Journal of Convex Analysis
Volume 2 (1995), No.1/2, 41–57

Quadratic Growth and Stability in Convex Programming Problems with Multiple Solutions

J. Frédéric Bonnans
INRIA-Rocquencourt, Domaine de Voluceau,
B.P. 105, 78153 Rocquencourt, France.
e-mail: frederic.bonnans@inria.fr

Alexander D. Ioffe[1]
Department of Mathematics, Technion Israel Institute of Technology,
Haifa 3200 - Israel.
e-mail: ioffe@techunix.technicon.ac.il

Received 9 June 1994
Revised manuscript received 23 February 1995

Dedicated to R. T. Rockafellar on his 60th Birthday

Given a convex program with C^2 functions and a convex set S of solutions to the problem, we give a second order condition which guarantees that the problem does not have solutions outside of S. This condition is interpreted as a characterization for the quadratic growth of the cost function. The crucial role in the proofs is played by a theorem describing a certain uniform regularity property of critical cones in smooth convex programs. We apply these results to the discussion of stability of solutions of a convex program under possibly nonconvex perturbations.

Keywords : Convexity, Lagrangian, composite functions, critical cone, quadratic growth, stability, multiple solutions.

1. Introduction

It is well known that in nonlinear programming problems with qualified constraints the standard second order sufficient condition is equivalent to the following estimate for the cost function f_0 on the feasible set:

$$f_0(x) \geq f_0(x) + r \cdot \operatorname{dist}^2(x, x_*),\ \forall x \text{ in a neighborhood of } x_*$$

where $r > 0$ and x_* is the solution (e.g. [2]). The latter property which we call the second order growth condition can easily be extended to problems without constraint qualification and with multiple non–isolated solutions. In either case the second order growth condition

[1] Supported by the US–Israel Binational Science Foundation grant no. 90–00455

ISSN 0944-6532 / \$ 2.50

is instrumental in obtaining many important results of sensitivity analysis (e.g. estimates for directional derivatives of the value function or for Hölder stability of solutions) [3, 4, 10, 15].

However, if the problem has a solution set with non-isolated points, no simple characterization of the growth condition is known (see [2] for details), except in the case of linearly independent gradients of active constraints [18]. The purpose of this paper is to show that at least for one important class of problems, namely problems of smooth convex programming, a simple second order characterization of the growth condition can be given. This characterization seems to be new and of an interest.

It is well known furthermore that to get Lipschitz stability of solutions of the perturbed problem a strengthening of the standard second order sufficient condition is needed in the regular case (qualified constraints and an isolated solution) [5, 15]. It was shown in [10] that this strengthened second order condition can also be expressed in terms of a "strong" second order growth condition that can be likewise extended to non–qualified problems with non–isolated solutions. We show further in this paper that in smooth convex problems (with arbitrary, not necessarily convex, dependence on the parameter) a simple second order characterization exists also for the strong growth condition.

As in [2] we start with the problem of minimizing the function

$$f(x) := \max_{1\leq i\leq m} f_i(x),$$

where $x \in \mathbb{R}^n$. Let S denote a *compact* set of minimum points of f : $S \subset \operatorname{argmin} f$. We assume that $S \neq \emptyset$ and denote by c the minimal value of f. We say that f satisfies the *quadratic growth condition* on S if

$$\exists\beta > 0 \ : \ f(x) \geq c + \beta \operatorname{dist}(x, S)^2, \ \forall x \text{ in a neighborhood of } S. \tag{QGC}$$

Of course, if (QGC) holds, then, at least locally, S coincides with the set of minima of f. The main result (Theorem 2.3) gives a complete characterization of (QGC) in case when all functions are convex and twice continuously differentiable. The theorem is stated in section 2 and proved in section 5. In section 3 we consider problems depending on a parameter and prove Theorem 3.4 containing a characterization for the strong growth condition.

In section 4, on the way to the proof of Theorem 2.3, we establish an estimate to the set of solution of a system of linear inequalities $a_i(x) \cdot h \leq 0$ in case when $a_i(x)$ are gradients of convex C^1 functions. This result, for this specific case, extends the well known Robinson's estimate for the distance to a parametrically dependent set of solutions of linear inequalities [13] in the sense that no regularity (Mangasarian-Fromovitz, or Slater) assumption at the reference point is made. Finally in section 6 we consider constrained optimization problems and state the analogues of the two main theorems in this case.

