP(\varepsilon))$.

A further improvement was obtained by Liu (1995b); since (WCR) is implied by (CR), which was used in Theorem 13, and (WCR) plus (MFCQ) is implied by (SMFCQ), which was used in Theorem 12, the next theorem directly extends both Theorem 12 and Theorem 13.

Theorem 14 (Liu (1995b)). Suppose that the conditions (KKT), (MFCQ), (WCR) and (GSSOSC) hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon^0))$. Then. for ε in some neighborhood of ε^0, there exists a locally unique continuous local minimum $x(\varepsilon)$ of $(NLP(\varepsilon))$. Moreover, $x(\varepsilon)$ is directionally differentiable at ε^0 in any direction $s \neq 0$ and its directional derivative z uniquely solves $(QP(u,w))$, as in Theorem 13.

The quoted paper of Liu (1995b) gives several useful informations on the assumptions assuring also the Lipschitz continuity of $x(\varepsilon)$, however, as in Theorem 13, Theorem 14 does *not* specify which multiplier $(u,w) \in E(x^0, \varepsilon^0)$ which solves $(QP(u,w))$ indeed gives the calculation of the directional derivative z of $x(\varepsilon)$. This difficulty is solved by Ralph and Dempe (1995) by making use of the *Constant Rank Condition* (CR). First we recall two basic concepts. We recall that the directional derivative of $x(\varepsilon)$ at ε^0 in the direction $d \in \mathbb{R}^n$ is defined as

$$\mathcal{D}x(\varepsilon^0, d) = \lim_{t \longrightarrow 0^+} \frac{x(\varepsilon^0 + td) - x(\varepsilon^0)}{t}$$

when this limit exists. Then $x(\varepsilon)$ is said to be *B-differentiable* at ε^0 if the above limit holds uniformly for all d of unit length. In other words, one has

$$x(\varepsilon^0 + d) = x(\varepsilon^0) + \mathcal{D}x(\varepsilon^0, d) + o(\|d\|).$$

See Robinson (1987), Ralph and Dempe (1995), Liu (1995b). The function $x(\varepsilon)$ is said to be PC^1 *near* ε^0 if it is continuous and there is a finite family of $\mathcal{C}^1$ functions $x^1(\varepsilon), ..., x^N(\varepsilon)$ defined in a neighborhood of ε^0, such that

$$x(\varepsilon) \in \left\{x^1(\varepsilon), ..., x^N(\varepsilon)\right\}$$

for each ε in that neighborhood. The PC^1 property is a kind of piecewise smoothness. It appears that piecewise differentiability is a stronger property than the B-differentiability, which in turn is stronger than directional differentiability. Moreover, PC^1-functions are locally Lipschitz continuous; see Dempe (2002).

Ralph and Dempe (1995) choose the multipliers vector (u,w) as a member of the set

$$S(\varepsilon, d) \equiv \arg\max_{(u,w)} \left\{u^\top \nabla_\varepsilon g(x,\varepsilon)d + w^\top \nabla_\varepsilon h(x,\varepsilon)d : (u,w) \in M(x,\varepsilon)\right\}.$$

The main result of Ralph and Dempe is the following theorem.

Theorem 15. Suppose that in $(NLP(\varepsilon))$ the functions f, g and h are twice continuously differentiable near (x^0, ε^0), where x^0 is a local solution of $(NLP(\varepsilon^0))$. Furthermore, suppose that the Mangasarian-Fromovitz C.Q. holds at (x^0, ε^0), the General Strong Second Order Optimality Conditions hold at (x^0, ε^0) and the (CR) condition holds at (x^0, ε^0). Then, there exist neighborhoods U of x^0 and V of ε^0 and a function $x(\cdot)$ mapping V to U such that

1) $x(\cdot)$ is continuous and for each $\varepsilon \in V$, $x(\varepsilon)$ is the unique local solution of $(NLP(\varepsilon))$ in V. More precisely:

2) $x(\cdot)$ is a PC^1-function, hence locally Lipschitz and B-differentiable;

3) the directional derivative $\mathcal{D}x(\varepsilon, d)$ is a piecewise linear function such that for each $\varepsilon \in V$, $d \in \mathbb{R}^n$ and $(u, w) \in S(\varepsilon, d)$, $\mathcal{D}x(\varepsilon, d)$ is the unique solution of the convex quadratic program

$$\begin{cases} \min\limits_z z^\top \nabla_x^2 \mathcal{L} z + 2z^\top \nabla_{\varepsilon x}^2 \mathcal{L} d & \\ \qquad \text{subject to:} & \nabla_x g_i z + \nabla_\varepsilon g_i d = 0,\ i \in I^+(u, w), \\ & \nabla_x g_i z + \nabla_\varepsilon g_i d \leqq 0,\ i \in I(x^0, \varepsilon^0) \diagdown I^+(u, w), \\ & \nabla_x h_j z + \nabla_\varepsilon h_j d = 0,\ j = 1, ..., p. \end{cases}$$

Thus, Ralph and Dempe (1995) have provided, for the case of multiplier degeneracy, a practical application of quadratic programming to compute the directional derivative $\mathcal{D}x(\varepsilon, d)$ for any d. Indeed, note that $S(\varepsilon, d)$ is nonempty because $M(x, \varepsilon)$ is nonempty and compact: the linear programming methods provide a practical way to calculate an element of $S(\varepsilon, d)$. Finally, we point out that also Ralph and Dempe, similarly to Liu (Theorem 14), observe that the condition (CR) can be weakened to the condition (WCR).

5. Differential Stability of the Optimal Value Function

We define the *optimal value function or marginal function* $f^*(\varepsilon)$ for $(NLP(\varepsilon))$ as follows

$$f^*(\varepsilon) = \begin{cases} \inf\limits_{x \in K(\varepsilon)} f(x, \varepsilon), & \text{if } K(\varepsilon) \neq \varnothing, \\ +\infty, & \text{if } K(\varepsilon) = \varnothing. \end{cases}$$

Rather good surveys of differential stability properties of the optimal value function $f^*(\varepsilon)$ are contained in the works of Fiacco (1983a,b), Fiacco and Hutzler (1982a,b), Fiacco and Ishizuka (1990).

One of the earliest characterizations of differential stability of the optimal value function of a mathematical programming problem was provided by Danskin (1966, 1967). Consider the problem

$$\min_{x \in K} f(x, \varepsilon)$$

where K is a subset of $\mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^r$.

Theorem 16 (Danskin). Let K be nonempty and compact and let f and the partial derivatives $\partial f / \partial \varepsilon_i$ be continuous. Then, at any point $\varepsilon \in \mathbb{R}^r$ and for any direction $d \in \mathbb{R}^r$ the directional derivative of f^* exists and is given by

$$\mathcal{D}f^*(\varepsilon, d) = \min_{x \in S(\varepsilon)} \nabla_\varepsilon f(x, \varepsilon) d,$$

where $S(\varepsilon) = \{x \in K : f(x, \varepsilon) = f^*(\varepsilon)\}$.

Considering the problem where also K depends on a parameter vector ε and is specified through the inequalities $g_i(x, \varepsilon) \leqq 0$, $i = 1, ..., m$, and, moreover, assuming that f and the constraints g_i are all convex, Hogan (1973a) has shown that $\mathcal{D}f^*(\varepsilon, d)$ exists and is finite for all $d \in \mathbb{R}^r$. The next theorem presents the details for this result for

$$f^*(\varepsilon) = \inf_x \{f(x, \varepsilon) : x \in M,\ g(x, \varepsilon) \leqq 0\},$$

where M is a subset of $\mathbb{R}^n$. The Lagrangian function for this problem is defined as

$$\mathcal{L}(x, u, \varepsilon) = f(x, \varepsilon) + u^\top g(x, \varepsilon).$$

For convenience and without loss of generality, we shall assume that the parameter value of interest is $\varepsilon = 0$, unless otherwise stated.

Theorem 17 (Hogan). Let M be a closed and convex set. Suppose that f and g_i are convex on M for each fixed ε and are continuously differentiable on $M \times N(0)$, where $N(0)$ is a neighborhood of $\varepsilon = 0$ in $\mathbb{R}^r$. If

$$S(0) \equiv \{x \in M : g(x, 0) \leqq 0 \text{ and } f^*(0) = f(x, 0)\}$$

is nonempty and bounded, if $f^*(0)$ is finite, and if there is a point $\bar{x} \in M$ such that $g(\bar{x}, 0) < 0$ (Slater constraint qualification), then $\mathcal{D}f^*(0, d)$ exists and is finite for all $d \in \mathbb{R}^r$ and

$$\begin{aligned} \mathcal{D}f^*(0, d) &= \min_{x \in S(0)} \max_{u \in M(x,0)} \nabla_\varepsilon \mathcal{L}(x, u, 0) d = \\ &= \min_{x \in S(0)} \max_{u \in M(x,0)} \left\{ (\nabla_\varepsilon f(x, 0) + u^\top \nabla_\varepsilon g(x, 0)) d \right\}, \end{aligned}$$

where $M(x, 0)$ is the set of optimal Lagrange multipliers for the given $x \in S(0)$.

For convex programs the set $M(x, 0)$ is the same for all $x \in S(0)$. This does *not* allow us to drop the minimization over $x \in S(0)$ in the expression for $\mathcal{D}f^*(0, d)$ however, since the quantity $\nabla_\varepsilon \mathcal{L}(x, u, 0) d$ being minimized, will generally depend on x. However, it does allow for considerable economy in the calculation of $\mathcal{D}f^*(0, d)$, since the constraint set $M(x, 0)$, associated with the inner maximization problem. does not depend on x.

There are other results on the directional differentiability of f^* : see, e. g., the papers of Bondarevsky and others (2016), Gauvin and Dubeau (1982), Fiacco and Hutzler (1982a,b), Fiacco and Ishizuka (1990), Rockafellar (1984), Gauvin and Janin (1990), Janin (1984), Gauvin (1988), Gauvin and Tolle (1977), Shapiro (1988), Wang and Zhao (1994).

The next result is a slight generalization of previous results of Gauvin and Dubeau (1982) and of Fiacco (1983b). We recall first the notion of *uniformly compact set* (see, e. g., Bank and others (1982), Berge (1963), Hogan (1973b)).

Definition 2. Let T be a metric space (e. g. a subset of $\mathbb{R}^r$) and let be given the point-to-set map $\Gamma : T \longrightarrow 2^{\mathbb{R}^n}$ (i. e. $\Gamma(\varepsilon)$ is a subset of $\mathbb{R}^n$ for each $\varepsilon \in T$). Then Γ is said to be *uniformly compact* near $\bar{\varepsilon} \in T$ if the set

$$\bigcup_{\varepsilon \in N(\bar{\varepsilon})} \Gamma(\varepsilon)$$

is bounded for some neighborhood $N(\bar{\varepsilon})$ of $\bar{\varepsilon}$.

Theorem 18 (Fiacco (1983b)). For $(NLP(\varepsilon))$ suppose that $K(\varepsilon) \neq \varnothing$ and that the constraint map (or feasible solution map) K is uniformly compact near $\varepsilon \in \mathbb{R}^r$. Suppose further that (MFCQ) holds at each $x \in S(\varepsilon)$, where

$$S(\varepsilon) = \{x \in K(\varepsilon) : f(x, \varepsilon) = f^*(\varepsilon)\}$$

is the *optimal solution map* of $(NLP(\varepsilon))$. Then f^* is locally Lipschitz near ε, and for any $d \in \mathbb{R}^r$ we have

$$\begin{aligned} &\inf_{x \in S(\varepsilon)} \min_{(u,w) \in M(x,\varepsilon)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon) d \leqq \\ &\leqq \lim_{\beta \longrightarrow 0^+} \inf \frac{1}{\beta} [f^*(\varepsilon + \beta d) - f^*(\varepsilon)] \leqq \\ &\leqq \lim_{\beta \longrightarrow 0^+} \sup \frac{1}{\beta} [f^*(\varepsilon + \beta d) - f^*(\varepsilon)] \leqq \\ &\leqq \inf_{x \in S(\varepsilon)} \max_{(u,w) \in M(x,\varepsilon)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon) d. \end{aligned}$$

Furthermore, if f and g are convex on $\mathbb{R}^n \times \{\varepsilon\}$, and if h is affine on $\mathbb{R}^n \times \{\varepsilon\}$, then f^* is directionally differentiable at ε, and

$$\mathcal{D}f^*(\varepsilon, d) = \min_{x \in S(\varepsilon)} \max_{(u,w) \in M(x,\varepsilon)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon) d, \tag{2}$$

where, under the assumptions, $M(x, \varepsilon) = M(\varepsilon)$ =constant for $x \in S(\varepsilon)$.

