JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 73, No. 1, APRIL 1992

Directional Derivatives of Optimal Solutions in Smooth Nonlinear Programming[1]

J. F. BONNANS[2]

Communicated by A. V. Fiacco

Abstract. We consider a smooth nonlinear program subject to perturbations in the right-hand side of the constraints. We do not assume that the unique solution of the original problem satisfies any qualification hypothesis. We suppose instead that the direction of perturbation satisfies the hypothesis of Gollan. We study the variation of the cost and, with the help of some second-order sufficiency conditions, obtain some conditions satisfied by the first term of the expansion of the solution. These conditions may vary depending on the existence of a Lagrange multiplier for the original problem.

Key Words. Sensitivity, stability, nonlinear programming, calmness.

1. Introduction

This paper is concerned with the perturbation of a standard nonlinear program with smooth data. The first results in this direction where obtained by applying the implicit function theorem; see Fiacco (Ref. 1) for a review of these results. Then, the hypothesis of linear independence of the gradients of active constraints, strict complementary, and second-order sufficiency conditions were needed. Jittorntrum (Ref. 2) relaxed the strict complementarity hypothesis. With the help of a strong second-order sufficiency condition, he obtained the directional differentiability of a local solution. Using a semistrong second-order sufficiency condition, the author generalized this result by studying in Ref. 3 situations where a finite number of computable

[1] A preliminary version of this paper has been presented at the workshop "Analyse Quantitative de la Sensibilité en Optimisation," Sainte Adèle, Québec, Canada, February 19–23, 1989. The author thanks A. Shapiro and the referees for their kind advices.

[2] Directeur de Recherche, INRIA, Domaine de Voluceau, Rocquencourt, France.

directions contain the limit points of the ratio of solutions w.r.t. perturbation. In Ref. 3, an algorithm is given that allows one to compute effectively the solutions of the perturbed problem. Recently, Shapiro (Ref. 4), using the Mangasarian and Fromovitz qualification hypothesis (Ref. 5) and some second-order sufficiency conditions, gave a second-order analysis of the optimal value and formulated an extended quadratic problem for which directional derivatives of the solutions are optimal. On the other hand, following an idea of Gollan (Refs. 6, 7), Gauvin and Janin (Refs. 8, 9, 10) studied the directional derivative of the value functions and the magnitude of the variation in the solution.

Our aim in this paper is to improve these results in two directions. First, when the original problem has no Lagrange multiplier (a situation that has not been much considered), we give an estimate of the variation of the optimal cost (which is of the order of the square root of the perturbation) and formulate a subproblem with quadratic constraints whose solutions have a close connection with the variations in the solution. When some Lagrange multipliers exist, we obtain some relations concerning the second-order analysis of the cost and the first-order variation of the solutions; these results extend those of Shapiro (Ref. 4) and of Gauvin and Janin (Ref. 8).

The paper is organized as follows. In Section 2, we give some notations and preliminary results. Section 3 (resp., Section 4) is concerned with the case when no Lagrange multiplier exists (resp., there exist some Lagrange multipliers). Section 5 presents two examples in which the results of Section 3 and Section 4 allow one to compute effectively the first term of the expansion of the solutions.

2. Preliminaries

This paper is concerned with the following problem:

$$(\mathrm{P}(u)) \qquad \min \; f(x), \quad \text{s.t.} \quad x \in \mathbb{R}^n, \quad g_I(x) \leq u_I, \quad g_J(x) = u_J,$$

with $I = (1, \ldots, q)$, $J = (q+1, \ldots, p)$, $u \in \mathbb{R}^p$, f and g being three times continuously differentiable. In order to shorten the notations, for y, z in $\mathbb{R}^p$ we will write $y \ll z$ whenever $y_I \leq z_I$ and $y_J = z_J$. More generally, for K such that $K \subset I$, we will write

$$y \overset{K}{\ll} z$$

whenever

$$y_i \leq z_i, \qquad i \in K,$$

$$y_J = z_J.$$

The set of feasible points of P(u) is

$$F(u) := \{x \in \mathbb{R}^n;\, g(x) \ll u\},$$

and the set of solutions of P(u) is

$$S(u) := \{x \in F(u);\, f(x) = \inf \mathrm{P}(u)\}.$$

For any x in $\mathbb{R}^n$, we will say that $\lambda = (\lambda_0, \ldots, \lambda_p) \in \mathbb{R}^{p+1}$ is a multiplier associated to x for problem P(u) if $\lambda \neq 0$ and (x, λ) satisfies the following first-order optimality conditions:

$$\lambda_0 \nabla f(x) + \sum_{i=1}^{p} \lambda_i \nabla g_i(x) = 0,$$

$$g(x) \ll u,$$

$$\lambda_0 \geq 0, \qquad \lambda_I \geq 0,$$

$$\lambda_i(g_i(x) - u_i) = 0, \qquad \forall i \in I.$$

It is well known that to each local solution of P(u) is associated at least one multiplier. We denote by $\Lambda(u, x)$ the set of multipliers. Subsets of $\Lambda(u, x)$ of interest are the set of degenerate multipliers,

$$\Lambda_0(u, x) = \{\lambda \in \Lambda(u, x);\, \lambda_0 = 0\},$$

and the set of Lagrange multipliers,

$$\Lambda_1(u, x) = \{\lambda \in \Lambda(u, x);\, \lambda_0 = 1\}.$$

These degenerate multipliers and Lagrange multipliers can often be identified without confusion to elements of $\mathbb{R}^p$; e.g., for z in $\mathbb{R}^p$, by $\lambda' z$ we will mean $\sum_{i=1}^{p} \lambda_i z_i$. We will denote $\Lambda_i(0, \bar{x})$ by Λ_i $(i = 0, 1)$.

In all this paper, we will assume that the two following hypotheses hold:

$$S(0) = \{\bar{x}\}, \tag{1}$$

i.e., P(0) has a unique solution, and a compactness condition,

$$\exists \alpha > 0, \qquad \beta > 0,$$

$$\text{s.t.} \quad \text{if } \|u\| \leq \beta \text{ and } x \in S(u), \text{ then } \|x\| \leq \alpha. \tag{2}$$

As a consequence, we have that, if for some sequence $u^k \to 0$ we have $x^k \in S(u^k)$, then $x^k \to \bar{x}$.