2. Characterization of quadratic growth

From now on we assume that S is a *convex compact* set and that $f_i(x)$, $i = 1, \cdots, m$ are convex. We denote by $\varphi'(x; d)$ the directional derivative of φ at x along the direction d and by $\mathcal{S}^m$ the standard simplex of $\mathbb{R}^m$, i.e.

$$\mathcal{S}^m := \{\lambda \in \mathbb{R}^m_+ \ ; \ \sum_{i=1}^m \lambda_i = 1\}.$$

We start by introducing some basic notation and recalling some known facts. With each $x \in S$, we associate the collection of active indices

$$I(x) := \{i = 1, \cdots, m \ : \ f_i(x) = f(x)\},$$

the critical cone at x

$$C(x) := \{d \in R^n \ : \ f_i'(x; d) \leq 0, \ \forall i \in I(x)\},$$

the contingent cone to S at x (the lim sup is taken in the sense of Painlevé-Kuratowski)

$$T_S(x) := \limsup_{t \downarrow 0} t^{-1}(S - x),$$

and the set of Lagrange multipliers:

$$\Lambda(x) := \{\lambda \in \mathcal{S}^m \ : \ \lambda_i = 0, \ i \notin I(x), \ \sum_{i=1}^m \lambda_i f_i'(x; d) \geq 0, \ \forall d \in \mathbb{R}^n\}.$$

The Lagrangian associated with f is

$$\mathcal{L}(x, \lambda) := \sum_{i=1}^m \lambda_i f_i(x).$$

As S is convex, the set $T_S(x)$ coincides with the tangent cone to S at x in the sense of convex analysis, which is the collection of vectors d such that $\text{dist}(x + td, S) = o(t)$ when $t \searrow 0$.
As all functions f_i are convex, the condition $\sum \lambda_i f_i'(x; d) \geq 0$, $\forall d \in \mathbb{R}^n$, is equivalent to the fact that $\mathcal{L}(., \lambda)$ attains its minimum at x.
For any convex set D we denote by $\text{aff}\, D$ and $\text{ri}\, D$ the affine hull and relative interior of D, and by $\text{rb}D$ its relative boundary : $\text{rb}D := D \setminus \text{ri}\, D$.
The two propositions that follow contain known facts. We prove them mainly because the proofs are short and useful for further discussions.

Proposition 2.1. (e.g. [12]) *Let D be a convex set on which the convex function φ is constant. If $x \in \text{ri}\, D$, then $\partial\varphi(x) \subset \partial\varphi(u)$ for any $u \in D$. In particular, $\partial\varphi(x)$ is constant over* $\text{ri}\, D$.

Proof. Given $x \in \text{ri}\, D$, $y \in \partial\varphi(x)$ and $u \in D$, we have $\varphi(u) - \varphi(x) \geq y \cdot (u - x)$, i.e. $y \cdot (u - x) \leq 0$. This being valid for all $u \in D$, as $x \in \text{ri}\, D$, we must have $y \cdot u = y \cdot x$. It follows that, for any $x' \in \mathbb{R}^n$

$$\varphi(x') - \varphi(u) = \varphi(x') - \varphi(x) \geq y \cdot (x' - x) = y \cdot (x' - u),$$

i.e. $y \in \partial\varphi(u)$. We have proved that $\partial\varphi(x) \in \partial\varphi(u)$. As $x \in D$, it follows that $\partial\varphi(x) = \cap_{u\in D}\partial\varphi(u)$ is constant over $\operatorname{ri} D$. $\square$

Proposition 2.2. *Let the functions f_i be convex and let S be a convex set on which f is constant. Then there are sets I, C and T such that, for all $x \in \operatorname{ri} S$:*

$$I(x) = I,\ C(x) = C,\ T_S(x) = T.$$

Proof. We first prove that when $x \in \operatorname{ri} S$, then $I(x) = \cap_{x'\in S} I(x')$. Indeed, assume that $i \in I(x)$ and $i \notin I(x')$ for some $x' \in S$, then $f_i'(x, x'-x) \le f_i(x') - f_i(x) < 0$. It follows that $f_i'(x, x-x') > 0$, in contradiction to the fact that $f(x) = \max_{i\le 1\le n} f_i(x)$ is constant on some segment $[x, x+\epsilon(x-x')]$ with $\epsilon > 0$.
We prove that $C(x)$ is constant on $\operatorname{ri} S$. Pick $i \in I$, it follows from Proposition 2.1 that $\partial f_i(x)$ is constant over $\operatorname{ri} S$. Consequently, $f_i'(x;.)$ is constant over $\operatorname{ri} S$, whence $C(x)$ is itself constant over $\operatorname{ri} S$.
For $T_S(x)$ the statement is trivial: at every $x \in \operatorname{ri} S$ this is the linear subspace parallel to $\operatorname{aff} S$. $\square$