We note that the two linits appearing in the thesis of Theorem 18 define, respectively, the *Dini lower directional derivative* of f^* at ε in the direction d and the *Dini upper directional derivative* of f^* at ε in the direction d (right hand-side Dini directional derivatives).

There are at least two consequences of Theorem 18. Under stronger constraint qualifications than (MFCQ) at $x \in S(\varepsilon)$, for example (SMFCQ) or (LICQ), $M(x, \varepsilon)$ becomes a singleton set, say $\{u(x), w(x)\}$, so that (2) reduces to

$$\mathcal{D}f^*(\varepsilon, d) = \min_{x \in S(\varepsilon)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon) d.$$

Another special case is the jointly convex case: if f and g are jointly convex and h is jointly affine on $\mathbb{R}^n \times \mathbb{R}^r$, then for each $(u, w) \in M(x, \varepsilon)$, the quantity $\nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon)$ does not depend on $x \in S(\varepsilon)$ (see Hogan (1973a)) and hence (2) becomes

$$\mathcal{D}f^*(\varepsilon, d) = \max_{(u,w) \in M(x,\varepsilon)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon) d, \tag{3}$$

where $x \in S(\varepsilon)$. Since, in this case, f^* is convex (see Fiacco and Kyparisis (1986)), relation (3) means that, for each $x \in S(\varepsilon)$ and $(u, w) \in M(x, \varepsilon)$, $\nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon)$ is a subgradient of f^* at ε.

Besides the above quoted paper of Fiacco and Kyparisis (1986), other results on generalized convexity and concavity of the optimal value function f^* are given by Choo and Chew (1985) and by Kyparisis and Fiacco (1987).

We note that the same symbol f^* is often used (see, e. g., Fiacco (1983a)) to denote a *local* optimal value function, in many results; in this case f^* is defined as

$$f^*(\varepsilon) = f\left[x(\varepsilon), \varepsilon\right]$$

where $x(\varepsilon)$ is an isolated local minimum of $(NLP(\varepsilon))$. The following classical result was obtained by Armacost and Fiacco (1975,1978,1979) and by Fiacco (1976, 1983a).

Theorem 19. Let the assumptions of Theorem 6 be satisfied, i. e. the functions of $(NLP(\varepsilon))$ are $\mathcal{C}^2$ in (x, ε) in some neighborhood of $(x^0\varepsilon^0)$, the (KKT), (SOSC), (SCS) and (LI) conditions hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon))$. Then, in a neighborhood of ε^0, the local optimal value function $f_\ell^*(\varepsilon)$ is twice continuously differentiable as a function of ε and:

(a) $f_\ell^*(\varepsilon) = \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right].$

(b) $\nabla_\varepsilon f_\ell^*(\varepsilon) = \nabla_\varepsilon \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right] = \nabla_\varepsilon f + \sum_{i=1}^m u_i(\varepsilon)\nabla_\varepsilon g_i + \sum_{j=1}^p w_j(\varepsilon)\nabla_\varepsilon h_j =$
$= \nabla_\varepsilon f + u(\varepsilon)^\top \nabla_\varepsilon g + w(\varepsilon)^\top \nabla_\varepsilon h.$

(c)
$$\begin{aligned}
\nabla_\varepsilon^2 f_\ell^*(\varepsilon) &= \nabla_\varepsilon \left[\nabla_\varepsilon \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right]^\top\right] = \\
&= \nabla_x \left\{\nabla_\varepsilon f^\top + \textstyle\sum_{i=1}^m u_i(\varepsilon)\nabla_\varepsilon g_i + \sum_{j=1}^p w_j(\varepsilon)\nabla_\varepsilon h_j\right\} \nabla_\varepsilon x(\varepsilon) + \\
&+ \left\{\nabla_\varepsilon^2 f + \textstyle\sum_{i=1}^m u_i(\varepsilon)\nabla_\varepsilon^2 g_i + \sum_{j=1}^p w_j(\varepsilon)\nabla_\varepsilon^2 h_j\right\} + \\
&+ \left\{\textstyle\sum_{i=1}^m \nabla_\varepsilon g_i^\top \nabla_\varepsilon u_i(\varepsilon) + \sum_{j=1}^p \nabla_\varepsilon h_j^\top \nabla_\varepsilon w_j(\varepsilon)\right\} = \\
&= \nabla_{x\varepsilon}^2 \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right] \nabla_\varepsilon x(\varepsilon) + \textstyle\sum_{i=1}^m \nabla_\varepsilon g_i^\top \nabla_\varepsilon u_i(\varepsilon) + \\
&+ \textstyle\sum_{j=1}^p \nabla_\varepsilon h_j^\top \nabla_\varepsilon w_j(\varepsilon) + \nabla_\varepsilon^2 \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right] = \\
&= \nabla_\varepsilon^2 \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right] + \left[N(\varepsilon)\right]^\top \left[M(\varepsilon)\right]^{-1} N(\varepsilon),
\end{aligned}$$

where the matrices $M(\varepsilon)$ and $N(\varepsilon)$ have been previously defined.

We note that the previous results extend what in economic theory is called "envelope theorem", often stated without the appropriate assumptions which assure its validity; see, e. g., Giorgi and Zuccotti (2008-2009) and the bibliographical references quoted in this paper. The following results are obtained by Jittorntrum (1984) and by Fiacco (1983a). see also Fiacco and Liu (1995).

Theorem 20. Suppose that the Karush-Kuhn-Tucker conditions hold at $x^0 \in K(\varepsilon^0)$ for $(NLP(\varepsilon^0))$ with some multipliers vector (u^0, w^0), that the Strong Second Order Sufficient Conditions (SSOSC) and the Linear Independence Condition(LI) hold at x^0. Then, in a neighborhood of ε^0 the local optimal value function f_ℓ^* is locally Lipschitz, once continuously differentiable and twice directionally differentiable. We have

(a) $f_\ell^*(\varepsilon) = \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right];$

(b) $\nabla_\varepsilon f_\ell^*(\varepsilon) = \nabla_\varepsilon \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right];$

$(c) \quad \mathcal{D}^2 f_\ell^*(\varepsilon, d) = d^\top \nabla^2_{y\varepsilon} \mathcal{L} \mathcal{D} y(\varepsilon, d) + d^\top \nabla^2_\varepsilon \mathcal{L} d,$

where $y(\varepsilon) = [x(\varepsilon), u(\varepsilon), w(\varepsilon)]^\top$ and

$$\mathcal{D}^2 f_\ell^*(\varepsilon, d) = \lim_{\beta \longrightarrow 0^+} \frac{1}{\beta^2} \left[f_\ell^*(\varepsilon + \beta d) - f_\ell^*(\varepsilon) - \beta \mathcal{D} f_\ell^*(\varepsilon, d) \right].$$

Jittorntrum (1984) also shows that, under the previous assumptions, the function $y(\cdot)$ is locally Lipschitz near ε^0, it is directionally differentiable at ε^0 and the directional derivative $\mathcal{D}y(\varepsilon^0, d)$ is a solution of a system analogous to (1). See also Bigelow and Shapiro (1974).

Following Fiacco and Kyparisis (1985) we now give some results due to Rockafellar (1984) and to the same previous two authors; these results are concerned with the *Hadamard directional differentiability* of $f^*(\varepsilon)$. The function $f^*(\varepsilon)$ is said to be *directionally differentiable in the Hadamard sense* if the limit

$$\lim_{\substack{d' \longrightarrow d \\ t \longrightarrow 0^+}} \frac{f^*(\varepsilon + t d') - f^*(\varepsilon)}{t} = \mathcal{D}_H f^*(\varepsilon, d)$$

exists.

Before stating the next results we need the following definitions.

$$S(\varepsilon) = \{ x \in K(\varepsilon) : f(x, \varepsilon) = f^*(\varepsilon) \},$$

i.e. $S(\varepsilon)$ is the set of optimal solutions of $(NLP(\varepsilon))$, already defined.

$$M_0(x, \varepsilon) = \left\{ \begin{array}{c} (u, w) : u_i g_i(x, \varepsilon) = 0, \ u_i \geqq 0, \ i = 1, ..., m, \\ \sum_{i=1}^m u_i \nabla_x g_i(x, \varepsilon) + \sum_{j=1}^p w_j \nabla_x h_j(x, \varepsilon) = 0 \end{array} \right\}.$$

$$Y(x, \varepsilon) = \left\{ \begin{array}{c} (u, w) \in M(x, \varepsilon) : \text{ the (SOSC) hold at } x \text{ with} \\ \text{multiplier vectors } (u, w) \text{ for } (NLP(\varepsilon)) \end{array} \right\}.$$

Note that for $x \in K(\varepsilon)$, the Mangasarian-Fromovitz C.Q. holds at x if and only if $M_0(x, \varepsilon) = \{0\}$.

Theorem 21 (Rockafellar (1984)). Let $f^*(\varepsilon^0)$ be finite and an appropriate boundedness assumption be satisfied (see Rockafellar (1984)). Suppose that every optimal solution $x \in S(\varepsilon^0)$ satisfies $M_0(x, \varepsilon^0) = \{0\}$ (i. e. the (MFCQ) holds) and that

$$relint M(x, \varepsilon^0) \subset Y(x, \varepsilon^0).$$

Then f^* possesses finite one-sided directional derivatives at ε^0 in the Hadamard sense, and for every direction d it holds

$$\mathcal{D}_H f^*(\varepsilon, d) = \min_{x \in S(\varepsilon^0)} \max_{(u,w) \in M(x, \varepsilon^0)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon^0) d.$$

Under the assumptions of Theorem 9 (Kojima (1980)), the function $x(\varepsilon)$ is not locally Lipschitz near ε^0 (see Robinson (1982)) and f_ℓ^* is not in general differentiable. However, Fiacco and Kyparisis (1985), exploiting some results of Rockafellar (1984), and assuming the conditions of Theorem 9, prove the directional differentiability of $f_\ell^*(\varepsilon)$ near ε^0.

Theorem 22 (Fiacco and Kyparisis (1985)). Suppose that the Karush-Kuhn-Tucker conditions hold at x^0 for $(NLP(\varepsilon^0))$; suppose further that at x^0 the (GSSOSC) and (MFCQ) conditions hold. Then, in a neighborhood of ε^0 the local optimal value function $f_\ell^*(\varepsilon)$ has finite one-sided directional derivatives in the Jadamard sense and for every direction d it holds

$$\mathcal{D}_H f^*(\varepsilon, d) = \max_{(u,w)\in M(x,\varepsilon)} \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon^0)d.$$

6. Sensitivity Results for Right-Hand Side Perturbations

In some books on mathematical programming (e. g. Bertsekas (2016), Florenzano and Le Van (2001), Luenberger and Ye (2008), Ruszczynski (2006)), but also in some books on mathematical economics (e. g. Intriligator (1971), Novshek (1993), Silberberg and Suen (2001), Simon and Blume (1994), Sydsaeter, Hammond, Seierstad and Strom (2005)), sensitivity analysis is performed on the problem

$$(NLP_1(\varepsilon)): \qquad \begin{cases} \min f(x) \\ \text{subject to:} \quad g_i(x) \leqq \varepsilon_i, \ i = 1, ..., m, \\ \qquad\qquad\quad h_j(x) = \varepsilon_{j+m}, \ j = 1, ..., p, \end{cases}$$

where $x \in \mathbb{R}^n$.