Let us define the set of active constraints

$$I(u, x) = \{i \in I;\ g_i(x) = u\}.$$

We will denote $I(0, \bar{x})$ by $\bar{I}$.

We will say that a point x in $F(u)$ satisfies the MF (Mangasarian and Fromovitz) hypothesis for some subset $K \subset I$ if the following holds:

(i) $\{\nabla g_i(x)\}_{i \in J}$ are linearly independent;
(ii) there exists d in $\ker g'_J(x)$ satisfying $g'_K(x)d < 0$.

If this relation is satisfied for $K = I(u, x)$ and $x \in F(u)$, we will say that x satisfies MF for problem P(u); this is the classical hypothesis of Mangasarian and Fromovitz (Ref. 5).

The case when $\bar{x}$ satisfies MF for problem P(0) has been studied in detail by Shapiro (Ref. 4). We will relax this hypothesis. However, in order to ensure the nonemptiness of the feasible set of the perturbed problem, we will assume that the following hypothesis due to Gollan (see Refs. 6, 7) holds:

$$\lambda' u > 0, \qquad \text{for all } \lambda \text{ in } \Lambda_0. \tag{3}$$

A condition equivalent to (3) is given below.

Lemma 2.1. Hypothesis (3) is equivalent to the following requirement:

(i) $\{\nabla g_i(\bar{x})\}_{i \in J}$ are linearly independent;
(ii) $\exists \bar{d}$ in $\mathbb{R}^n$ and $\gamma > 0$ s.t.
$g'_J(\bar{x})\bar{d} = u_J$,
$g'_i(\bar{x})\bar{d} + \gamma \leq u_i$, for all i in $\bar{I}$.

Proof. Hypothesis (3) says that the following system has no nonnull solution:

$$\sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) = 0,$$

$$\sum_{i=1}^{p} \lambda_i u_i \leq 0,$$

$$\lambda_I \geq 0,$$

$$\lambda_i = 0, \qquad \text{if } g_i(\bar{x}) < 0,\ \forall i \in I.$$

This is equivalent to saying that the following system has no nonnull solution (here, $\theta \in \mathbb{R}$):

$$\sum_{i=1}^{p} \lambda_i \begin{bmatrix} \nabla g_i(\bar{x}) \\ -u_i \end{bmatrix} + \theta \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0,$$

$$\lambda_I \geq 0, \qquad \theta \geq 0,$$

$$\lambda_i = 0, \qquad \text{if } g_i(\bar{x}) < 0, \ \forall i \in I.$$

But this is the dual of the MF condition applied to the system (here, $\sigma \in \mathbb{R}$)

$$g'(\bar{x})d - \sigma u \overset{I}{\ll} 0,$$

$$-\sigma \leq 0.$$

Hence, (3) is equivalent to

$$\left\{\begin{pmatrix} \nabla g_i(\bar{x}) \\ -u_i \end{pmatrix}\right\}_{i \in J} \quad \text{linearly independent,}$$

i.e., $(d, \sigma) \to g'_J(\bar{x})d - \sigma u_J$ surjective and there exists $(\bar{d}, \bar{\sigma})$ in $\mathbb{R}^n \times \mathbb{R}$ satisfying

$$g'_J(\bar{x})\bar{d} - \bar{\sigma} u_J = 0,$$

$$g'_i(\bar{x})\bar{d} - \bar{\sigma} u_i < 0, \qquad \text{for all } i \text{ in } \bar{I},$$

$$-\bar{\sigma} < 0.$$

The above conditions are an easy consequence of (i), (ii); hence, we just have to prove that they imply (i), (ii). As $\bar{\sigma} > 0$, multiplying $(\bar{d}, \bar{\sigma})$ by $1/\bar{\sigma}$, we find (ii). The first condition says that, for any e in $\mathbb{R}^p$, there exists (d, σ) such that

$$g'_J(\bar{x})d - \sigma u_J = e.$$

From this and $g'_J(\bar{x})\bar{d} - \bar{\sigma} u_J = 0$, we deduce that

$$g'_J(\bar{x})(d - (\sigma/\bar{\sigma})\bar{d}) = e.$$

Hence, $g'_J(\bar{x})$ is surjective, i.e., (i) holds. □

From Lemma 2.1, we will deduce a nice property concerning $F(\sigma u)$ when u satisfies (3). We define the two following closed sets:

$$\tau = \{d \in \mathbb{R}^n; \exists \sigma^k \searrow 0, x^k \in F(\sigma^k u); (x^k - \bar{x})/\sigma^k \to d\},$$

$$T = \{d \in \mathbb{R}^n; g'(\bar{x})d \overset{I}{\ll} u\}.$$

Lemma 2.2. Hypothesis (3) implies that $\tau = T$. In addition, for any d in T, there exists $x(\sigma) = \bar{x} + \sigma d + O(\sigma^2)$ in $F(\sigma u)$ for $\sigma > 0$ small enough.

Proof. The inclusion $\tau \subset T$ is easily obtained. Conversely, pick d in T. For $\beta > 0$ given, consider

$$y(\sigma) = \bar{x} + \sigma[(1-\beta\sigma)d + \beta\sigma\bar{d}],$$

where $\bar{d}$ satisfies the conditions of Lemma 2.1. Then, for i in $\bar{I}$, we have

$$\begin{aligned} g_i(y(\sigma)) &= \sigma g_i'(\bar{x})[(1-\beta\sigma)d + \beta\sigma\bar{d}] + O(\sigma^2) \\ &\le \sigma(1-\beta\sigma)u_i + \beta\sigma^2(u_i - \gamma) + O(\sigma^2) \\ &= \sigma u_i - \beta\gamma\sigma^2 + O(\sigma^2). \end{aligned}$$