It is well known (e.g. [14]) that $\Lambda(x)$ does not actually depend on x, that is, there exists Λ such that, for all $x \in S$

$$\Lambda(x) \equiv \Lambda.$$

This can be easily verified if all functions f_i are continuously differentiable. Indeed, by Propositions 2.1 and 2.2 this is true for $x \in \operatorname{ri} S$. It is clear furthermore that $x \mapsto \Lambda(x)$ is an u.s.c. set-valued map, so it remains to check that $\lambda_j = 0$ whenever $x \in S$ and $j \in I(x)\backslash I$. Pick $x' \in \operatorname{ri} S :$ we have $f_j'(x, x'-x) \le f_j(x') - f_j(x) < 0$. As $f_i'(x, x'-x) = 0$ for $i \in I$, we conclude that the inequality $\sum \lambda_i f_i'(x, x'-x) \ge 0$ may hold for a $\lambda = (\lambda_1, \cdots, \lambda_m) \in \Lambda(x)$ only if $\lambda_j = 0$ for $j \in I(x)\backslash I$.

We can now formulate the main theorem. Set

$$\varphi_x(d) := \max_{\lambda\in\Lambda} \nabla_x^2 \mathcal{L}(x,\lambda)(d,d).$$

This function is strongly related to the second-order variation of the cost along a critical direction (see Section 4). Define the normal cone to S at x, which is the polar of $T_S(x)$, and (for $\varepsilon > 0$) the set of approximate normal and critical directions as

$$\begin{aligned} N_S(x) &:= \{v : v \cdot d \le 0,\ \forall d \in T_S(x)\}, \\ N_S^\varepsilon(x) &:= \{v : \operatorname{dist}(v, N_S(x)) \le \varepsilon\|v\|\}, \\ C^\varepsilon(x) &:= \{h : \operatorname{dist}(h, C(x)) \le \varepsilon\|h\|\}. \end{aligned}$$

Theorem 2.3. *Let S be a non–empty convex compact set on which f attains its minimum. Assume that all functions f_i are convex and twice continuously differentiable. Then the following properties are equivalent:*
(i) (QGC) *holds.*
(ii) *There exists $\beta > 0$ such that*

$$\varphi_x(d) \ge \beta \operatorname{dist}(d, T_S(x))^2,\ \forall x \in S,\ \forall d \in C(x)\backslash T_S(x);$$

(iii) *There exists* $\beta > 0$ *such that, for all* $\varepsilon \in (0, 1)$

$$\varphi_x(d) \geq (1-\varepsilon)^2 \beta \|d\|^2, \ \forall x \in S, \ \forall d \in C(x) \cap N_S^\varepsilon(x).$$

(iv) *There exist* $\beta > 0$, $\varepsilon > 0$ *such that*

$$\varphi_x(d) \geq \beta \|d\|^2, \ \forall x \in S, \ \forall d \in C^\varepsilon(x) \cap N_S(x).$$

By Theorem 2.3, the condition below is necessary for quadratic growth:

$$\varphi_x(d) > 0, \ \forall d \in C(x) \backslash T_S(x), \ \forall x \in S. \tag{2.1}$$

This condition, in turn, implies a first-order geometric condition on S, which is therefore itself a necessary condition for quadratic growth.

Proposition 2.4. *Under the assumption of* Theorem 2.3 *except for the compactness of* S, *the condition* (2.1) *implies*
(i) $C(x) \backslash T_S(x) \subset C \backslash T, \ \forall x \in S.$
(ii) $I(x) \backslash I \neq \phi$ *on* $\operatorname{rb} S$.