This is due to the possibility, under suitable assumptions, not always respected in economic analysis, to interpret the optimal Lagrange or Karush-Kuhn-Tucker multipliers as a first-order measure of the sensitivity of the optimal value of the objective function with respect to right-hand changes of the constraints ("shadow prices" in economic analysis). Let us assume that all functions involved in $(NLP_1(\varepsilon))$ are twice continuously differentiable in their domain. The results of the previous sections are then easily fitted to problem $(NLP_1(\varepsilon))$, in order to obtain the well-known "Lagrange multipliers sensitivity results" and also the (not so well-known) second-order changes in the optimal value function for problem $(NLP_1(\varepsilon))$. See Fiacco (1980, 1983a). The Lagrangian function for $(NLP_1(\varepsilon))$ is:

$$\mathcal{L}_1(x, u, w, \varepsilon) = f(x) + \sum_{i=1}^{m} u_i(g_i(x) - \varepsilon_i) + \sum_{j=1}^{p} w_j(h_j(x) - \varepsilon_{j+m}).$$

Theorem 23. If

(1) The second-order sufficient conditions for a local minimum of $(NLP_1(0))$ hold at x^0 with associated multipliers vectors u^0 and w^0;

(2) The gradients $\nabla_x g_i(x^0)$, $i \in I(x^0)$, and $\nabla_x h_j(x^0)$, $j = 1, ..., p$, are linearly independent (i. e. the (LICQ) holds);

(3) It holds $u_i^0 > 0$, for $i \in I(x^0)$ (i. e. the Strict Complementarity Slackness Condition holds),

then:

(a) x^0 is an isolated local minimum point for $(NLP_1(0))$ and the associated multipliers vectors u^0 and w^0 are unique;

(b) For ε in a neighborhood of 0 there exists a unique continuously differentiable vector function $y(\varepsilon) = [x(\varepsilon), u(\varepsilon), w(\varepsilon)]^\top$ satisfying the second-order sufficient conditions for a local minimum of problem $(NLP_1(\varepsilon))$ such that $y(0) = [x^0, u^0, w^0]^\top$ and hence, $x(\varepsilon)$ is a locally unique minimum of $(NLP_1(\varepsilon))$ with associated multiplers vectors $u(\varepsilon)$ and $w(\varepsilon)$;

(c) For ε in a neighborhood of 0,

$$f^*(\varepsilon) = \mathcal{L}_1^*(\varepsilon) \equiv \mathcal{L}_1 \left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right];$$

(d) Strict complementarity and linear independence of the binding constraint gradients hold at $x(\varepsilon)$ for ε near 0;

(e) For ε in a neighborhood of 0, the gradient of the optimal value function $f_\ell^*(\varepsilon)$ is

$$\nabla_\varepsilon f_\ell^*(\varepsilon) = \left[-u(\varepsilon)^\top, -w(\varepsilon)^\top\right]; \tag{4}$$

(f) For ε in a neighborhood of 0, the Hessian matrix of the optimal value function is

$$\nabla_\varepsilon^2 f_\ell^*(\varepsilon) = -\begin{bmatrix} \nabla_\varepsilon u(\varepsilon) \\ \nabla_\varepsilon w(\varepsilon) \end{bmatrix}. \tag{5}$$

Equations (4) and (5) give us the means, at least in theory, to calculate the gradient $\nabla_\varepsilon f_\ell^*(\varepsilon)$ and the Hessian matrix $\nabla_\varepsilon^2 f_\ell^*(\varepsilon)$ of the optimal value function for problem $(NLP_1(\varepsilon))$. In practice, estimates of the values of the Karush-Kuhn-Tucker multiplers are usually available and we can thus estimate $\nabla_\varepsilon f_\ell^*(\varepsilon)$. However, we may not generally have available a functional relationship for the Karush-Kuhn-Tucker multipliers and therefore may not be able to calculate the gradient of the multipliers (the Hessian matrix of the optimal value function) directly. Armacost and Fiacco (1975, 1976) in unpublished papers and subsequently Armacost and Fiacco (1979), Fiacco (1980, 1983a) have presented a general procedure for obtaining the derivatives of a Karush-Kuhn-Tucker triple for $(NLP(\varepsilon))$, based on an observation developed by Fiacco and McCormick (1968) and Fiacco (1976). This procedure has been specified also for problem $(NLP_1(\varepsilon))$ by Armacost and Fiacco (1976) and by Fiacco (1983a). The same procedure has been applied by Fiacco (1983a) to a parametric linear programming problem of the type

$$\begin{aligned} &\min f(x, \varepsilon) = c_1\varepsilon_1 x_1 + ... + c_n\varepsilon_n x_n \\ &\text{subject to:} \quad g_i(x, \varepsilon) = g_i(x) - \varepsilon_i \geqq 0, \end{aligned}$$

$i = 1, ..., m$, where $m \geqq n$, the c_i are constants and the g_i are linear.

It is also possible to get general results (i. e. results for $(NLP(\varepsilon))$) from the right-hand side parametric problem $(NLP_1(\varepsilon))$. In fact, as noted by Fiacco (1983a) and Rockafellar (1984), any parametric problem of the form given by $(NLP(\varepsilon))$ may be formulated as an *equivalent* right hand-side parametric problem:

$$(NLP_1(\varepsilon)) : \quad \begin{cases} \min\limits_{x} f(x) & \\ \text{subject to:} & g(x) \leqq \varepsilon^1, \\ & h(x) = \varepsilon^2, \end{cases}$$

by simply redefining ε in $(NLP(\varepsilon))$ to be a variable and introducing a new parameter α such that $\varepsilon = \alpha$. This results in the problem

$$(NLP(\alpha)) : \quad \begin{cases} \min\limits_{(x,\varepsilon)} f(x, \varepsilon) & \\ \text{subject to:} & g(x, \varepsilon) \leqq 0, \\ & h(x, \varepsilon) = 0, \\ & \varepsilon = \alpha, \end{cases}$$

which is clearly equivalent to $(NLP(\varepsilon))$ and in the form of $(NLP_1(\varepsilon))$.

We now give some basic results for $(NLP_1(\varepsilon))$ in the case all functions f and g_i of this problem are *convex* and all h_j are *linear.* See, e. g., Fiacco and Kyparisis (1986), Florenzano and Le Van (2001), Horst (1984a, b), Hogan (1973a). Let us rewrite $(NLP_1(\varepsilon))$ in a more convenient form:

$$(NLP_1(\alpha, \beta)) : \quad \begin{cases} \min f(x) & \\ \text{subject to:} & g_i(x) \leqq \alpha_i, \ i = 1, ..., m, \\ & h_j(x) = \beta_j, \ j = 1, ..., p. \end{cases}$$

Let us suppose that f and g_i, $i = 1, ..., m$, are convex functions from $\mathbb{R}^n$ into $\mathbb{R} \cup \{+\infty\}$ and $h_j(x)$, $j = 1, ..., p$, are linear (affine), i. e. $h_j(x) = A_j x - b_j$, $j = 1, ..., p$, where A_j denotes a vector of $\mathbb{R}^n$ (the j-th row vector of a matrix A of order (p, n)).

Theorem 24. Under the above assumptions the following results hold:

(i) The optimal value function f^* is convex.

(ii) If α, α' are such that $\alpha_i \leqq \alpha'_i$, $\forall i = 1, ..., m$, then $f^*(\alpha, \beta) \geqq f^*(\alpha', \beta)$.

Proof.

(i) Let

$$E(\alpha, \beta) = \left\{ \begin{array}{ll} x \in \mathbb{R}^n : & g_i(x) \leqq \alpha_i, \ i = 1, ..., m; \\ & h_j(x) = \beta_j, \ j = 1, ..., p \end{array} \right\}.$$

One can easily check that for $\lambda \in [0, 1]$ we have the inclusion

$$\lambda E(\alpha, \beta) + (1 - \lambda) E(\alpha', \beta') \subset E(\alpha(\lambda), \beta(\lambda))$$

where $\alpha(\lambda) = \lambda\alpha + (1-\lambda)\alpha'$ and $\beta(\lambda) = \lambda\beta + (1-\lambda)\beta'$. Consequently, if $x \in E(\alpha,\beta)$, $x' \in E(\alpha',\beta')$, $\lambda \in [0,1)$, then

$$f^*(\alpha(\lambda),\beta(\lambda)) \leqq f(\lambda x + (1-\lambda)x') \leqq \lambda f(x) + (1-\lambda)f(x').$$

Hence

$$f^*(\alpha(\lambda),\beta(\lambda)) \leqq \lambda f^*(\alpha,\beta) + (1-\lambda)f^*(\alpha',\beta').$$

(ii) Since $E(\alpha,\beta) \subset E(\alpha',\beta)$ if $\alpha_i \leqq \alpha_i'$, $\forall i = 1,...,m$, one has

$$f^*(\alpha,\beta) = \inf_{x\in E(\alpha,\beta)} f(x) \geqq \inf_{x\in E(\alpha',\beta)} f(x) = f^*(\alpha',\beta). \qquad \square$$

We recall (see Florenzano and Le Van (2001) and the classical book of Rockafellar (1970)) that if f is a convex function on $\mathbb{R}^n$, the element x^* of $\mathbb{R}^n$ is called a *subgradient* of f at x if

$$f(z) \geqq f(x) + x^*(z-x), \ \forall z \in \mathbb{R}^n.$$

The set of subgradients of f at x is called *subdifferential* of f at x and denoted by $\partial f(x)$. If $\partial f(x) \neq \varnothing$ we say that f is *subdifferentiable* at x. The following proposition is basic to clarify the links between differentiable convex functions and subdifferentiable convex functions.

Theorem 25. Let f be a convex function on $\mathbb{R}^n$. Then f is differentiable at x if and only if $\partial f(x) = \{x^*\}$. In this case $\nabla f(x) = x^*$.

The following two propositions are proved, e. g., in Dhara and Dutta (2012) and in Florenzano and Le Van (2001). With reference to $(NLP_1(\alpha,\beta))$, it is convenient to assume, at least for the original non perturbed problem $(NLP_1(0,0))$, that the same admits a solution and that the Slater constraint qualification is verified for the said problem:

- The vectors A_j, $j = 1,...,p$, are linearly independent (i. e. the matrix A has rank p) and there exists $\bar{x}$ such that $A\bar{x} = b$ and $g_i(\bar{x}) < 0$, for all $i = 1,...,m$.

This assumption assures that the set of multipliers $M(x^0)$ for $(NLP_1(0,0))$ is a nonempty compact set of $\mathbb{R}^m_+ \times \mathbb{R}^p$ and that the feasible set of $(NLP_1(\alpha,\beta))$ is not empty for (α,β) in a neighborhood of $(0,0)$.

Theorem 26. If x^0 is a solution of $(NLP_1(0,0))$, then:

(i) The set of multipliers at x^0 is equal to $-\partial f^*(0,0)$.

(ii) Further, if the Slater constraint qualification holds for $(NLP_1(0,0))$, then f^* is continuous at the origin and the set of multipliers at x^0 is a convex and compact set.

We remark that no differentiability assumption was made in Theorem 26; hence the terms "set of multipliers" can be referred to the set of multipliers for a saddle point of the Lagrangian function of $(NLP_1(\alpha,\beta))$ or for optimality conditions expressed in terms of subdifferentials. Florenzano and Le Van (2001) give the following result, as a corollary to Theorem 26.

Theorem 27. Let x^0 be a solution of $(NLP_1(0,0))$. Then the optimal value function f^* is differentiable at $(0,0)$ if and only if there exists a unique multipliers vector (u,w) at x^0. In this case $\nabla f^*(0,0) = \left[-u^\top, -w^\top\right]$.

We remark that the multipliers vector (u,w) will be unique if the functions of $(NLP_1(\alpha,\beta))$ are differentiable and the gradients $\nabla g_i(x^0)$, $i \in I(x^0)$, $\nabla h_j(x^0)$, $j = 1,...,p$, are linearly independent.

If the multipliers vector is not unique, the results of the above theorem have to be modified. The following result, expressed perhaps in a more convenient form than the previous result, can be found, e. g., in Hiriart-Urruty (2008). We rewrite the convex problem $(NLP_1(\alpha,\beta))$ in the form:

$$(P(\alpha,\beta)): \quad \begin{cases} \min f(x) \\ \text{subject to:} \quad g_i(x) \leqq \alpha_i, \ i = 1,...,m, \\ \qquad\qquad\quad A_j(x) - b_j = \beta_j, \ j = 1,...,p, \end{cases}$$

where A_j denotes the j-th row of the matrix A of order (p,n). Besides the convexity of f and every g_i, $i = 1,...,m$, it is assumed that $(P(0,0))$, i. e. the non perturbed problem, admits a solution and that $(P(0,0))$ verifies the Slater constraint qualification at a feasible point $\bar{x}$.

Theorem 28 (Hiriart-Urruty (2008)). Under the above assumptions there exists the directional derivative of f^* at $(0,0)$ in the direction (α,β) of $\mathbb{R}^m_+ \times \mathbb{R}^p$ and is given by

$$\max_{(u,w)\in M(x^0)} \left\{(-u^\top\alpha) + (-w^\top\beta)\right\}.$$

This shows that the graph of f^* is not necessarily smooth at $(0,0)$ when the set $M(x^0)$ is not a singleton: when $M(x^0)$ is a singleton, i. e. $M(x^0) = \{(u,w)\}$, then f^* is differentiable at $(0,0)$ and we have

$$\nabla f^*(0,0) = \left[-u^\top, -w^\top\right].$$

These questions have been treated also by Horst (1984a,b), with reference to a convex optimization problem with no equality constraints. This author considers the problem

$$\begin{cases} \max f(x) \\ g_i(x) \leqq b_i, \quad i = 1,...,m, \end{cases}$$

where $(-f)$ and each g_i are real-valued convex and differentiable functions defined on $\mathbb{R}^n$.