Here, when $\beta\sigma < 1$, we have

$$\|(1-\beta\sigma)d + \beta\sigma\bar{d}\| \le \max(\|d\|, \|\bar{d}\|).$$

Hence, there exists $c_1 > 0$ such that, if $\sigma < 1/\beta$, the term $O(\sigma^2)$ above satisfies $O(\sigma^2) \le c_1\sigma^2$. Taking $\beta > 0$ large enough, we obtain, when $\beta\sigma < 1$,

$$g_i(y(\sigma)) \le \sigma(u_i - \sigma\beta\gamma/2).$$

Similarly, for i in J and $\beta\sigma < 1$, we have

$$g_i(y(\sigma)) = \sigma u_i + O(\sigma^2),$$

with $|O(\sigma^2)| \le c_2\sigma^2$ and c_2 independent on β. Hence, from the implicit function theorem, there exists $r(\sigma) = O(\sigma^2)$ such that $x(\sigma) = y(\sigma) + r(\sigma)$ satisfies $g_J(x(\sigma)) = \sigma u_J$; and, for β large enough and σ small, we will still obtain

$$g_i(x(\sigma)) \le \sigma(u_i - \sigma\beta\gamma/2), \qquad i \in \bar{I}.$$

Hence,

$$g_i(x(\sigma)) \le \sigma u_i, \qquad i \in \bar{I},$$

i.e.,

$$x(\sigma) = \bar{x} + \sigma d + O(\sigma^2)$$

is in $F(\sigma u)$ as desired. □

Remark 2.1. If (3) holds, using $\bar{d}$ we can construct a path $x(\sigma)$ with $x(\sigma) \in F(\sigma u)$. This proves in particular that $F(\sigma u) \neq \varnothing$.

3. The Case $\Lambda_1 = \emptyset$

In this part, we present a study of the perturbation problem when $\Lambda_1 = \emptyset$. A first result in this direction is the following theorem.

Theorem 3.1. If hypotheses (1)–(3) hold and $\Lambda_1 = \emptyset$, then

$$\limsup_{\sigma \searrow 0} [\inf P(\sigma u) - f(\bar{x})]/\sigma^{1/2} < 0.$$

Proof. As in Lemma 2.2, we construct a feasible path, but on a different basis. The homogeneous linear program

$$\min_{d \in \mathbb{R}^n} \nabla f(\bar{x})' d; \; g'(\bar{x}) d \overset{\bar{I}}{\ll} 0$$

has $d=0$ as a feasible point, but its infimum is $-\infty$; otherwise, $d=0$ would be a solution and Λ_1 would not be empty. Consequently, there exists d^0 satisfying

$$\nabla f(\bar{x})' d^0 = -1, \qquad g'(\bar{x}) d^0 \overset{\bar{I}}{\ll} 0.$$

For α given in $]0, 1]$, we consider the path

$$y(\sigma) = \bar{x} + \alpha \sigma^{1/2} d^0 + \sigma \bar{d}.$$

Here, $\bar{d}$ is given by Lemma 2.1. As $|\alpha| \leq 1$, we have

$$g_J(y(\sigma)) = g_J'(\bar{x})(\alpha \sigma^{1/2} d^0 + \sigma \bar{d}) + \alpha^2 O(\sigma) + O(\sigma^2)$$
$$= \alpha^2 O(\sigma) + O(\sigma^2),$$

with $|O(\sigma)| \leq c_1 \sigma$ and $O(\sigma^2) \leq c_1 \sigma^2$, for some $c_1 > 0$ not depending on α. Hence, there exists a mapping

$$\mathbb{R}^+ \to \mathbb{R}^n, \qquad \sigma \to r(\sigma),$$

with $r(\sigma) = \alpha^2 O(\sigma) + O(\sigma^2)$, such that the path $x(\sigma) = y(\sigma) + r(\sigma)$ satisfies

$$g_J(x(\sigma)) = 0.$$

For $i \in \bar{I}$, we similarly obtain

$$g_i(x(\sigma)) \leq \sigma(u_i - \gamma) + \alpha^2 O(\sigma) + O(\sigma^2).$$

If α is small enough, $-\sigma\gamma + \alpha^2 O(\sigma) < 0$. In that case, for $\alpha > 0$ small enough, $x(\sigma)$ is feasible for $P(\sigma u)$ and

$$[\inf P(\sigma u) - f(\bar{x})]/\sigma^{1/2} \leq [f(x(\sigma)) - f(\bar{x})]/\sigma^{1/2} \to -\alpha.$$

This proves the theorem. □

We now consider the weak second-order sufficiency condition. For this, we have to define the critical cone at $\bar{x}$ for problem P(0),

$$C = \{d \in \mathbb{R}^n;\ \nabla f(\bar{x})' d \leq 0;\ g'(\bar{x}) d \overset{\bar{I}}{\ll} 0\},$$

and the following mapping:

$$H(x, \lambda) := \lambda_0 \nabla^2 f(x) + \sum_{i=1}^{p} \lambda_i \nabla^2 g_i(x).$$

The weak second-order sufficiency conditions is as follows:

$$\text{for all } d \text{ in } C - \{0\}, \text{ there exists some multiplier } \lambda \text{ associated to } \bar{x} \text{ such that } d' H(\bar{x}, \lambda) d > 0. \tag{4}$$

Theorem 3.2. If hypotheses (1)–(4) hold and if $\Lambda_1 = \emptyset$, then any solution $x(\sigma)$ of problem $P(\sigma u)$ satisfies

$$x(\sigma) = \bar{x} + O(\sigma^{1/2}).$$

Proof. Let $x(\sigma)$ be a solution of $P(\sigma u)$. We may always write $x(\sigma) = \bar{x} + \alpha(\sigma) d(\sigma)$, with $\alpha(\sigma) \in \mathbb{R}^+$ and $\|d(\sigma)\| = 1$. If, for a given value of σ, $\alpha(\sigma) = 0$, the desired estimate is obtained for that value of σ; otherwise, we have

$$g(x(\sigma)) = g(\bar{x}) + \alpha(\sigma) g'(\bar{x}) d(\sigma) + [\alpha(\sigma)^2/2] d(\sigma)' \nabla^2 g(\bar{x}) d(\sigma) + O(\alpha(\sigma)^3) \overset{\bar{I}}{\ll} \sigma u.$$

Let λ be a degenerate multiplier associated to $\bar{x}$. Multiplying the above equalities or inequalities by λ_i and summing over i, we obtain, as $\lambda_i \geq 0$,

$$[\alpha(\sigma)^2/2] d(\sigma)' H(\bar{x}, \lambda) d(\sigma) + O(\alpha(\sigma)^3) \leq \sigma \lambda' u;$$

hence, if $\alpha(\sigma) \neq 0$,

$$d(\sigma)' H(\bar{x}, \lambda) d(\sigma) \leq [2\sigma/\alpha(\sigma)^2] \lambda' u + O(\alpha(\sigma)).$$