Proof. (i) Assuming the contrary, we find $x \in S$ and $d \in C(x) \backslash T_S(x)$ with $d \notin C \backslash T$. As $C(x) \subset C$, it follows that d must belong to $T \backslash T_S(x)$. Take a sequence of $x^k \in \operatorname{ri} S$ converging to x. As $d \in T$, we have $\varphi_{x^k}(d) = 0$, hence $\varphi_x(d) = \lim \varphi_{x^k}(d) = 0$ in contradiction with (2.1).
(ii) If S is a singleton, then the relation is trivially true as $\operatorname{rb} S$ is empty. Otherwise pick $u \in \operatorname{ri} S$, $x \in \operatorname{rb} S$, and set $d := x - u$. It is clear that $d \in T \backslash T_S(x)$; hence $d \notin C \backslash T$. If $I(x) = I$, then $C(x) = C \supset T$, whence $d \in C(x) \backslash T_S(x)$, in contradiction with (i). □

From (ii) we deduce in particular that if f is a C^2 function, and its set of minima S satisfies the quadratic growth condition, then $\operatorname{rb} S$ is empty, which means that S is an affine space.

3. Parametric problems

We consider the family of problems

$$\min_{x \in \mathbb{R}^n} f(x, \theta) \tag{P_θ}$$

where $x \in \mathbb{R}^n$, $\theta \in \mathbb{R}_+$ and $f(x,\theta) := \max_{i=1,\cdots,m} f_i(x,\theta)$, with each $f_i, i = 1, \cdots, m$ of class $\mathcal{C}^2$.
We view (P_θ), for $\theta > 0$, as a perturbation of the original problem (P_0). With (P_θ) is associated the Lagrangian

$$\mathcal{L}(\lambda, x, \theta) := \sum_{i=1}^{n} \lambda_i f_i(x, \theta).$$

We denote by $v(\theta)$, $S(\theta)$ the value function and the set of solutions of (P_θ).

We say that a mapping $\theta \mapsto x(\theta) : \mathbb{R}_+ \to \mathbb{R}^n$, is an $\varepsilon(\theta)$-optimal path if $f(x(\theta),\theta) \leq v(\theta) + \varepsilon(\theta)$. Here $\varepsilon(\theta)$ may take the value $o(\varepsilon)$, $O(\varepsilon^2)$, etc.. Similarly, we say that a sequence $\{x_n\}$ is $\varepsilon(\theta_n)$-optimal, where $\theta_n \searrow 0$, if $f(x_n,\theta_n) \leq v(\theta_n) + \varepsilon(\theta_n)$. Denote

$$\mu := \min_{x\in S} \max_{\lambda\in\Lambda} \mathcal{L}_\theta(\lambda, x, 0),$$

$$\bar{\Lambda}(x) := \{\lambda \in \Lambda : \mathcal{L}_\theta(\lambda, x, 0) \geq \mu\},$$

where $\mathcal{L}_\theta$ is the derivative of $\mathcal{L}$ with respect to the parameter. Let further $\bar{S}$ be the collection of points at which the above minimum is attained, that is

$$\overline{S} := \operatorname*{argmin}_{x\in S} \max_{\lambda\in\Lambda} \mathcal{L}_\theta(\lambda, x, 0).$$

As $\overline{S}$ is the set of minima of a continuous function over a compact set, it is nonempty and compact.

Proposition 3.1. (Gol'stein [6]) *Assume that $S(\theta)$ is non empty and uniformly bounded for θ close to* 0. *Then the value function $v(\theta) := \inf_{x\in\mathbb{R}^n} f(x,\theta)$ has a right derivative at 0 that is equal to μ. In addition, every $\varepsilon(\theta_n)$-optimal sequence $\{x_n\}$ is bounded and all its limit points belong to $\overline{S}$.*

The following standard estimate will be useful later.

$$\begin{cases} \text{There exists } a > 0 \text{ such that, if } \theta > 0 \text{ is small enough} \\ v(\theta) \leq v(0) + \mu\theta + a\theta^2. \end{cases} \tag{3.2}$$

We say (cf. [10]) that the *strong quadratic growth* condition is satisfied if

$$\begin{cases} \text{There exists } \beta > 0,\ \gamma > 0 \text{ such that, if } (x,\theta) \\ \text{is sufficiently close to } \overline{S} \times \{0\},\ \theta \geq 0, \text{ then} \\ f(x,\theta) \geq v(0) + \mu\theta + \beta \operatorname{dist}^2(x,S) - \dfrac{\gamma}{2}\theta^2. \end{cases} \tag{SQG}$$

Proposition 3.2. [10] *If the strong quadratic growth condition holds, then any $O(\theta^2)$-optimal path x_θ satisfies* $\operatorname{dist}(x_\theta, S) = O(\theta)$.