We denote by $K(b)$ the feasible set of the said problem and by $f^*(b)$ its optimal value function:

$$f^*(b) = \begin{cases} \sup\limits_{x\in K(b)} f(x), & \text{if } K(b) \neq \varnothing, \\ -\infty, & \text{if } K(b) = \varnothing. \end{cases}$$

Theorem 29 (Horst). Assume that in the above problem the Slater constraint qualification is satisfied: there exists $\bar{x} \in \mathbb{R}^n$ such that $g_i(\bar{x}) < b_i$, $i = 1,...,m$. Moreover, let the set $U(b)$ of optimal solutions of the associated dual problem

$$\min_{u\in\mathbb{R}^m_+} \left\{ \sup_{x\in\mathbb{R}^n} \left(f(x) + u^\top(b - g(x))\right) \right\}$$

be compact. Then it holds

$$\frac{\partial f^*(b)}{\partial b_i^+} = \min_{u\in U(b)} u_i, \quad i = 1, ..., m,$$
$$\frac{\partial f^*(b)}{\partial b_i^-} = \max_{u\in U(b)} u_i, \quad i = 1, ..., m,$$

where $\frac{\partial f^*(b)}{\partial b_i^+}$ denotes the right-hand side partial derivative and $\frac{\partial f^*(b)}{\partial b_i^-}$ denotes the left-hand side partial derivative of $f^*(b)$. The compactness of the set $U(b)$ of optimal dual solutions is assumed to assure that the minima and maxima in the above relations exist. It could be replaced by: "suppose that

$$\min_{u\in U(b)} u_i \quad \text{and} \quad \max_{u\in U(b)} u_i$$

exist". Note that this will always be the case if we have only finitely many dual solutions.

From this result it is then possible to obtain the classical limitations, for the maximization problem considered,

$$\frac{\partial f^*(b)}{\partial b_i^+} \leqq u_i \leqq \frac{\partial f^*(b)}{\partial b_i^-}, \quad i = 1, ..., m.$$

See, e. g., Balinski and Baumol (1968), Gale (1967), Takayama (1977, 1985), Uzawa (1958).

For general sensitivity results on the optimal value function$f^*(\alpha, \beta)$ when the related problem is *not* convex, the reader is referred to Gauvin (1979,1980), Gauvin and Tolle (1977), Gauvin and Janin (1990), Geoffrion (1971) and the other authors quoted in this section.

7. Notes on Sensitivity for Linear Programming Problems

The literature concerning stability, sensitivity and post-optimal analysis in linear programming problems is rather extensive: almost all good books on linear programming treat these subjects. For (non exhaustive) surveys and interesting papers the reader is referred to Akgül (1984), Aucamp and Steinberg (1982), Gal (1984, 1979, 1986), Gauvin (1980a,b, 1995b), Gauvin (2001), Horst (1984a,b), Jansen and others (1997), Peterson (1970), Ward and Wendell (1990), Williams (1963). A good treatment of sensitivity properties of the optimal value function in linear programming problems (with no references to algorithmic questions), is contained in the book of Achmanov (1984). This author takes into consideration the primal problem

$$(P): \quad \begin{cases} \max c^\top x \\ Ax \leqq b \end{cases}$$

and its associated dual problem

$$(D): \quad \begin{cases} \min b^\top p \\ A^\top p = c, \ p \geqq 0, \end{cases}$$

where A is of order (m, n), $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $p \in \mathbb{R}^m$. X^0 and Y^0 are, respectively, the set of optimal solutions of (P) and (D). It is convenient to assume $X^0 \neq \varnothing$, $Y^0 \neq \varnothing$.

We recall some basic facts about the primal problem and its dual problem. If we denote by X and y, respectively, the feasible set of (P) and (D), it holds:

(i) On the basic assumption $X \neq \varnothing$, the primal problem (P) has a solution if and only if $Y \neq \varnothing$.

(ii) On the basic assumption $Y \neq \varnothing$, the dual problem (D) has a solution if and only if $X \neq \varnothing$.

With $f^*(b)$ we denote the optimal value function of (P) and we shall write $f^*(b, c)$ if we want to put into evidence the dependence of $f^*(\cdot)$ from vector c.

Let us consider the sets

$$B = \{z \in \mathbb{R}^m : \text{ there exists } x \in \mathbb{R}^n \text{ such that } z \geqq Ax\},$$

$$C = \left\{w \in \mathbb{R}^n : \text{ there exists } p \in \mathbb{R}^m, \ p \geqq 0, \text{ such that } w = A^\top p\right\}.$$

The duality theorem for linear programs inplies that the function $f^*(b, c)$ is defined on $B \times C = \{(z, w) : z \in B, w \in C\}$.

Theorem 30. The function $f^*(b, c)$ is positively homogeneous of first degree in both variables b and c, i. e.

$$f^*(\lambda b, c) = \lambda f^*(b, c); \quad f^*(b, \lambda c) = \lambda f^*(b, c), \ \forall \lambda > 0.$$

Theorem 31. The function $f^*(b, c)$ is concave with respect to b and convex with respect to c.

If we consider the usual function $f^*(b)$, the two previous theorems become, respectively:

Theorem 30 bis. The function $f^*(b)$ is positively homogeneous of first degree, i. e. $f^*(b) = \lambda f^*(b), \ \forall \lambda > 0$.

Theorem 31 bis. The function $f^*(b)$ is concave on its domain B (convex set).

An important property of $f^*(b)$ is its continuity.

Theorem 32. The function $f^*(b)$ is continuous on its domain B.

As for what concerns sensitivity results on $f^*(b)$, we have the following statement (Achmanov) which characterizes the dual variables as "shadow prices".

Theorem 33. If the dual problem (D) has a *unique* solution p^*, then the optimal value function $f^*(\cdot)$ of (P) is differentiable at the point b and we have $\partial f^*(b)/\partial b_i = p_i^*$, $i = 1, ..., m$, i. e.

$$\nabla_b f^*(b) = (p^*)^\top.$$

The conditions which assure uniqueness of solutions in a linear programming problem are studied by Mangasarian (1979).

Following again Achmanov, we now see the differentiability properties of the function $f^*(b)$ when the dual (D) has not a unique solution. When the primal solution is "degenerate" there may be multiple dual optimal solutions and in this case the thesis of Theorem 33 obviously cannot hold. Let us consider the set of the *extreme points (or vertices)* of the set Y^0, the set of solutions of (P), which is a polyhedral convex set, and suppose that this set is formed by k elements: $\left\{p^1, p^2, ..., p^k\right\}$. Let us consider the following convex cones of $\mathbb{R}^m$

$$K_r = \left\{s \in \mathbb{R}^m : sp^r - sp^i \leqq 0, \quad i = 1, 2, ..., k\right\},$$

for every $r = 1, 2, ..., k$.

Obviously $K_r \neq \varnothing : b \in K_r$, $r = 1, 2, ..., k$, as $bp^r = bp^i$.

Theorem 34. Let be $s \in \mathbb{R}^m$, $\|s\| = 1$, a vector defining a direction at the point b, $s \in K_r$. Then, the number sp^r is the directional derivative of the function $f^*(\cdot)$ at b in the direction s.

Remark 2. If we suppose that at least a dual solution p^* verifies the property $p_i^* = 0$ for an index i, we can deduce that if the i-th component b_i of vector b is modified, and the other components are not modified (they are fixed), then the maximum value $f^*(b)$ does not change.

A result similar to Theorem 34 of Achmanov is given by Horst (1984a,b) and by Gauvin (1980a). We report the version of Horst. This author considers the classical linear programming problem with nonnegative variables

$$(PL): \quad \begin{cases} \max c^\top x \\ Ax \leqq b \\ x \geqq 0, \end{cases}$$

where A is of order (m, n), $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and its dual

$$(DL): \quad \begin{cases} \min b^\top u \\ A^\top u \geqq c \\ u \geqq 0. \end{cases}$$

Let us denote by $U(b)$ the set of optimal solutions of the dual linear problem (DL). Then the results of the previous Theorem 29 apply to (PL) and (DL). More precisely, we have the following theorem.

Theorem 35 (Horst)**.** Let the set $U(b)$ be nonempty and compact. Then we have

$$\frac{\partial f^*(b)}{\partial b_i^+} = \min_{u \in U(b)} u_i; \qquad \frac{\partial f^*(b)}{\partial b_i^-} \max_{u \in U(b)} u_i,$$

where $\frac{\partial f^*(b)}{\partial b_i^+}$ denotes the right-side partial derivative of f^*, $\frac{\partial f^*(b)}{\partial b_i^-}$ denotes the left-side partial derivative and u_i denotes the i-th component of u.

Let the set $U(b)$ be compact with k extreme point solutions $\bar{u}^j$, $1 \leqq j \leqq k$. Then it is well known that $U(b)$ consists of all convex combinations of the extreme point solutions $\bar{u}^j$, that is

$$U(b) = \left\{u \in \mathbb{R}^m : u = \sum_{j=1}^{k} \lambda_j \bar{u}^j, \quad \sum_{j=1}^{k} \lambda_j = 1, \ \lambda_j \geqq 0, \ j = 1, ..., k\right\}.$$

This last one is a compact set. Moreover, in this case the quantities

$$\min_{u \in U(b)} ; \quad \max_{u \in U(b)} u_i$$

which appear in Theorem 35 can be replaced by

$$\min_{1 \leqq j \leqq k} \bar{u}_i^j ; \quad \max_{1 \leqq j \leqq k} \bar{u}_i^j ,$$

since $u_i = u^\top e^i$, where e^i denotes the i-th unit vector, is a linear function that attains its minimum and maximum at one of the extreme points $\bar{u}^j$ of $U(b)$. See also Aucamp and Steinberg (1979, 1982),

From the above results it is easy to draw the inequalities (see, e. g., Nikaido (1968))

$$\frac{\partial f^*(b)}{\partial b_i^+} = \min_{u \in U(b)} u_i \leqq u_i \leqq \max_{u \in U(b)} u_i = \frac{\partial f^*(b)}{\partial b_i^-}, \quad i = 1, ..., m.$$

A paper which applies to a nonlinear programming problem with linear constraints, the results obtained by Fiacco (1983a) and other authors for the general nonlinear parametric programming problem, is due to Kyparisis (1987a). In particular, this author takes into consideration the problem

$$\begin{cases} \min\limits_x f(x, v) & \\ \quad \text{s. t.} & Ax - b(v) \geqq 0 \\ & Dx - c(v) = 0, \end{cases}$$

where $v \in \mathbb{R}^r$ is the perturbation parameters vector, $f : \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}$ is of class $\mathcal{C}^2$ in (x, v), $b : \mathbb{R}^r \longrightarrow \mathbb{R}^m$, $c : \mathbb{R}^r \longrightarrow \mathbb{R}^p$, A is a matrix of order (p, n). Then this author takes into consideration the problem with fixed linear constraints

$$P(v^*) : \quad \begin{cases} \min\limits_x f(x, v) & \\ \quad \text{s. t.} & Ax - b \geqq 0 \\ & Dx - c = 0. \end{cases}$$

With reference to $P(v^*)$ Kyparisis proves the following results. Suppose that the (KKT) conditions hold at x^* for $P(v^*)$ and define the set

$$C = \{z : \nabla_x f(x^*, v^*) z = 0, \ A^* z \geqq 0, \ Dz = 0\},$$

where A^* is the submatrix of A which comprises rows A_k, $k \in I^*$, $I^* = \{k : A_k x^* = b_k\}$.

Theorem 36 (Kyparisis). Suppose that (KKT) conditions hold at x^* for $P(v^*)$ with some (u^*, v^*), and that the Second Order Sufficient Conditions hold at x^*. Then, there are neighborhoods U of x^* and V of v^*, and a Lipschitz continuous function $x : V \longrightarrow U$, such that, for each $v \in V$, $x(v)$ is the unique solution in U of $P(v^*)$. Moreover, the one-sided directional derivative of $x(v)$ at v^* exists for any direction $s \neq 0$ and is given by $\bar{x}$ which uniquely solves the following system:

$$\left[\nabla_x^2 f(x^*, v^*)\bar{x} + \nabla_{vx}^2 f(x^*, v^*)s\right]^\top \bar{x} = 0,$$

$$\nabla_x^2 f(x^*, v^*)\bar{x} + \nabla_{vx}^2 f(x^*, v^*)s \in C^+,$$

$\bar{x} \in C$, where $C^+ = \left\{z : z^\top x \geqq 0,\ \forall x \in C\right\}$ is the positive dual cone of C.

8. Notes on Sensitivity for Variational Inequality Problems

Parallel to the stability and sensitivity developments in nonlinear programming problems, stability and sensitivity studies for variational inequalities have been expanding rapidly. The results obtained before 1990 heve been surveyed by Kyparisis (1990a). Other works on this subject are Gowda and Pang (1994), King and Rockafellar (1992), Kyparisis (1987b, 1990b, 1992), Mordukhovich (1994), Pang (1993), Liu (1995a,b), Qiu and Magnanti (1992), Tobin (1986).