If the conclusion does not hold, we have, for some d limit point of $d(\sigma)$ (hence $\|d\| = 1$), when $\sigma \searrow 0$,

$$d' H(\bar{x}, \lambda) d \leq 0, \qquad \text{for all } \lambda \text{ in } \Lambda_0. \tag{5}$$

But from

$$g(x(\sigma)) = \alpha(\sigma) g'(\bar{x}) d(\sigma) + O(\alpha(\sigma)) \overset{\bar{I}}{\ll} \sigma u,$$

we deduce that

$$g'(\bar{x})d \overset{\bar{I}}{\ll} 0.$$

Also Theorem 3.1 implies that, for σ small enough,

$$0 \geq [f(x(\sigma)) - f(\bar{x})]/\alpha(\sigma) = \nabla f(\bar{x})'d(\sigma) + O(\alpha(\sigma)),$$

hence $\nabla f(\bar{x})'d \leq 0$. This implies that d is in C. As $d \neq 0$ and $\|d\| = 1$, Inequality (5) is in contradiction with (4). This proves the theorem. □

We note that the conclusion of Theorem 3.1 implies that

$$\liminf \|x(\sigma u) - \bar{x}\|/\sigma^{1/2} > 0,$$

whenever $x(\sigma u) \in S(\sigma u)$ and $\sigma \searrow 0$. Hence, from Theorem 3.1 and Theorem 3.2, we deduce the following corollary.

Corollary 3.1. Under the hypotheses of Theorem 3.2, there exists $m > 0$, $M > 0$ such that the solutions $\bar{x}(\sigma u)$ of $\mathrm{P}(\sigma u)$ satisfy, for $\sigma > 0$ small enough,

$$m\sigma^{1/2} \leq \|\bar{x}(\sigma u) - \bar{x}\| \leq M\sigma^{1/2}.$$

We now try to characterize the limit points of $\sigma^{-1/2}(\bar{x}(\sigma u) - \bar{x})$, with $\bar{x}(\sigma u) \in S(\sigma u)$. Let d^0 be such a limit point. Expanding f and g up to the first order and using Theorem 3.1, we find, using the arguments in the proof of Theorem 3.2, that d^0 is a critical direction, as observed in Ref. 8. Also, normalizing if necessary the multipliers associated to $\bar{x}(\sigma u)$, we find that, for the extracted sequence $\{\sigma^k\}$ corresponding to d^0, they have a limit point λ^0 in Λ_0. If $\lambda_i^0 \neq 0$, the constraint i is active for σ^k small, hence $g_i'(\bar{x})d^0 = 0$. In summary, d^0 is an element of the following set:

$$C^0 = \{d \in C, d \neq 0; \exists \lambda \in \Lambda_0; \lambda_i > 0 \Rightarrow g_i'(\bar{x})d = 0\}.$$

For any d in C^0, we denote

$$I(d) = \{i \in \bar{I}; g_i'(\bar{x})d = 0\}.$$

We now consider the following subproblem:

$$(\mathrm{SP}(u)) \qquad \min_{d^0, d^1} \quad \nabla f(\bar{x})'d^0,$$

$$\text{s.t.} \quad d^0 \in C^0,\ d^1 \in \mathbb{R}^n,$$

$$g'(\bar{x})d^1 + (1/2)(d^0)'\nabla^2 g(\bar{x})d^0 \overset{I(d^0)}{\ll} u.$$

Here, $d'\nabla^2 g(\bar{x})d$ is a vector whose ith component is $d'\nabla^2 g_i(\bar{x})d$. We actually prove that any limit point d^0 of $\sigma^{-1/2}(\bar{x}(\sigma u)-\bar{x})$ is associated to some d^1 such that (d^0, d^1) is a solution of SP(u). There may be some parasitic solutions to SP(u), i.e., solutions that do not correspond to a limit point of the type $\sigma^{-1/2}(\bar{x}(\sigma u)-\bar{x})$, but we prove that any such parasitic solution gives rise to a suboptimal path.

Theorem 3.3. Under the hypotheses of Theorem 3.2, the following holds:

(i) Any d^0, limit point of $\sigma^{-1/2}(\bar{x}(\sigma u)-\bar{x})$, with $\bar{x}(\sigma u)\in S(\sigma u)$, is associated to some d^1 in $\mathbb{R}^n$ such that (d^0, d^1) is a solution of SP(u).

(ii) To each feasible solution d^0, d^1 in SP(u) is associated a path $x(\sigma)=\bar{x}+\sigma^{1/2}d^0+o(\sigma^{1/2})$ with $x(\sigma)\in F(\sigma u)$. If in addition (d^0, d^1) is a solution of SP(u), then

$$f(x(\sigma))=\inf(\mathrm{P}(\sigma u))+o(\sigma^{1/2}).$$

Proof. (a) Let d^0 be the limit of $(\sigma^k)^{-1/2}(x^k-\bar{x})$ with $\sigma^k \searrow 0$ and $x^k\in S(\sigma^k u)$. We already know that $d^0\in C^0$. We may write

$$x^k=\bar{x}+(\sigma^k)^{1/2}(d^0+\delta^k),$$

with $\delta^k\in\mathbb{R}^n$, $\delta^k\to 0$. Expanding the constraints up to the first order in σ, we get

$$g_i(x^k)=(\sigma^k)^{1/2}g_i'(\bar{x})\delta^k+(\sigma^k/2)(d^0)'\nabla^2 g_i(\bar{x})d^0+o(\sigma^k)\overset{I(d^0)}{\ll}\sigma u_i,$$

hence

$$g_i'(\bar{x})[\delta^k/(\sigma^k)^{1/2}]+(1/2)(d^0)'\nabla^2 g_i(\bar{x})d^0+[o(\sigma^k)/\sigma^k]\overset{I(d^0)}{\ll}u_i.$$

But the set

$$\{v\in\mathbb{R}^p;\ \exists d^1;\ g'(\bar{x})d^1\overset{I(d^0)}{\ll}v\}$$

is closed; see, e.g., Bonnans and Launay (Ref. 11, Part 4). Therefore, $u-\frac{1}{2}(d^0)'\nabla^2 g(\bar{x})d^0$ is in this set. We conclude that there exists $d^1\in\mathbb{R}^n$ such that (d^0, d^1) is feasible for SP(u).