Proof. Combining (SQG) and (3.2), we get

$$\beta \operatorname{dist}^2(x_\theta, S) \leq (a+\gamma)\theta^2 + O(\theta^2) = O(\theta^2),$$

from which the conclusion follows. □

We observe that (SQG) can be interpreted as a quadratic growth condition for an auxiliary problem. Indeed, (SQG) is obviously equivalent to

$$f(x,\theta) - \mu\theta + \gamma\theta^2 - v(0) \geq \beta \operatorname{dist}^2(x,S) + \frac{\gamma}{2}\theta^2.$$

It follows that $\boldsymbol{S} := S \times \{0\}$ is the set of local minima of the quantity at the left-hand-side of the above inequality over $\mathbb{R}^n \times [0, \infty[$ or, equivalently, S is the set of local minima of

$$\mathbf{f}(x, \theta) := \max(f(x, \theta) - \mu\theta + \gamma\theta^2 - v(0), -\theta).$$

In order to state second order conditions we compute the critical cone and set of Lagrange multipliers associated with the problem of minimizing $\mathbf{f}$.

Lemma 3.3. *The set $\boldsymbol{S} := S \times \{0\}$ is a set of local minima of the function $f(x, \theta)$ with common value zero. The critical cone associated with $(x, 0) \in \boldsymbol{S}$ is is defined by*

$$\boldsymbol{C}(x) = \{(d, \eta);\ \ f'(x, 0)(d, \eta) - \mu\eta = 0;\ \ \eta \geq 0\}, \tag{3.3}$$

and the set of Lagrange multipliers $\Lambda(x)$ is defined by

$$\Lambda(x) = \{(\alpha\hat{\lambda}, 1 - \alpha) :\ \hat{\lambda} \in \bar{\Lambda}(x)\ ;\ \alpha = (\mathcal{L}_\theta(\hat{\lambda}, x) - \mu + 1)^{-1} \in (0, 1]\}.$$

Proof. The critical cone is the set of directions in which the directional derivative of the cost is null, and is easily checked to be characterized by (3.3). Writing the optimality system, we find that the Lagrange multipliers satisfy

$$\begin{cases} \Sigma_{i=1}^m \lambda_i \nabla_x f_i(x, 0) = 0, \\ \\ \Sigma_{i=1}^m \lambda_i \left(\frac{\partial f_i}{\partial \theta}(x, 0) - \mu \right) - \xi = 0, \\ \\ \Sigma_{i=1}^m \lambda_i + \xi = 1, \\ \\ \lambda \geq 0,\ \xi \geq 0,\ \lambda_i(f_i(x_i, 0) - v_0) = 0, \end{cases} \tag{3.4}$$

where ξ is the multiplier associated with the constraint $-\theta \leq 0$. Set $\alpha := \Sigma_{i=1}^m \lambda_i$. It follows from the second and third relations above that $0 < \alpha \leq 1$. Rewrite the three first relation as

$$\begin{cases} \nabla_x \mathcal{L}(\lambda, x) = 0, \\ \\ \mathcal{L}_\theta(\lambda, x) - \mu\alpha = \xi, \\ \\ \alpha + \xi = 1. \end{cases} \tag{3.5}$$

Then $\xi = 1 - \alpha$ and

$$\hat{\lambda} := \lambda/\alpha \in \Lambda \quad \text{and} \quad \alpha(\mathcal{L}_\theta(\hat{\lambda}, x) - \mu) = \xi \geq 0,$$

i.e. in fact $\hat{\lambda} \in \bar{\Lambda}(x)$; the result follows. $\square$

Note that for all $x \in S$, $\boldsymbol{C}(x) \supset C(x) \times \{0\}$ and

$$\begin{aligned} x \in \overline{S} &\Rightarrow \Lambda(x) = \bar{\Lambda}(x) \times \{0\}, \\ x \in S \backslash \overline{S} &\Rightarrow \boldsymbol{C}(x) = C(x) \times \{0\}. \end{aligned}$$

Our next result gives a characterization of (SQG) in terms of the second-order expansion of the data, even if the data are not jointly convex with respect to x and θ. Define

$$\bar{\varphi}_x(d) := \max_{\lambda \in \bar{\Lambda}(x)} \nabla^2_x \mathcal{L}(\lambda, \S, \prime)(\lceil, \rceil).$$

We say that the *strong second-order condition* is satisfied if

$$\begin{cases} \text{There exist } \beta > 0,\ \varepsilon > 0 \text{ such that} \\ \bar{\varphi}_x(d) \geq \beta \|d\|^2,\ \forall x \in \bar{S}^\varepsilon,\ \forall d \in C^\varepsilon(x) \cap N_S(x), \end{cases} \tag{SSOC}$$

where $\overline{S}^\varepsilon = \{x \in S : d(x, \overline{S}) \leq \varepsilon\}$.