Variational inequalities were originally introduced in infinite-dimensional spaces (see, e. g., Kinderlehrer and Stampacchia (1980), Baiocchi and Capelo (1984)). However, subsequently finite-dimensional versions have been found to have applications to transportation planning, regional science, socio-economic analysis, energy modeling, game theory, etc. An excellent survey of finite-dimensional variational inequalities is given by Harker and Pang (1990). Also the two volumes of Facchinei and Pang (2003) are a good reference work on variational inequalities. In the present section we shall follow mainly the paper of Crespi and Giorgi (2002).

We recall that the general parametric variational inequality problem is defined as:

- Find $x^* \in K(\varepsilon)$ such that

$$F(x^*, \varepsilon)(x - x^*) \geqq 0$$

for all $x \in K(\varepsilon)$, where F is a mapping from $\mathbb{R}^n \times \mathbb{R}^r$ to $\mathbb{R}^n$, $K(\cdot)$ is a feasible point-to-set map from $\mathbb{R}^r$ to $\mathbb{R}^n$ (i. e. it assigns to each vector ε the feasible set $K(\varepsilon)$), and $\varepsilon \in \mathbb{R}^r$ is the parameters vector. Direct analysis of this problem is in general difficult, unless the feasible set $K(\varepsilon)$ has a specific functional form. For example, $K(\varepsilon)$ is a nonempty closed convex subset of $\mathbb{R}^n$ or $K(\varepsilon)$ is specified in a functional form. We shall consider this last case; more precisely, we are concerned with the case where $K(\varepsilon)$ is given by

$$K(\varepsilon) = \{x \in \mathbb{R}^n : g(x, \varepsilon) \leqq 0,\ h(x, \varepsilon) = 0\},$$

where g and h are mappings from $\mathbb{R}^n \times \mathbb{R}^r$ into $\mathbb{R}^m$ and $\mathbb{R}^p$, respectively, and $\varepsilon \in \mathbb{R}^r$ is the vector of perturbation parameters. We shall call $(VI(\varepsilon))$ this last variational inequality problems.

Since $(VI(\varepsilon)$ and $(NLP(\varepsilon))$ are traditionally linked together, it has been quite common to translate classical results of the latter problem to the former one. For instance, we shall recall that a function $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ can be seen as the gradient of some function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, i. e. $F(x) = \nabla f(x)$, provided that the Jacobian matrix $\nabla f(x)$ is symmetric for all $x \in dom(f)$. If that is the case (the variational inequality is then said to be *integrable*), we have that the Karush-Kuhn-Tucker type conditions for the variational inequality coincide with the ordinary Karush-Kuhn-Tucker conditions for the optimization problem.

We assume that the functions F, g and h are $\mathcal{C}^2$ in both variables (x, ε) at least in a neighborhood of (x^*, ε^*), local solution of the unperturbed problem $(VI(\varepsilon^*))$. For $(VI(\varepsilon))$ we shall introduce a set of KKT type conditions, by means of the following Lagrangian function (see Kyparisis (1987b, 1990a)):

$$\mathcal{L}_D(x, \lambda, \mu, \varepsilon) = F(x, \varepsilon) + \nabla_x g(x, \varepsilon) + \mu \nabla h_x(x, \varepsilon).$$

Let x^* be a solution or also a local solution of $(VI(\varepsilon^*))$; we say that the *generalized Karush-Kuhn-Tucker* (GKKT) conditions hold at x^* if

$$\mathcal{L}_D(x^*, \lambda, \mu, \varepsilon^*) = 0,$$

$$\lambda_i \geqq 0,\ \lambda_i g_i(x^*, \varepsilon^*) = 0,\ i = 1, ..., m,$$
$$g_i(x^*, \varepsilon^*) \leqq 0,\ i = 1, ..., m;\ h_j(x^*, \varepsilon^*) = 0,\ j = 1, ..., p.$$

It is well-known that the (GKKT) conditions hold at the optimal point x^* for $(VI(\varepsilon^*))$ if some constraint qualification holds: e. g. the gradients $\nabla g_i(x^*)$, $i \in I(x^*)$, and $\nabla h_j(x^*)$, $j = 1, ..., p$, are linearly independent.

We call a point x^* that satisfies (GKKT) conditions with some λ and μ a *stationary point* for $(VI(\varepsilon))$. It is also known (Tobin (1986)) that if for $\varepsilon = \varepsilon^*$, the functions g_i, $i = 1, ..., m$, are convex in x and the functions $h_j(x)$, $j = 1, ..., p$, are affine in x, then any point x^* which satisfies (GKKT) (i. e. x^* is a stationary point for $(VI(\varepsilon))$), is a solution of $(VI(\varepsilon^*))$.

Conditions (LI), (MFCQ), (SMFCQ), (CR) are applied to $(VI(\varepsilon^*))$ without any change. Some second order conditions for $(VI(\varepsilon))$ are introduced.

Let x^* be a stationary point for $(VI(\varepsilon^*))$.

- The *second order conditions* (SOC) hold at x^* with (λ, μ) if

$$z^\top \nabla_x \mathcal{L}_D(x^*, \lambda, \mu, \varepsilon^*) z > 0$$

for all $z \neq 0$ such that

$$\begin{aligned} \nabla g_i(x^*, \varepsilon^*) z &\leqq 0,\ \forall i \in I(x^*), \\ \nabla g_i(x^*, \varepsilon^*) z &= 0,\ \forall i \in I^+(\lambda, \mu), \\ \nabla h_j(x^*, \varepsilon^*) z &= 0,\ \forall j = 1, ..., p, \end{aligned}$$

where

$$I^+(\lambda, \mu) = \{i : 1 \leqq i \leqq m,\ \lambda_i > 0\}$$

for each pair of multipliers (λ, μ).

• The *strong second order conditions* (SSOC) hold at x^* with (λ, μ) if

$$z^\top \nabla_x \mathcal{L}_D(x^*, \lambda, \mu, \varepsilon^*) z > 0$$

for all $z \neq 0$ such that

$$\begin{aligned} \nabla g_i(x^*, \varepsilon^*) z &= 0, \ \forall i \in I^+(\lambda, \mu), \\ \nabla h_j(x^*, \varepsilon^*) z &= 0, \ \forall j = 1, ..., p. \end{aligned}$$

• The *modified strong second order conditions* (MSSOC) hold at x^* with (λ, μ) if

$$z^\top \nabla_x F(x^*, \varepsilon^*) z > 0$$

for all vectors $z \neq 0$ as in (SSOC).

• The *general modified strong second order conditions* (GMSSOC) hold at x^* if (MSSOC) hold at x^* with (λ, μ), *for all* $(\lambda, \mu) \in M(x^*, \varepsilon^*)$, where $M(x, \varepsilon)$ denotes the set of multipliers (λ, μ) which satisfy (GKKT) at (x, ε).

We are now ready to introduce some counterparts of sensitivity results for $(NLP(\varepsilon))$ to variational inequality problem $(VI(\varepsilon))$.

Theorem 37 (Tobin (1986), Kyparisis (1987b), Qiu and Magnanti (1992)). For $(VI(\varepsilon^*))$ suppose that (GKKT) conditions hold at x^* with (λ, μ); suppose further that the (LI) constraint qualification and the Strict Complementary Slackness Condition (SCS) hold at x^*. Finally, suppose that the (SOC) hold at x^* with (λ, μ). Then:

a) For ε in a neighborhood of ε^* there exists a locally unique once continuously differentiable vector function $y(\varepsilon) = [x(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon)]^\top$ satisfying (GKKT), (SOC), (SC) and (LI) conditions at $x(\varepsilon)$ with $(\lambda(\varepsilon), \mu(\varepsilon))$ for $(VI(\varepsilon))$ such that $y(\varepsilon^*) = [x^*, \lambda, \mu)]^\top$.

b) Under additional assumption that, for any fixed ε, all g_i are convex in x and all h_j are affine in x, $x(\varepsilon)$ is also a locally unique solution to $(VI(\varepsilon))$ with associated unique multipliers vectors $\lambda(\varepsilon)$ and $\mu(\varepsilon)$.

Theorem 38 (Kyparisis (1987b)). Suppose that x^* verifies (GKKT), (SSOC), and (LI) conditions with multipliers (λ, μ) for $(VI(\varepsilon^*))$. Then:

(a) For ε in a neighborhood of ε^*, there exists a locally unique continuous vector function $y(\varepsilon) = [x(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon)]^\top$ satisfying (GKKT) as well as (SSOC) and (LI) conditions for $(VI(\varepsilon))$ and such that $x(\varepsilon) \in K(\varepsilon)$ and $y(\varepsilon^*) = [x^*, \lambda, \mu)]^\top$.

(b) There exist $c > 0$ and $d > 0$ such that for all ε with $\|\varepsilon - \varepsilon^*\| < d$ it follows

$$\|y(\varepsilon) - y(\varepsilon^*)\| \leqq c \|\varepsilon - \varepsilon^*\| .$$

(c) Under additional assumptions that for any fixed ε, all g_i are convex in x and all h_j are affine in x, $x(\varepsilon)$ is also a locally unique solution to $(VI(\varepsilon))$ with associated unique multipliers vectors $\lambda(\varepsilon)$ and $\mu(\varepsilon)$.

Theorem 39 (Kyparisis (1990)). Let (GKKT), (MFCQ), (CR) and (GMSSOC) conditions hold at x^*. Moreover, assume for any fixed ε, all g_i, $i = 1, ..., m$, are convex functions in x and all h_j, $j = 1, ..., p$, are affine functions in x. Then, for any ε in a suitable neighborhood of ε^*, there exists (locally) a unique solution $x(\varepsilon)$ of $(VI(\varepsilon))$ and this solution is a continuous function of ε. Moreover, $x(\varepsilon)$ is directionally differentiable at ε^* with respect to any direction $d \neq 0$ and its directional derivative $\mathcal{D}x(\varepsilon^*, d)$ uniquely solves, for some (λ, μ) in the set of extreme points of $M(x^*, \varepsilon^*)$, the following linear variational inequality:

Find $z^* \in LK(\lambda, \mu)$ so that $\forall z \in LK(\lambda, \mu)$,

$$[\nabla_x \mathcal{L}_D(x^*, \lambda, \mu, \varepsilon^*) z^* + \nabla_\varepsilon \mathcal{L}_D(x^*, \lambda, \mu, \varepsilon^*) d] (z - z^*) \geqq 0 \tag{6}$$

where

$$LK(\lambda, \mu) = \left\{ \begin{array}{l} z : \nabla_x g_i(x^*, \varepsilon^*) z + \nabla_\varepsilon g_i(x^*, \varepsilon^*) d = 0, \ i \in I^+(\lambda, \mu), \\ \nabla_x g_i(x^*, \varepsilon^*) z + \nabla_\varepsilon g_i(x^*, \varepsilon^*) d \geqq 0, \ i \in I(x^*, \varepsilon^*) \diagdown I^+(\lambda, \mu), \\ \nabla_x h_j(x^*, \varepsilon^*) z + \nabla_\varepsilon h_j(x^*, \varepsilon^*) d = 0, \ j = 1, ..., p. \end{array} \right\}.$$

Crespi and Giorgi (2002) have improved the previous result for a special case.

Theorem 40. Let the assumptions of Theorem 39 hold and let $f : \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}^n$ be convex and such that $F = \nabla f$. Then the conclusions of Theorem 39 still hold true and, moreover:

1. $x(\cdot)$ is (locally) $\mathcal{PC}^1$ and then locally Lipschitz and B-differentiable.
2. $\mathcal{D}x(\varepsilon^*, d)$ uniquely solves (6) for all $(\lambda, \mu) \in S(\varepsilon^*, d)$, where

$$S(\varepsilon^*, d) \equiv \arg\max_{\lambda, \mu} \{\lambda \nabla_x g(x^*, \varepsilon^*) d + \mu \nabla_x h(x^*, \varepsilon^*) d \mid (\lambda, \mu) \in M(x^*, \varepsilon^*)\}.$$

For the proof see Crespi and Giorgi (2002).

A similar result can be proved in a more general setting. First, we note that (6) can be seen as the first order condition of the following quadratic programming problem:

$$\min_{z \in LK(\lambda, \mu)} \frac{1}{2} z^\top \nabla_x \mathcal{L}_D(x, \lambda, \mu, \varepsilon) z + z^\top \nabla_\varepsilon \mathcal{L}_D(x, \lambda, \mu, \varepsilon) d.$$

Define

$$\Phi(x, \varepsilon, y) = \frac{1}{2} F(x, \varepsilon)(x - y) + \frac{1}{2} F(y, \varepsilon)(x - y),$$

for $y \in K(\varepsilon)$.