(b) We now prove the first statement of (ii), and the end of the proof. For this purpose, to a feasible solution (d^0, d^1) of SP(u), we associate the path

$$x(\sigma)=\bar{x}+(\sigma-\sigma^{1.3})^{1/2}d^0+(\sigma-\sigma^{1.3})d^1+\sigma^{1.3}\bar{d}+r(\sigma),$$

where $\bar{d}$ is again given by Lemma 2.1 and $r(\sigma)$ is a small correction term. If $r(\sigma)=0$, we get, using the fact that (d^0, d^1) is feasible for SP(u),

$$\begin{aligned} g_J(x(\sigma)) &= (\sigma - \sigma^{1.3})[g_J'(\bar{x})d^1 + (1/2)(d^0)'\nabla^2 g_J(\bar{x})d^0] \\ &\quad + \sigma^{1.3} g_J'(\bar{x})\bar{d} + O(\sigma^{3/2}) \\ &= \sigma u_J + O(\sigma^{3/2}). \end{aligned}$$

As $\{\nabla g_i(\bar{x})\}_{i\in J}$ is linearly independent, we may find $r(\sigma) = O(\sigma^{3/2})$ such that $g_J(x(\sigma)) = \sigma u_J$. Expanding $g_i(x(\sigma))$ as before, for $i \in I(d^0)$ we find that

$$g_i(x(\sigma)) \leq \sigma u_i - \sigma^{1.3}\gamma + O(\sigma^{3/2});$$

hence, $g_i(x(\sigma)) < \sigma u_i$ for σ small. If $i \in \bar{I} - I(d^0)$, then $g_i'(\bar{x})d^0 < 0$, so that $g_i(x(\sigma)) < \sigma u_i$ is also satisfied. We have proved that $x(\sigma) \in F(\sigma u)$ when σ is small. In addition, we obviously have

$$f(x(\sigma)) = f(\bar{x}) + \sigma^{1/2}\nabla f(\bar{x})'d^0 + o(\sigma^{1/2}).$$

This result, Theorem 3.1, and point (a) of the proof imply that $\nabla f(\bar{x})'d^0$ must be minimum when d^0 is as in point (i). Conversely, if (d^0, d^1) is any solution of SP(u), we deduce the last statement of Theorem 3.3 with the above expansion of $f(x(\sigma))$. □

Remark 3.1.

(i) From Theorem 3.3, we deduce that, when $\sigma \searrow 0$,

$$\sigma^{-1/2}(\inf P(\sigma u) - \inf(P(0)) \to \inf(SP(u)).$$

(ii) The proof of Theorem 3.3 gives us a means to compute effectively (e.g., numerically) a suboptimal path.

(iii) If all solutions of SP(u) have the same component d^0, then $\sigma^{1/2}(x(\sigma) - \bar{x})$ converges toward d^0 whenever $x(\sigma)$ is in $S(\sigma u)$.

(iv) Assume that the hypotheses of Theorem 3.3 hold, except perhaps the second-order sufficiency condition. Using part (b) of the proof of Theorem 3.3, we deduce that

$$\lim\sup(P(\sigma u) - P(0))/\sigma^{1/2} \leq \inf SP(u).$$

In particular, if $\inf SP(u) = -\infty$, we deduce that, in this case,

$$\lim_{\sigma \searrow 0} [\inf P(\sigma u) - \inf P(0)]/\sigma^{1/2} = -\infty.$$

(v) Rockafellar (Ref. 13) observed that, if a convex program has no Lagrange multiplier, then the directional derivative of the cost must be $-\infty$ in some directions. Clarke (Ref. 14) defines calmness of P(0) as

$$\lim_{u \to 0} \inf(P(u) - P(0))/\|u\| > -\infty.$$

Observing that calmness is equivalent to the requirement that, for some $r > 0$, the minimum of the penalty function

$$f(x) + r[\|g_I(x)^+\| + \|g_J(x)\|]$$

is attained at $\bar{x}$, he deduces that calmness is not satisfied if no Lagrange multiplier exists. Extensions of this result are given in Rockafellar (Ref. 14). In Gollan (Ref. 7), we find the more precise result that

$$\lim_{\sigma \searrow 0} [\mathrm{P}(\sigma u) - \mathrm{P}(0)]/\sigma = -\infty$$

if (3) is satisfied.

4. The Case $\Lambda_1 \neq \emptyset$

We assume in this section that $\Lambda_1 \neq \emptyset$. Let us define

$$\Lambda^*(u) = \{\lambda \in \Lambda_1;\ -\lambda' u = \max\{-\lambda' u,\ \lambda \in \Lambda_1\}\}.$$

As observed by Gauvin and Janin (Ref. 8), if $\Lambda_1 \neq \emptyset$, hypothesis (3) is equivalent to the requirement that $\Lambda^*(u)$ is nonempty and bounded. We will use the strengthened second-order sufficiency condition [see Shapiro (Ref. 4)]:

$$\text{for any } d \text{ in } C - \{0\}, \text{ there exists } \lambda \text{ in } \Lambda^*(u) \text{ s.t. } d' H(\bar{x}, \lambda) d > 0. \tag{6}$$

The stability of the perturbed solution (Theorem 4.1 below) under such hypothesis has already been obtained by Shapiro (Ref. 4) when the hypothesis of Mangasarian and Fromovitz holds and by Gauvin and Janin (Ref. 10) under the hypothesis of Gollan.

Theorem 4.1. We assume that (1)–(3) and (6) hold and that $\Lambda_1 \neq \emptyset$. Then, if $x(\sigma)$ is in $S(\sigma u)$ and $\sigma \to 0$, one has

$$\limsup_{\sigma \searrow 0} \|x(\sigma u) - \bar{x}\|/\sigma < +\infty.$$

Our goal is now to compute the directional derivatives of the solutions. We define the Lagrangian associated to P(0) as

$$L(x, \lambda) := f(x) + \lambda' g(x).$$

We quote the following result.

Lemma 4.1. Under the hypotheses of Theorem 4.1, the following holds:

(i) If K is a limit point of the set of active constraints of the solutions of $P(\sigma u)$ when $\sigma \searrow 0$, then $\bar{x}$ satisfies the MF hypotheses for the set K.