We observe that, as $\bar{\varphi}_x(d) \leq \varphi_x(d)$, (SSOC) implies (QGC). If S is a singleton, then (SSOC) reduces to the strong second-order condition of Shapiro [15].

Theorem 3.4. *Conditions* (SSOC) *and* (SGQ) *are equivalent.*

Proof. We prove first that (SSOC) implies (SGQ). If (SQG) does not hold, then there exists $x_n \to x \in \overline{S}$, $\theta_n \downarrow 0$ with

$$f(x_n, \theta_n) < v(0) + \mu\theta_n + \frac{1}{n}\operatorname{dist}(x_n, S)^2 - n\theta_n^2. \tag{3.6}$$

Let u_n be the projection of x_n onto S and $d_n := {\theta_n}^{-1}(x_n - u_n)$.

Case 1. $\theta_n = o(\operatorname{dist}(x_n, S))$, for a subsequence. Extracting this subsequence if necessary we may assume that this relation holds for the entire sequence so that $\|d_n\| \to \infty$. From Theorem 4.1 we deduce that given any $\varepsilon > 0$, $d_n \in C^\varepsilon(u_n)$, $u_n \in S^\varepsilon$ for n large. As $\|d_n\| \to \infty$ we have for all $\lambda \in \bar{\Lambda}(u_n)$

$$\begin{aligned} f(x_n, \theta_n) &\geq \mathcal{L}(\lambda, x_n, \theta_n) \\ &= \mathcal{L}(\lambda, u_n, 0) + \theta_n \mathcal{L}_\theta(\lambda, u_n, 0) \\ &\quad + \frac{\theta_n^2}{2} \nabla^2_x \mathcal{L}(\lambda, u_n, 0) d_n d_n + o(\theta_n^2 \|d_n\|^2). \end{aligned}$$

As $\lambda \in \bar{\Lambda}(u_n)$ we have $\mathcal{L}_\theta(\lambda, u_n, 0) \geq \mu$, so that

$$f(x_n, \theta_n) \geq v(0) + \theta_n \mu + \frac{\theta_n^2}{2} \nabla^2_x \mathcal{L}(\lambda, u_n, 0) d_n d_n + o(\theta_n^2 \|d_n\|^2).$$

Maximizing over $\lambda \in \bar{\Lambda}(u_n)$ we get with (3.6), as $\operatorname{dist}(x_n, S)/\theta_n = \|d_n\|$,

$$\bar{\varphi}_{u_n}(d_n) \leq \frac{1}{n}\|d_n\|^2 - n + o(\|d_n\|^2),$$

a contradiction with (SSOC).

Case 2. $\theta_n = O(\mathrm{dist}(x_n, S))$. Then

$$f_i(x_n, \theta_n) = f_i(u_n, 0) + \theta_n \frac{\partial f_i}{\partial \theta}(u_n, 0)(d^n, 1) + O(\theta_n^2),$$

hence

$$\begin{aligned} f(x_n, \theta_n) \;&\geq\; \max_{\lambda \in \Lambda} \mathcal{L}(x, \lambda, \theta_n), \\ &\geq v(0) + \theta_n \max_{\lambda \in \Lambda} \mathcal{L}'(u_n, \lambda, 0)(d^n, 1) + O(\theta_n^2), \\ &\geq v(0) + \theta_n \mu + O(\theta_n^2), \end{aligned}$$

in contradiction with (3.6) (in which we use $\mathrm{dist}(x_n, S) = O(\theta^n)$).