Definition 3. A function $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is said to be *monotone* if

$$(F(x) - F(y))(x - y) \geqq 0, \quad \forall x, y \in dom(F).$$

Theorem 41 (Crespi and Giorgi (2002)). Let F be continuous and monotone in x^*. Then x^* solves $(VI(\varepsilon^*))$ if and only if $x^* \in \arg\min_{x \in K(\varepsilon^*)} \Phi(x, \varepsilon^*, x^*)$.

For the proof see Crespi and Giorgi (2002).

Finally, we quote the following results due to Crespi and Giorgi (2002) and which may be considered as an extension to variational inequalities of similar results which hold for $(NLP(\varepsilon))$.

Theorem 42. Under the same assumptions of Theorem 39 and assuming F monotone, we have, besides the thesis of Theorem 39, that:

1. $x(\cdot)$ is (locally) $\mathcal{PC}^1$ and then locally Lipschitz and B-differentiable.
2. $\mathcal{D}x(\varepsilon^*, d)$ uniquely solves (6) for all $(\lambda, \mu) \in S(\varepsilon^*, d)$.

9. Estimates of Sensitivity Results for (NLP(ε)) by Means of Penalty Functions

In this final section we summarize the main results given by Fiacco (1976, 1983a) and by Armacost and Fiacco (1975, 1978, 1979) on the use of penalty methods in order to estimate some sensitivity results for a parametric nonlinear programming problem $(NLP(\varepsilon))$. The above quoted authors consider $(NLP(\varepsilon))$ with inequalities of the form $g_i(x, \varepsilon) \geqq 0$, $i = 1, ..., m$, i. e. the problem

$$(NLP(\varepsilon)) : \quad \begin{cases} \min f(x, \varepsilon) \\ \text{subject to:} \quad g_i(x, \varepsilon) \geqq 0, \ i = 1, ..., m, \\ \qquad\qquad\qquad h_j(x, \varepsilon) = 0, \ j = 1, ..., p. \end{cases}$$

In the present section we shall consider this form of $(NLP(\varepsilon))$. For this problem the Lagrange function is defined as

$$\mathcal{L}(x, u, w, \varepsilon) = f(x, \varepsilon) - \sum_{i=1}^{m} u_i g_i(x, \varepsilon) + \sum_{j=1}^{p} w_j h_j(x, \varepsilon).$$

For the reader's convenience we recall the assumptions made in order to state Theorem 6, assumptions that we impose also in the present section.

A1) The functions defining problem $(NLP(\varepsilon))$ are twice continuously differentiable in (x, ε) in a neighborhood of $(x^*, 0)$.

A2) The second order sufficient conditions for a local minimum of problem $(NLP(0))$ hold at x^* with associated Lagrange-Kuhn-Tucker multipliers u^* and w^*.

A3) The gradients $\nabla_x g_i(x^*, 0)$, for all i such that $g_i(x^*, 0) = 0$, and $\nabla_x h_j(x^*, 0)$, $j = 1, ..., p$, are linearly independent.

A4) Strict complementary slackness (SCS) holds at $(x^*, 0)$, i. e. $u_i^* > 0$ for all i such that $g_i(x^*, 0) = 0$.

We recall also the basic local characterizations of a Karush-Kuhn-Tucker triplet, given by Theorem 6 (Fiacco (1976)). If assumptions A1), A2), A3) and A4) hold for problem $(NLP(\varepsilon))$ at $(x^*, 0)$, then:

(a) $\quad x^*$ is a local isolated minimizing point of problem $(NLP(0))$ and the associated multipliers u^* and w^* are unique.

(b) $\quad$ For ε in a neighborhood of 0, there exists a unique, once continuously differentiable vector function $y(\varepsilon) = [x(\varepsilon), u(\varepsilon), w(\varepsilon)]^\top$ satisfying the second order sufficient conditions for a local minimum of problem $(NLP(\varepsilon))$ such that $y(0) = [x^*, u^*, w^*]^\top = y^*$ and hence, $x(\varepsilon)$ is a locally unique local minimum of problem $(NLP(\varepsilon))$ with associated unique multipliers $u(\varepsilon)$ and $w(\varepsilon)$.

(c) $\quad$ For ε near 0, the set of active inequalities is unchanged, strict complementary slackness holds for $u_i(\varepsilon)$ for i such that $g_i(x(\varepsilon), \varepsilon) = 0$, and the active constraint gradients are linearly independent at $x(\varepsilon)$.

We recall that when $y(\varepsilon) = [x(\varepsilon), u(\varepsilon), w(\varepsilon)]^\top$ is available, then $\nabla_\varepsilon y(\varepsilon)$ can be calculated near $\varepsilon = 0$ by the expression

$$\nabla_\varepsilon y(\varepsilon) = [M(\varepsilon)]^{-1} N(\varepsilon)$$

where $M(\varepsilon)$ is the Jacobian with respect to (x, u, w) of the following Karush-Kuhn-Tucker system (satisfied by $y(\varepsilon)$ near $\varepsilon = 0$, thanks to the previous conclusion (b)):

$$\begin{gathered} \nabla_x \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right] = 0, \\ u_i(\varepsilon) g_i\left[x(\varepsilon), \varepsilon\right] = 0, \ i = 1, ..., m, \\ h_j\left[x(\varepsilon), \varepsilon\right] = 0, \ j = 1, ..., p, \end{gathered}$$

and $N(\varepsilon)$ is the negative of the Jacobian of the Karush-Kuhn-Tucker system with respect to ε.

The authors quoted above use the *logarithmic-quadratic-barrier-penalty function* for problem $(NLP(\varepsilon))$, considered by Fiacco and McCormick (1968), and defined as

$$W(x, \varepsilon, r) = f(x, \varepsilon) - r \sum_{i=1}^{m} \ln(g_i(x, \varepsilon)) + \frac{1}{2r} \sum_{j=1}^{p} h_j^2(x, \varepsilon).$$

Under assumptions A1), A2, A3) and A4, the following facts are known for problem $(NLP(0))$ from penalty function theory (see Fiacco and McCormick (1968)):

(1) $\quad$ For $r > 0$ and small, there exists a unique once continuously differentiable vector function $x(0, r)$ such that $x(0, r)$ is a locally unique minimizing point of $W(x, 0, r)$ in

$$R^{\#}(0) = \{x : g_i(x, 0) > 0, \ i = 1, ..., m, \text{ and } h_j(x, 0) = 0, \ j = 1, ..., p\}$$

and such that $x(0, r) \longrightarrow x(0, 0) = x^*$, x^* local isolated minimizing point of $(NLP(0))$.

(2)

$$\lim_{r \longrightarrow 0} r \sum_{i=1}^{m} \ln(g_i(x(0, r), 0) = 0.$$

(3)

$$\lim_{r \longrightarrow 0} \frac{1}{2r} \sum_{j=1}^{p} h_j^2(x(0, r), 0) = 0.$$

(4)
$$\lim_{r \longrightarrow 0} W(x(0,r),0,r) = f(x^*,0).$$

The following theorem extends these results for problem $(NLP(\varepsilon))$, where ε is allowed to vary in a neighborhood of 0, and provides a basis for approximating the sensitivity information associated with problem $(NLP(\varepsilon))$.

The notation $\nabla_x^2 W$ denotes the matrix of second partial derivatives of W with respect to x.

Theorem 43 (Fiacco (1976)). If assumptions A1), A2), A3) and A4) hold for $(NLP(\varepsilon))$, then in a neighborhood about $(\varepsilon, r) = (0,0)$ there exists a unique once continuously differentiable vector function
$$y(\varepsilon, r) = [x(\varepsilon,r), u(\varepsilon,r), w(\varepsilon,r)]^\top$$
satisfying
$$\begin{aligned} \nabla_x \mathcal{L}(x,u,w,\varepsilon) &= 0, \\ u_i g_i(x,\varepsilon) &= r, \;\; i = 1,...,m, \\ h_j(x,\varepsilon) &= w_j r, \;\; j = 1,...,p, \end{aligned}$$
with $y(0,0) = (x^*, u^*, w^*)^\top$ and such that, for any (ε, r) near $(0,0)$ and $r > 0$, $x(\varepsilon,r)$ is a locally unique unconstrained local minimizing point of $W(x,\varepsilon,r)$, $g_i(x(\varepsilon,r),\varepsilon) > 0$, $i = 1,...,m$, and $\nabla_x^2 W(x(\varepsilon,r),\varepsilon,r)$ is positive definite.

Just as before, we can immediately differentiate the above equations with respect to ε, now to obtain the perturbed formula for the parameter derivative $\nabla_\varepsilon y(\varepsilon)$. The analogous reasoning applies and we can obtain
$$\nabla_\varepsilon y(\varepsilon,r) = [M(\varepsilon,r)]^{-1} N(\varepsilon,r)$$
where M and $-N$ are the Jacobians of the perturbed KKT system with respect to (x,u,w) and ε, respectively.

Corollary 1 (Fiacco (1976)). If assumptions A1), A2), A3) and A4) hold for $(NLP(\varepsilon))$, then for any ε near 0,

(a)
$$\lim_{r \longrightarrow 0^+} y(\varepsilon,r) = y(\varepsilon,0) = y(\varepsilon),$$
the Kuhn-Tucker triplet characterized in Theorem 6 and previously recalled in the present section.

(b)
$$\lim_{r \longrightarrow 0^+} \nabla_\varepsilon y(\varepsilon,r) = \nabla_\varepsilon y(\varepsilon,0) = \nabla_\varepsilon y(\varepsilon).$$

This result motivates use of $\nabla_\varepsilon y(\varepsilon,r)$ to estimate $\nabla_\varepsilon y(\varepsilon)$, when ε is near 0 and r is near 0, once $y(\varepsilon,r)$ is available. Theorem 43 provides the basis for an efficient calculation of $\nabla_\varepsilon y(\varepsilon)$, since, at a local solution point (x,r) of $W(x,\varepsilon,r)$, it follows that
$$\nabla_x W\left[x(\varepsilon,r),\varepsilon,r\right] = 0, \tag{7}$$

we can differentiate (7) with respect to ε to obtain

$$\nabla_x^2 W\left[x(\varepsilon,r),\varepsilon,r\right]\nabla_\varepsilon x(\varepsilon,r)+\nabla_\varepsilon(\nabla_x W\left[x(\varepsilon,r),\varepsilon,r\right])=0. \tag{8}$$

By Theorem 43, $\nabla_x^2 W$ is positive definite for (ε,r) near $(0,0)$ and $r>0$, so $\nabla_x^2 W$ has an inverse and

$$\nabla_\varepsilon x(\varepsilon,r)=-(\nabla_x^2 W\left[x(\varepsilon,r),\varepsilon,r\right])^{-1}\nabla_{\varepsilon x}^2 W\left[x(\varepsilon,r),\varepsilon,r\right].$$

Also, since

$$u_i(\varepsilon,r)=r\diagup g_i(x(\varepsilon,r),\varepsilon),\ i=1,...,m, \tag{9}$$

and

$$w_j(\varepsilon,r)=h_j(x(\varepsilon,r),\varepsilon)\diagup r,\ j=1,...,p, \tag{10}$$

for (ε,r) near $(0,0)$ and $r>0$, these equations can be differentiated with respect to ε to obtain

$$\nabla_\varepsilon u_i(\varepsilon,r)=-(r\diagup g_i^2)\left[\nabla_x g_i(x(\varepsilon,r),\varepsilon)\nabla_\varepsilon x(\varepsilon,r)+\partial g_i(x(\varepsilon,r),\varepsilon)\diagup\partial\varepsilon\right], \tag{11}$$

$$\nabla_\varepsilon w_j(\varepsilon,r)=(1\diagup r)\left[\nabla_x h_j(x(\varepsilon,r),\varepsilon)\nabla_\varepsilon x(\varepsilon,r)+\partial h_j(x(\varepsilon,r),\varepsilon)\diagup\partial\varepsilon\right]. \tag{12}$$

Solving (8) and calculating (11) and (12) then yields the components of $\nabla_\varepsilon y(\varepsilon,r)$, which can be used to estimate $\nabla_\varepsilon y(\varepsilon)$ for (ε,r) near $(0,0)$.

The next results extend this theory to an analysis of the optimal value function of problem $(NLP(\varepsilon))$ along the Karush-Kuhn-Tucker point trajectory $\left[x(\varepsilon),u(\varepsilon),w(\varepsilon)\right]^\top$.

The optimal value function is defined as:

$$f^*(\varepsilon)=f\left[x(\varepsilon),\varepsilon\right] \tag{13}$$

and the "optimal value Lagrangian" is defined as:

$$\mathcal{L}^*(\varepsilon)=\mathcal{L}\left[x(\varepsilon),u(\varepsilon),w(\varepsilon),\varepsilon\right],$$

where now

$$\mathcal{L}(x,u,w,\varepsilon)=f(x,\varepsilon)-\sum_{i=1}^m u_i g_i(x,\varepsilon)+\sum_{j=1}^p w_j h_j(x,\varepsilon).$$

The logarithmic-quadratic loss penalty function considered in the present section can also be used to provide estimates of the first and second order sensitivity of the optimal value function. Let the optimal value penalty function be defined as $W^*(\varepsilon,r)=W(x(\varepsilon,r),\varepsilon,r)$.