(ii) The set of Lagrange multipliers associated to $x(\sigma u)$ remains bounded and its limit points are in $\Lambda^*(u)$.

(iii) For $\sigma > 0$ small enough, one has

$$f(x(\sigma u)) = \max\{L(x(\sigma u), \lambda) - \sigma\lambda' u; \lambda \in \Lambda^*(u)\}.$$

Proof. Let $\bar{d}$ be given by Lemma 2.1. Define

$$w(\sigma) := \bar{d} - (x(\sigma u) - \bar{x})/\sigma.$$

If $x(\sigma u)$ has K as a set of active constraints, then for all i in K, by Theorem 4.1,

$$\sigma u_i = g_i(x(\sigma u)) = g_i'(\bar{x})(x(\sigma u) - \bar{x}) + o(\sigma).$$

Hence,

$$g_i'(\bar{x})w(\sigma) = g_i'(\bar{x})\bar{d} - (\sigma u_i + o(\sigma))/\sigma \leq -\gamma + o(\sigma)/\sigma.$$

We may consider a sequence $\sigma^k \searrow 0$. Then, $\{w(\sigma^k)\}$ is bounded by Theorem 4.1. Any limit point w satisfies

$$g_i'(\bar{x})w \leq -\gamma, \qquad i \in K,$$

and similarly we can prove that $g_j'(\bar{x})w = 0$. This proves (i).

From (i) and the fact that the MF condition is stable by perturbation [see Robinson (Ref. 15)], we deduce that the set of Lagrange multipliers associated to $x(\sigma u)$ is nonempty and uniformly bounded when $\sigma \searrow 0$. If $\lambda(\sigma)$ is such a multiplier, we have

$$\lambda_i(\sigma)(g_i(x)(\sigma u) - \sigma u_i) = 0, \qquad \text{for all } i \text{ in } I.$$

Hence, if $\bar{\lambda}$ is a limit point of $\{\lambda(\sigma)\}$, we also have (discussing the different cases)

$$\bar{\lambda}_i(g_i(x(\sigma u)) - \sigma u_i) = 0.$$

Consequently, for this Lagrange multiplier,

$$f(x(\sigma u)) = L(x(\sigma u), \bar{\lambda}) - \sigma\bar{\lambda}' u. \tag{7}$$

On the other hand, if λ is a Lagrange multiplier, from $\lambda_I \geq 0$ we deduce that

$$\lambda_i(g_i(x(\sigma u)) - \sigma u) \leq 0, \qquad \text{for all } i \text{ in } I \cup J;$$

hence,

$$f(x(\sigma u)) \geq L(x(\sigma u), \lambda) - \sigma\lambda' u, \qquad \forall \lambda \in \Lambda_1.$$

This and (7) prove that

$$f(x(\sigma u), \lambda) = \max\{L(x, \lambda) - \sigma\lambda' u;\ \lambda \in \Lambda_1\}.$$

Hence, to prove (ii) and (iii), we just have to prove that $\bar{\lambda}$ is in $\Lambda^*(u)$.

From Theorem 4.1, relation (7), and the first-order optimality system of (P), we deduce that

$$f(x(\sigma u)) = f(\bar{x}) - \sigma\bar{\lambda}' u + O(\sigma^2);$$

hence, for the values of σ for which $x(\sigma u)$ has K as a set of active constraints,

$$\lim_{\sigma \searrow 0} (f(x(\sigma u)) - f(\bar{x}))/\sigma = -\bar{\lambda} u,$$

by Theorem 4.1. Notice that $(x(\sigma u) - \bar{x})/\sigma$ has at least a limit point $\tilde{d}$ which is a feasible point of the linear program

$$\min_d \nabla f(\bar{x})' d, \qquad g'(\bar{x}) d \overset{\bar{I}}{\ll} u. \tag{8}$$

But the dual of (8) is

$$\max -\lambda' u, \qquad \lambda \in \Lambda_1.$$

The relation

$$(d/d\sigma) f(x(\sigma u)) = \nabla f(\bar{x})' \tilde{d} = -\bar{\lambda}' u$$

hence implies that $\tilde{d}$ is a solution of (8) and that $\bar{\lambda}$ is a solution of its dual, i.e., $\bar{\lambda}$ is in $\Lambda^*(u)$. □

The above relations will allow us to expand $f(x(\sigma u))$ up to the second order. Indeed, let us define

$$Q^*(d) := \max\{d' H(\bar{x}, \lambda) d;\ \lambda \in \Lambda^*(u)\}.$$

Then, we have the following proposition.

Proposition 4.1. Under the hypotheses of Theorem 4.1, let $\{\sigma^k\}$ be such that $\sigma^k \searrow 0$, that x^k is a solution of $\mathrm{P}(\sigma^k u)$, and that $(x^k - \bar{x})/\sigma^k \to d$. Then,

$$F(x^k) = f(\bar{x}) - \sigma^k \lambda' u + [(\sigma^k)^2/2] Q^*(d) + o((\sigma^k)^2).$$

Proof. From Theorem 4.1, Lemma 4.1(iii), and the boundedness of $\Lambda^*(u)$, we get, for any $\bar{\lambda}$ in $\Lambda^*(u)$,

$$F(x^k)=\max\{f(\bar{x})+(1/2)(x^k-\bar{x})'H(\bar{x},\lambda)(x^k-\bar{x}),\ \lambda\in\Lambda^*(u)\}$$
$$-\sigma^k\bar{\lambda}'u+o((\sigma^k)^2).$$

With $x^k-\bar{x}=\sigma^k d+o(\sigma^k)$, we get the result. □

We now try to characterize the limit points of $\sigma^{-1}(\bar{x}(\sigma u)-\bar{x})$ when $\sigma\searrow 0$, with $\bar{x}(\sigma u)\in S(\sigma u)$. We consider the following sets:

$$F^*(u):=\{x\in F(u);\ \exists\lambda\in\Lambda^*(u);\ \lambda_i>0\Rightarrow g_i(x)=u_i,\ i\in I\},$$
$$C^*(u):=\{d\in\mathbb{R}^n;\ d \text{ solves } (8)\}.$$

By Lemma 4.1, we have $S(\sigma u)\subset F^*(\sigma u)$ for σ small. The set $C^*(u)$ can be characterized using the multipliers that are solutions of the dual of (8),

$$C^*(u)=\{d\in\mathbb{R}^n;\ g'(x)d\overset{\bar{I}}{\ll}u;\ \exists\lambda\in\Lambda^*(u);\ \lambda_i>0\Rightarrow g_i'(\bar{x})d=u_i,\ i\in I\}.$$

Here, we see that $C^*(u)$ is a candidate for the linearization of the set-valued mapping $u\to F^*(u)$ at $u=0$ and at point $\bar{x}$. We will use the following directional qualification property:

$$(\mathrm{DQ}^*(u))\qquad \forall d\in C^*(u),\ \exists x(\sigma u)=\bar{x}+\sigma d+o(\sigma),\ x(\sigma u)\in F^*(\sigma u).$$

Theorem 4.2. Under the hypotheses of Theorem 4.1 and $\mathrm{DQ}^*(u)$, the following holds:

(i) Let $\tilde{d}$ be a limit point of $(\sigma)^{-1}(\bar{x}(\sigma u)-\bar{x})$, with $\bar{x}(\sigma u)\in S(\sigma u)$. Then, $\tilde{d}$ is a solution of the pseudo-quadratic problem

$$\min Q^*(d),\qquad d\in C^*(u). \tag{9}$$

(ii) Conversely, to each solution $\tilde{d}$ of this problem is associated a path $x(\sigma u)=\bar{x}+\sigma\tilde{d}+o(\sigma)$, with $x(\sigma u)\in F^*(\sigma u)$ and

$$f(x(\sigma u))=\inf \mathrm{P}(\sigma u)+o((\sigma)^2).$$

Proof. Let $\tilde{d}$ be as in point (i). Then, $\tilde{d}$ is in $C(u)$. From the proof of Lemma 4.1, we get

$$\nabla f(\bar{x})'\tilde{d}=-\bar{\lambda}'u,\qquad \text{for any } \bar{\lambda} \text{ in } \Lambda^*(u).$$

This implies that $\tilde{d}$ is in $C^*(u)$, hence is feasible for (9).

Let d be in $C^*(u)$, and let $x(\sigma)$ be the path associated by $\mathrm{DQ}^*(u)$. Expanding $f(x(\sigma))$ as in Proposition 4.1, we get [here again, $\bar{\lambda}$ is any element of $\Lambda^*(u)$]

$$f(x(\sigma))=f(\bar{x})-\sigma\bar{\lambda}'u+[(\sigma)^2/2]Q^*(d)+o((\sigma)^2).$$

With Proposition 4.1, this implies that $Q^*(\cdot)$ must be minimum at $\tilde{d}$. The rest of the theorem follows. □

Remark 4.1. Assume that $\Lambda^*(u)$ reduces to a singleton $\{\bar{\lambda}\}$, which will be the case for almost all directions u (for the Lebesgue measure). In this case, (9) is a true quadratic problem. Applying the characterization of Kyparisis (Ref. 16) to the problem (8) and defining

$$K := J \cup \{i \in I;\ \bar{\lambda}_i > 0\},$$

we see that, d^0 being in $C^*(u)$, the following is satisfied:

(i) $\{\nabla g_i(\bar{x}),\ i \in K\}$ is linearly independent.
(ii) $\exists d^\# \in \ker g'_K(\bar{x});\ g'_i(\bar{x})d^\# < 0,\ \forall i \in I(d^0) - K$.

From this, it can be easily deduced that $CQ^*(u)$ holds; hence, we may apply Theorem 4.2. In this case, our Theorem 4.2 reduces to Theorem 5.1 of Gauvin–Janin (Ref. 8).

Remark 4.2. We may compare Theorem 4.2 with the corresponding results of Shapiro (Ref. 4) in which the hypothesis of Mangasarian and Fromovitz is assumed to hold.

5. Examples

We give two two-dimensional examples for which the hypothesis of Mangasarian and Fromovitz is not satisfied, the first with $\Lambda_1 = \emptyset$. We show how our theorems allow us to compute effectively suboptimal paths on these examples.

Example 5.1. The problem is

$$\min \quad x_1,$$

$$\text{s.t.} \quad x_1^2 - x_2 \le u_1, \qquad x_1^2 + x_2 \le u_2.$$

One has $F(0) = \{0\}$. Hence, the original problem has the unique solution $\bar{x} = 0$ and (2) is easy to check. We easily obtain $\Lambda_1 = \emptyset$ and

$$\Lambda_0 = \left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix};\ \alpha \in \mathbb{R}^{-*} \right\}.$$

Consequently, (3) will be satisfied iff $u_1 + u_2 > 0$. We assume that this is the case. We also obtain

$$C = \{d \in \mathbb{R}^2;\ d_1 \le 0,\ d_2 = 0\}, \qquad (1/2)d'H(\bar{x}, \lambda)d = 2\alpha d_1^2,$$

with $\lambda=(\alpha,\alpha)'$, $\alpha>0$. This implies (4). We may apply Theorem 3.3. Here, $C^0=C$ and $I(d)=\{1,2\}$, for all d in C^0. Hence, SP(u) reduces to

$$\min_{d^0,d^1} \quad d_1^0,$$

$$\text{s.t.} \quad d_1^0\leq 0, \qquad d_2^0=0, \qquad (d_1^0)^2-d_2^1\leq u_1, \qquad (d_1^0)^2+d_2^1\leq u_2.$$

The two last inequalities are equivalent to

$$(d_1^0)^2\leq \min(u_1+d_2^1,\, u_2-d_2^1).$$

In order to get d_1^0 minimum, we have to maximize the right-hand side with respect to d_2. The maximum is obtained when

$$d_2^1=(1/2)(u_2-u_1),$$

d_1^1 having any value and then

$$d_1^0=-\sqrt{[(u_2+u_1)/2]}$$

is optimal. Hence, there is a unique component d^0 for the solutions of SP(u). Consequently, the solutions $\bar{x}(\sigma u)$ of the perturbed problem satisfy

$$\bar{x}(\sigma u)=\bar{x}-\sigma^{1/2}\begin{bmatrix}\sqrt{[(u_1+u_2)/2]}\\ 0\end{bmatrix}+o(\sigma^{1/2}).$$