Theorem 44 (Armacost and Fiacco (1975)). If assumptions A1), A2), A3) and A4) hold for problem $(NLP(\varepsilon))$, then for (ε,r) near $(0,0)$ and $r>0$, $W^*(\varepsilon,r)$ is a twice continuously differentiable function of ε and

(a)

$$\lim_{r\longrightarrow 0^+}W^*(\varepsilon,r)=\mathcal{L}^*(\varepsilon)=f^*(\varepsilon);$$

(*b*)

$$\nabla_\varepsilon W^*(\varepsilon, r) = \nabla_x W \nabla_\varepsilon x + \nabla_\varepsilon W = \nabla_\varepsilon \mathcal{L}(x, u, w, \varepsilon) \mid_{(x,u,w)=(x(\varepsilon,r),u(\varepsilon,r),w(\varepsilon,r))};$$

(*c*)

$$\lim_{r \longrightarrow 0^+} \nabla_\varepsilon W^*(\varepsilon, r) = \nabla_\varepsilon \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right] = \nabla_\varepsilon f^*(\varepsilon); \tag{14}$$

(*d*)

$$\nabla^2_\varepsilon W^*(\varepsilon, r) = \nabla_\varepsilon(\nabla_\varepsilon \mathcal{L}\left[x(\varepsilon), u(\varepsilon), w(\varepsilon), \varepsilon\right]^\top);$$

(*e*)

$$\lim_{r \longrightarrow 0^+} \nabla^2_\varepsilon W^*(\varepsilon, r) = \nabla^2_\varepsilon f^*(\varepsilon).$$

These results provide a justification for estimating $f^*(\varepsilon)$, $\nabla_\varepsilon f^*(\varepsilon)$ and $\nabla^2_\varepsilon f^*(\varepsilon)$ by $W^*(\varepsilon, r)$, $\nabla_\varepsilon W^*(\varepsilon, r)$ and $\nabla^2_\varepsilon W^*(\varepsilon, r)$, respectively, when r is positive and small enough. Since Corollary 1 and continuity imply that

$$\lim_{r \longrightarrow 0^+} f(x(\varepsilon, r), \varepsilon) = f^*(\varepsilon),$$

another estimate of the optimal value function (13) is provided by $f^\#(\varepsilon, r) \equiv f(x(\varepsilon, r), \varepsilon)$ when $r > 0$ and small. Direct application of the chain rule for differentiation then yields, for $x = x(\varepsilon, r)$

$$\nabla_\varepsilon f^\#(\varepsilon, r) = \nabla_x f(x, \varepsilon) \nabla_\varepsilon x(\varepsilon, r) + \nabla_\varepsilon f(x, \varepsilon). \tag{15}$$

Under the given assumptions, continuity also assures that $\nabla_\varepsilon f^\#(\varepsilon, r) \longrightarrow \nabla_\varepsilon f^*(\varepsilon)$ as $r \longrightarrow 0^+$. Thus both $\nabla_\varepsilon f^\#(\varepsilon, r)$ and $\nabla_\varepsilon W^*(\varepsilon, r)$ are estimates of $\nabla_\varepsilon f^*(\varepsilon)$ for r sufficiently small. It should be noted that these estimates are functionally related since

$$\begin{aligned} \nabla_\varepsilon W^*(\varepsilon, r) &= \nabla_\varepsilon f^\#(\varepsilon, r) - \sum_{i=1}^m u_i(\nabla_x g_i \nabla_\varepsilon x(\varepsilon, r) + \nabla_\varepsilon g_i) + \\ &\quad + \sum_{j=1}^p w_j(\nabla_x h_j \nabla_\varepsilon x(\varepsilon, r) + \nabla_\varepsilon h_j)_{|x=x(\varepsilon,r)}. \end{aligned}$$

From this expression, it is clear that $\nabla_\varepsilon f^\#(\varepsilon, r)$ is the better estimate of $\nabla_\varepsilon f^*(\varepsilon)$, the remaining terms in $\nabla_\varepsilon W^*(\varepsilon, r)$ simply consisting "noise" that is eliminated as $r \longrightarrow 0^+$. However, by using the expression for $\nabla_\varepsilon W^*(\varepsilon, r)$ given by (14) $\nabla_\varepsilon W^*(\varepsilon, r)$ can be calculated *without necessitating* the calculation of $\nabla_\varepsilon x(\varepsilon, r)$, which is required to compute (15).

In summary, the basis for the estimation procedure, utilized by the authors quoted at the beginning of the present section, for a specific problem, say $(NLP(\varepsilon))$, is the minimization of the penalty function $W(x, \varepsilon, r)$ given by

$$W(x, \varepsilon, r) = f(x, \varepsilon) - r \sum_{i=1}^m \ln(g_i(x, \varepsilon)) + \frac{1}{2r} \sum_{j=1}^p h_j^2(x, \varepsilon).$$

This yields a point $x(\varepsilon, r)$ which may be viewed as an estimate of a (local) solution $x(\varepsilon)$ of problem $(NLP(\varepsilon))$. The estimate $f(x(\varepsilon, r), \varepsilon)$ of $f^*(\varepsilon)$ is immediately available when $x(\varepsilon, r)$ has been determined, and the two estimates of $\nabla_\varepsilon f^*(\varepsilon)$ given by (14) and (15) were already discussed. The associated optimal Lagrange multipliers $u(\varepsilon)$ and $w(\varepsilon)$ are estimated by using the relationships given in (9) and (10), respectively. For other considerations on these types of problems the reader may consult Fiacco (1983a).

References

S. ACHMANOV (1984), Programmation Linéaire, Editions MIR, Moscou.

M. AKGÜL (1984), *A note on shadow prices in linear programming,* J. Operational Research Soc., 35, 425-431.

R. L. ARMACOST and A. V. FIACCO (1975), *Second order parametric sensitivity analysis in NLP and estimates by penalty functions methods,* Technical Paper Serial T-324, Institute for Management Science and Engineering, The George Washington University, Washington D. C.

R. L. ARMACOST and A. V. FIACCO (1976), *NLP sensitivity analysis for RHS perturbations: a brief survey and second-order extensions,* Technical Paper Serial T-334, Institute for Management Science and Engineering, The George Washington University, Washington D. C.

R. L. ARMACOST and A. V. FIACCO (1978), *Sensitivity analysis for parametric nonlinear programming using penalty methods;* in Computers and Mathematical Programming, National Bureau of Standards Special Publication 502, 261-269.

R. L. ARMACOST and A. V. FIACCO (1979), *Sensitivity analysis of a well-behaved Kuhn-Tucker triple;* in A. Prekopa (Ed.), Survey of Mathematical Programming - Proceedings of the 9th International Mathematical Programming Symposium, Budapest, August 23-27, 1976, North Holland Publishing Company, Amsterdam, 121-134.

D. C. AUCAMP and D. I. STEINBERG (1979), *On the nonequivalence of shadow prices and dual variables,* Technical Report WUCW-79-11, Washington University Department of Computer Sciences, St Louis, Missouri.

D. C. AUCAMP and D. I. STEINBERG (1982), *The computation of shadow prices in linear programming,* J. Operational Research Soc., 33, 557-565.

A. AUSLENDER and R. COMINETTI (1990), *First and second order sensitivity analysis of nonlinear programs under directional constraint conditions,* Optimization, 21, 351-363.

C. BAIOCCHI and A. CAPELO (1984), Variational and Quasivariational Inequalities: Application to Free-Boundary Problems, J. Wiley, New York.

M. L. BALINSKI and W. J. BAUMOL (1968), *The dual in nonlinear programming and its economic interpretation,* Rev. of Economic Studies, 35, 237-256.

B. BANK, J. GUDDAT, D. KLATTE, B. KUMMER and K. TAMMER (1982), Non-Linear Parametric Optimization, Akademie-Verlag, Berlin.

E. BEDNARCZUK (1994), *Sensitivity in mathematical programming: a review,* Control and Cybernetics, 23, 589-604.

C. BERGE (1963), Topological Spaces, Macmillan, New York.

D. BERTSEKAS (2016), Nonlinear Programming (Third Edition), Athena Scientific, Belmont, Mass.

J. H. BIGELOW and N. Z. SHAPIRO (1974), *Implicit function theorems for mathematical programming and for systems of inequalities,* Mathematical Programming, 6, 141-156.

V. BONDAREVSKY, A, LESCHOV and L. MINCHENKO (2016), *Value functions and their directional derivatives in parametric nonlinear programming,* J. Optim. Theory Appl., 171, 440-464.

J. F. BONNANS and R. COMINETTI (1996), *Perturbed optimization in Banach spaces I: a general theory based on weak directional constraint qualifications; II: A theory based on strong directional constraint qualifications; III: Semi-infinite optimization,* SIAM J. Control Optim., 34, 1151-1171; 1172-1189 and 1555-1567.

J. F. BONNANS and A. SHAPIRO (1998), *Optimization problems with perturbations: a guided tour*, SIAM Rev., 40, 228-264.

J. F. BONNANS and A. SHAPIRO (2000), Perturbation Analysis of Optimization Problems, Springer-Verlag, New York.

E. V. CHOO and K. P. CHEW (1985), *Optimal value functions in parametric programming,* Z. Oper. Res., 29, 47-57.

G. P. CRESPI and G. GIORGI (2002), *Sensitivity analysis for variational inequalities,* Journal of Interdisciplinary Mathematics, 5, 165-176.

J. M. DANSKIN (1966), *The theory of max-min with applications,* SIAM J. Appl. Math., 14, 651-665.

J. M. DANSKIN (1967), The Theory of Max-Min, Springer Verlag, New York.

S. DEMPE (1993), *Directional differentiability of optimal solutions under Slater's condition,* Mathematical Programming, 59, 48-69.

S. DEMPE (2002), Foundations of Bilevel Programming, Kluwer Academic Publishers, Dordrecht, Boston and London.

A. DHARA and J. DUTTA (2012), Optimality Conditions in Convex Optimization. A Finite-Dimensional View, CRC Press, Boca Raton, London. New York.

W. E. DIEWERT (1984), *Sensitivity analysis in economics,* Comput. & Ops. Res., 11, 141-156.

J. DUGGAN and T. KALANDRAKIS (2007), *A note on sensitivity analysis for local solutions of non-linear programs,* Unpublished, University of Rochester.

J. P. EVANS and F. J. GOULD (1970), *Stability in nonlinear programming,* Op. Res., 18, 107-118.

F. FACCHINEI and J.-S. PANG (2003), Finite-Dimensional Variational Inequalities and Complementarity Problems (2 volumes), Springer, New York.

A. V. FIACCO (1976), *Sensitivity analysis for nonlinear programming using penalty methods,* Math. Programming, 10, 287-311.

A. V. FIACCO (1980), *Nonlinear programming sensitivity analysis results using strong second order assumptions;* in Numerical Optimization of Dynamic Systems (L. C. W. Dixon and G. P. Szegö, Eds.), North Holland, Amsterdam, 327-348.

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

A. V. FIACCO (1983b), *Optimal value continuity and differential stability bounds under the Mangasarian-Fromovitz constraint qualification;* in A. V. Fiacco (ed.), Mathematical Programming with Data Perturbations, Vol. 2, Marcel Dekker, New York, 65-90.

A. V. FIACCO and W. P. HUTZLER (1982a), *Basic results in the development of sensitivity and stability analysis in nonlinear programming;* in Mathematical Programming with Parameters and Multi-Level Constraints (J. E. Falk and A. V. Fiacco, Eds.), Special Issue of Computers and Operations Research, 9, 9-28.

A. V. FIACCO and W. P. HUTZLER (1982b), *Optimal value differential stability results for general inequality constrained differentiable mathematical programs;* in A. V. Fiacco (Ed.), Mathematical Programming with Data Perturbations, Vol. 1, Marcel Dekker, New York, 29-43.

A. V. FIACCO and Y. ISHIZUKA (1990), *Sensitivity and stability analysis for nonlinear programming,* Annals of Operations Research, 27, 215-235.

A. V. FIACCO and J. KYPARISIS (1985), *Sensitivity analysis in nonlinear programming under second order assumptions;* in A. Bagchi and H. Th. Jongen (eds.), Systems and Optimization. Proceedings of the Twente Workshop, Enschede, The Netherlands, April 16-18, 1984, Springer-Verlag, Berlin, 74-97.