We may check this result by computing the exact solution of the perturbed problem. They saturate the two constraints, hence there is a unique solution

$$\bar{x}(\sigma u)=\begin{bmatrix}-\sigma^{1/2}\sqrt{[(u_1+u_2)/2]}\\ \sigma(u_2-u_1)/2\end{bmatrix}.$$

Example 5.2. The problem is

$$\min \quad x_2,$$

$$\text{s.t.} \quad x_1^2-x_2\leq u_1, \qquad x_1^2+x_2\leq u_2.$$

We only changed the criterion. Again, $\bar{x}=0$, (1) and (2) are satisfied, Λ_0 is as before, and (3) is satisfied iff $u_1+u_2>0$, which is assumed to hold. Now,

$$\Lambda_1=\left\{\lambda=\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix};\ \lambda_1-\lambda_2=1,\ \lambda_1\geq 0,\ \lambda_2\geq 0\right\}$$

$$=\left\{\lambda=\begin{pmatrix}1+\lambda_2\\ \lambda_2\end{pmatrix};\ \lambda_2\geq 0\right\},$$

and $\Lambda^*(u)$ is the set of element of the form $(1+\lambda_2, \lambda_2)'$ for which λ_2 is solution of

$$\min(1+\lambda_2)u_1+\lambda_2 u_2, \qquad \lambda_2 \geq 0.$$

As $u_1+u_2>0$, we get

$$\Lambda^*(u)=\{\bar{\lambda}\}, \qquad \text{with } \bar{\lambda}=(1, 0)'.$$

Hence, $CQ^*(u)$ holds. Also,

$$C=\{d\in\mathbb{R}^2; d_2=0\},$$

$$C^*(u)=\{d\in\mathbb{R}^2; -d_2=u_1, d_2\leq u_2\}, \qquad (1/2)d'H(\bar{x}, \bar{\lambda})d=d_1^2.$$

Hence, (6) holds. We may apply Theorem 4.2. The quadratic problem to be considered is

$$\min \quad d_1^2,$$

$$\text{s.t.} \quad -d_2=u_1, \qquad d_2\leq u_2.$$

This problem has a unique solution $(0, -u_1)'$. We have proved that any optimal solution $\bar{x}(\sigma u)$ of $P(\sigma u)$ satisfies

$$\bar{x}(\sigma u)=\bar{x}+\sigma\begin{bmatrix}0\\-u_1\end{bmatrix}+o(\sigma).$$

In fact, the exact solution of $P(\sigma u)$ is just $(0, -\sigma u_1)'$.

Note Added in Proof

After acceptance of this paper I learned that the conclusion of Theorem 4.2 holds without the restrictive hypothesis $DQ^*(u)$ (Ref. 17). Also, the case with the existence of Lagrange multipliers and weak second-order conditions is dealt with in Ref. 18.

References

1. Fiacco, A. V., *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, New York, 1983.
2. Jittorntrum, K., *Solution Point Differentiability without Strict Complementarity in Nonlinear Programming*, Mathematical Programming Study, Vol. 21, pp. 127-138, 1984.

3. BONNANS, J. F., *A Semi-Strong Sufficiency Condition for Optimality in Nonconvex Programming and Its Connection to the Perturbation Problem*, Journal of Optimization Theory and Applications, Vol. 60, pp. 7–18, 1989.
4. SHAPIRO, A., *Sensitivity Analysis of Nonlinear Programs and Differentiability Properties of Metric Projections*, SIAM Journal on Control and Optimization, Vol. 26, pp. 628–645, 1988.
5. MANGASARIAN, O. L., and FROMOVITZ, S., *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints*, Journal of Mathematical Analysis and Applications, Vol. 7, pp. 37–47, 1967.
6. GOLLAN, B., *On the Marginal Function in Nonlinear Programming*, Mathematics of Operation Research, Vol. 9, pp. 208–221, 1984.
7. GOLLAN, B., *Perturbation Theory for Abstract Optimization Problems*, Journal of Optimization Theory and Applications, Vol. 35, pp. 417–441, 1981.
8. GAUVIN, J., and JANIN, R., *Directional Behavior of Optimal Solutions in Nonlinear Mathematical Programming*, Mathematics of Operations Research, Vol. 13, pp. 629–649, 1988.
9. GAUVIN, J., and JANIN, R., *Directional Derivative of the Value Function in Parametric Optimization*, Annals of Operations Research, Vol. 27, pp. 237–252, 1990.
10. GAUVIN, J., and JANIN, R., *Directional Lipschitzian Optimal Solutions and Directional Derivative for the Optimal Value Function in Nonlinear Mathematical Programming*, Analyse Non-Linéaire, Edited by H. Attouch, J. P. Aubin, F. Clarke, and I. Ekeland, Gauthiers-Villars, Paris, France, pp. 305–324, 1989.
11. BONNANS, J. F., and LAUNAY, G., *On the Stability of Sets Defined by a Finite Number of Equalities and Inequalities*, Journal of Optimization Theory and Applications, Vol. 70, pp. 417–428, 1991.
12. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
13. CLARKE, F. H., *A New Approach to Lagrange Multipliers*, Mathematics of Operations Research, Vol. 1, pp. 165–174, 1976.
14. ROCKAFELLAR, R. T., *Lagrange Multipliers and Subderivatives of Optimal Value Functions in Nonlinear Programming*, Mathematical Programming Study, Vol. 17, pp. 28–66, 1982.
15. ROBINSON, S. M., *Stability Theory for Systems of Inequalities, Part 2: Differentiable Nonlinear Systems*, SIAM Journal on Numerical Analysis, Vol. 13, pp. 497–513, 1976.
16. KYPARISIS, J., *On Uniqueness of Kuhn–Tucker Multipliers in Nonlinear Programming*, Mathematical Programming, Vol. 32, pp. 242–246, 1985.
17. AUSLENDER, A., and COMINETTI, R., *First and Second Order Sensitivity Analysis of Nonlinear Programs under Directional Constraint Qualification Condition*, Optimization, Vol. 31, pp. 351–363, 1990.
18. BONNANS, J. F., JOFFE, A. D., and SHAPIRO, A., *Expansion of Exact and Approximate Solutions in Nonlinear Programming*, French–German Conference in Optimization, Edited by D. Pallaschke, Lecture notes in Economics and Mathematical Systems, Springer Verlag, Berlin (to appear).