A. V. FIACCO and J. KYPARISIS (1986), *Convexity and concavity properties of the optimal value function in parametric nonlinear programming,* J. Optim. Theory Appl., 48, 95-126.

A. V. FIACCO and J. KYPARISIS (1988), *Computable bounds on parametric solutions of convex problems,* Mathematical Programming, 40, 213-221.

A. V. FIACCO and J. LIU (1993), *Degeneracy in NLP and the development of results motivated by its presence,* Annals of Operations Research, 46, 61-80.

A. V. FIACCO and J. LIU (1995), *Extensions of algorithmic sensitivity calculations in nonlinear programming using barrier and penalty functions,* Optimization, 32, 335-350.

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

M. F. FLORENZANO and C. LE VAN (2001), Finite Dimensional Convexity and Optimization, Springer-Verlag, Berlin.

T. GAL (1979), Postoptimal Analysis, Parametric Programming and Related Topics, McGraw-Hill, New York.

T. GAL (1984), *Linear parametric programming- a brief survey,* Math. Programming Study, 21, 43-68.

T. GAL (1986), *Shadow prices and sensitivity analysis in linear programming under degeneracy. State-of-the art-survey,* OR Spektrum, 8, 59-71.

D. GALE (1967), *A geometric duality theorem with economic applications,* Rev. of Economic Studies, 34, 19-24.

J. GAUVIN (1977), *A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming,* Mathematical Programming, 12, 136-138.

J. GAUVIN (1979), *The generalized gradient of a marginal function in mathematical programming,* Math. Oper. Res., 4, 458-463.

J. GAUVIN (1980a), *Shadow prices in nonconvex mathematical programming,* Math. Programming, 19, 300-312.

J. GAUVIN (1980b), *Quelques précisions sur les prix marginaux en programmation linéaire,* INFOR, 18, 68-73.

J. GAUVIN (1988), *Directional derivative for the value function in mathematical programming;* in F. H. Clarke, V. F. Dem'yanov and F. Giannessi (Eds.), Nonsmooth Optimization and Related Topics, Plenum Press, New York, 167-184.

J. GAUVIN (1995a), *Degeneracy, normality, stability in mathematical programming;* in R. Durier and C. Michelot (Eds.), Recent Developments in Optimization - Seventh French-german Conference on Optimization, Springer, Berlin, 136-141.

J. GAUVIN (1995b), Leçons de Programmation Mathématique, Editions de l'Ecole Polytechnique de Montréal, Montréal.

J. GAUVIN (2001), *Formulae for the sensitivity analysis in linear programming problems;* in M. Lassonde (Ed.), Approximation, Optimization and Mathematical Economics, Physica-Verlag, Heidelberg, 117-120.

J. GAUVIN and F. DUBEAU (1982), *Differential properties of the marginal function in mathematical programming,* Math. Programming Study, 19, 101-119.

J. GAUVIN and R. JANIN (1990), *Directional derivative of the value function in parametric optimization,* Annals of Operations Research, 27, 237-252.

J. GAUVIN and J. W. TOLLE (1977), *Differential stability in nonlinear programming,* SIAM J. Control Optim., 15, 294-311.

A. M. GEOFFRION (1971), *Duality in nonlinear programming: a simplified applications-oriented development,* SIAM Rev., 13, 1-37.

G. GIORGI and T. H. KJELDSEN (2014), Traces and Emergence of Nonlinear Programming, Birkäuser, Basel.

G. GIORGI and C. ZUCCOTTI (2008-2009), *Some results and remarks on the envelope theorem,* Quaderno di Ricerca - XII, Facoltà di Economia, Università di Pavia, Dipartimento di Ricerche Aziendali "Riccardo Argenziano", a. a. 2008-2009.

M. S. GOWDA and J. S. PANG (1994), *Stability analysis of variational inequalities and nonlinear complementarity problems, via the mixed linear complementarity problem and degree theory, Math. Oper. Res., 19, 831-879.*

H. J. GREENBERG and W. P. PIERSKALLA (1972), *Extensions of the Evans-Gould stability theorems for mathematical programs,* Operations Research, 20, 143-153.

S. P. HAN and O. L. MANGASARIAN (1979), *Exact penalty functions in nonlinear programming,* Mathematical Programming, 17, 251-269.

P. T. HARKER and J. S. PANG (1990), *Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications,* Mathematical Programming, 48, 161-220.

J.-B. HIRIART-URRUTY (2008), Les Mathématiques du Mieux Faire - Volume 1 - Premiers Pas en Optimisation, Ellipses Editions Marketing S. A., Paris.

W. W. HOGAN (1973a), *Directional derivatives for extremal value functions with applications to the completely convex case,* Oper. Res., 21, 188-209.

W. W. HOGAN (1973b), *Point-to-set maps in mathematical programming,* SIAM Rev., 15, 591-603.

W. W. HOGAN (1973c), *The continuity of the perturbation function of a convex program,* Op. Res., 21, 351-352.

R. HORST (1984a), *Shadow prices and equilibrium prices in linear and nonlinear decision models,* Methods of Operations Research, vol. 48, 525-543.

R. HORST (1984b), *On the interpretation of optimal value dual solutions in convex programming,* J. Op. Res. Soc., 35, 327-335.

M. D. INTRILIGATOR (1971), Mathematical Optimization and Economic Theory, Prentice Hall, Englewood Cliffs, N. J.

R. JANIN (1984), *Directional derivative of the marginal function in nonlinear programming,* Math. Programming Study, 21, 110-126.

B. JANSEN, J. J. DE JONG, C. ROOS and T. TERLAKY (1997), *Sensitivity analysis in linear programming: just be careful!,* European Journal of Operational Research, 101, 15-28.

K. JITTORNTRUM (1984), *Solution point differentiability without strict complementarity in nonlinear programming,* Math. Programming Study, 21, 127-138.

W. KARUSH (1939), Minima of functions of several variables with inequalities as side conditions, M. S. Thesis, Department of Mathematics, University of Chicago. Reprinted in G. Giorgi and T. Kjeldsen (2014).

A. J. KING and R. T. ROCKAFELLAR (1992), *Sensitivity analysis for nonsmooth generalized equations,* Math. Programming, 55, 193-212.

M. KOJIMA (1980), *Strongly stable stationary solutions in nonlinear programs;* in S. M. Robinson (Ed.), Analysis and Computation of Fixed Points, Academic Press, New York, 93-138.

D. KINDERLEHRER and G. STAMPACCHIA (1980), An Introduction to Variational Inequalities and Their Applications, Academic Press, New York.

H. W. KUHN and A. W. TUCKER (1951), *Nonlinear programming;* in J. Neyman (Ed.), Proceedings of the 2nd Berkeley Symposiun on Mathematical Statistics and Probability, University of California Press, Berkeley, Cal. Reprinted in G. Giorgi and T. Kjeldsen (2014).

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

J. KYPARISIS (1987a), *Sensitivity analysis for non-linear programs with linear constraints,* Operations Research Letters, 6, 275-280.

J. KYPARISIS (1987b), *Sensitivity analysis framework for variational inequalities,* Math. Programming, 38, 203-213.

J. KYPARISIS (1990a), *Sensitivity analysis for variational inequalities and nonlinear complementarity problems,* Annals of Oper. Res., 27, 143-174.

J. KYPARISIS (1990b), *Sensitivity analysis for nonlinear programs and variational inequalities with nonunique multipliers,* Math. Oper. Res., 15, 286-298.

J. KYPARISIS (1990c), *Solution differentiability for variational inequalities,* Mathematical Programming, 48, 285-301.

J. KYPARISIS (1992), *Partametric variational inequalities with multivalued solution sets,* Math. Oper. Res., 12, 341-364.

J. KYPARISIS and A. V. FIACCO (1987), *Generalized convexity and concavity of the optimal value function in nonlinear programming,* Math. Programming, 39, 285-304.

E. S. LEVITIN (1994), Perturbation Theory in Mathematical Programming and its Applications, John Wiley & Sons, Ltd., Chichester.

J. LIU (1995a), *Strong atability in variational inequalities,* SIAM J. Control and Optimization, 33, 725-749.

J. LIU (1995b), *Sensitivity analysis in nonlinear programs and variational inequalities via continuous selections,* SIAM J. Control and Optimization, 33, 1040-1060.

D. G. LUENBERGER and Y. YE (2008), Linear and Nonlinear Programming, Springer, New York.

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

O. L. MANGASARIAN (1979), *Uniqueness of solution in linear programming,* Linear Algebra and Its Appl., 25, 151-162.

G. P. McCORMICK (1967), *Second order conditions for constrained minima,* SIAM Journal on Applied Mathematics, 15, 641-652. Reprinted in G. Giorgi and T. Kjeldsen (2014).

G. P. McCORMICK (1976), *Optimality criteria in nonlinear programming;* in SIAM-AMS Proceedings N. 9, SIAM, Philadelphia, 27-38.

B. MORDUKHOVICH (1994), *Stability theory for parametric generalized equations and variational analysis via nonsmooth analysis,* Transactions of the American Mathematical Society, 343, 609-657.

H. NIKAIDO (1968), Convex Structures and Economics Theory, Academic Press, New York.

W. NOVSHEK (1993), Mathematics for Economists, Academic Press, New York.

J.-S. PANG (1993), *A degree-theoretic approach to parametric nonsmooth equations with multivalued perturbed solutions sets,* Mathematical Programming, 62, 359-383.

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

E. L. PETERSON (1970), *An economic interpretation of duality in linear programming,* J. Math. Anal. Appl., 30, 172-196.

Y. QIU and T. L. MAGNANTI (1992), *Sensitivity analysis for variational inequalities,* Math. Oper. Res., 17, 61-76.

D. Ralph and S. DEMPE (1995), *Directional derivatives of the solution of a parametric nonlinear program,* Math. Programming, 70, 159-172.

S. M. ROBINSON (1974), *Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear programming algorithms,* Mat. Programming, 7, 1-16.

S. M. ROBINSON (1977), *A characterization of stability in linear programming,* Operations Research, 25, 435-447.

S. M. ROBINSON (1980), *Strongly regular generalized equations,* Mathematics of Operations Research, 5, 43-62.

S. M. ROBINSON (1982), *Generalized equations and their solutions, Part II: Applications to nonlinear programming,* Math. Programming Study, 19, 200-221.

S. M. ROBINSON (1987), *Local structure of feasible sets in nonlinear programming. Part III. Stability and sensitivity,* Math. programming Stydy, 30, 45-66.

R. T. ROCKAFELLAR (1970), Convex Analysis, Princeton Univ. Press, Princeton.

R. T. ROCKAFELLAR (1984), *Directional differentiability of the optimal value function in a nonlinear programming problem,* Math. Programming Study, 21, 213-226.

A. RUSZCZYNSKI (2006), Nonlinear Optimization, Princeton Univ. Press, Princeton, N. J.

A. SHAPIRO (1985), *Second order sensitivity analysis and asymptotic theory of parameterized nonlinear program,* Math. Programming, 33, 280-299.

A. SHAPIRO (1988), *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections,* SIAM J. Control Optim., 26, 628-645.

E. SILBERBERG and W. SUEN (2001), The Structure of Economics: A Mathematical Analysis, 3rd edition, McGraw-Hill, New York.

C. P. SIMON and L. BLUME (1994), Mathematics for Economists, W. W. Norton & Co., New York.

A. L. SOYSTER (1981), *An objective function perturbation with economic interpretations,* Management Science, 27, 231-237.

J. E. SPINGARN (1980), *Fixed and variable constraints in sensitivity analysis,* SIAM J. Control Optim., 18, 297-310.

M. H. STERN and D. M. TOPKIS (1976), *Rates of stability in nonlinear programming,* Operations Research, 24, 462-476.

K. SYDSAETER, P. HAMMOND, A. SEIERSTAD, A. STROM (2005), Further Mathematics fo Economic Analysis, Prentice Hall, Pearson Education Ltd, Harlow.

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

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

R. L. TOBIN (1986), *Sensitivity analysis for variational inequalities,* J. Optim. Theory Appl., 48, 191-204.

H. UZAWA (1958), *A note on the Menger-Wieser theory of imputation,* Zeitschrift für Nationalökonomie, 18, 318-334.

G. WACHSMUTH (2013), *On LICQ and the uniqueness of Lagrange multipliers,* Operations Research Letters, 41, 78-80.

C. Y. WANG and F. ZHAO (1994), *Directional derivatives of optimal value functions in mathematical programming,* J. Optim. Theory Appl., 82, 397-404.

J. E. WARD and R. E. WENDELL (1990), *Approaches to sensitivity analysis in linear programming,* Annals of Operations Research, 27, 3-38.

A. C. WILLIAMS (1963), *Marginal values in linear programming,* J. Soc. Indust. Appl. Math., 11, 82-94.