This proves the implication. To prove the inverse implication (SGQ) $\Rightarrow$ (SSOC) we apply the second order necessary condition for quadratic growth (Theorem 3 of [2]) to the problem of minimizing f. As $\boldsymbol{C}(x) \supset C(x) \times \{0\}$ for all $x \in S$, we deduce, using Lemma 3.3, that whenever $x \in S$ is sufficiently close to $\overline{S}$, then for small $\varepsilon > 0$

$$\exists \beta > 0; \quad \max_{\lambda \in \bar{\Lambda}(x)} \nabla_x^2 \mathcal{L}(\lambda, x, 0) dd \geq \beta \|d\|^2, \ \forall d \in C(x) \cap N_S^{\varepsilon}(x),$$

which is equivalent to (SSOC). □

We end this section by some remarks.
The strong quadratic growth condition is sufficient for Lipschitz stability of solutions by Proposition 3.2. On the other hand, a necessary condition for Lipschitz stability is

$$\bar{\varphi}_x(d) \geq 0, \ \forall x \in \overline{S}, \ d \in C(x),$$

as observed by [17] (in the case of an isolated solutions, but the argument is obviously valid for nonisolated solutions). This means that (SQG) is, up to a second order analysis, the weakest possible sufficient condition for Lipschitz stability of solutions.

When the set of solutions is a singleton, it is possible to deduce from the Lipschitz stability of solutions the second-order expansion of the cost and the first-order expansion of pathes of approximate solutions ([1, 3, 4, 15]). In the case of nonisolated solutions, no similar theory has been developed yet.

The above necessary condition for Lipschitz stability is **contingent** in the sense that it depends of the perturbation also. A simple (though rather rough) **structural** (depending on unperturbed data only) sufficient condition for Lipschitz stability is the following. Define

$$\underline{\varphi}_x(d) := \min_{\lambda \in \Lambda} \mathcal{L}_{xx}(\lambda, x, 0)(d, d).$$

Then, as an easy consequence of Theorem 3.4, we obtain

Theorem 3.5. *A sufficient condition for Lipschitz stability is*

$$\exists \varepsilon > 0; \quad \exists \beta > 0; \quad \underline{\varphi}_x(d) \geq \beta \|d\|^2 \ ; \ d \in C^{\varepsilon}(x) \cap N_S(x) \ ; \ x \in S.$$

We end by a remark on the case when the data are jointly convex with respect to (x, u). Then, because of Proposition 2.1, $\mathcal{L}_\theta(\lambda, x, 0)$ is constant over S. It follows that $\overline{S} = S$ and $\bar{\Lambda}(x) = \Lambda$ for all $x \in S$. As a consequence, condition (SSOC) coincide with condition (iv) of Theorem 2.3. It follows by Theorems 2.3 and 3.4 that (QGC) and (SQG) coincide in that case.

4. Studying the critical cone

Our purpose in this section is to prove the following fact closely related to the celebrated Hoffmann's lemma [7] and Robinson's stability theorem [13]. We denote $\alpha^+ := \max(0, \alpha)$.

Theorem 4.1. *Under the assumptions of* Theorem 2.3*, there exists* $\gamma > 0$ *such that*

$$\operatorname{dist}(d, C(x)) \leq \gamma f'(x; d)^+, \ \forall x \in S, \ \forall d \in \mathbb{R}^n.$$

The novelty of this result lies primarily in the fact that neither $f'(x; d)$ nor the set-valued map $C(x)$ are, in general, continuous.

Proof. We shall prove the theorem in several steps.
Step 1. We analyse the set C. Denote $a_i := \nabla f_i(x)$ for $x \in \operatorname{ri} S$. By Proposition 2.1, a_i does not depend on the choice of $x \in \operatorname{ri} S$. Define $A := (a_i)_{i \in I}$. Given $d \in \mathbb{R}^n$, by $A \cdot d$ we mean $(a_i \cdot d)_{i \in I}$. Then

$$C = \{d \in \mathbb{R}^n : A \cdot d \leq 0\}.$$

Let us look closer into the inequalities defining C. Set

$$I_0 := \{i \in I : a_i \cdot d = 0, \ \forall d \in C\}, \text{ and } I_1 := I \backslash I_0.$$

Define similarly $A_0 := (a_i)_{i \in I_0}$, and $A_1 := (a_i)_{i \in I_1}$, so that $A = A_0 \cup A_1$. We claim that there exists $d^0 \in C$ such that $A_1 \cdot d^0 < 0$. Indeed, with each $i \in I_1$ is associated $d^i \in C$ such that $a_i \cdot d^i < 0$, so that $d^0 := \sum_{i \in I_1} d^i$ is the desired vector.
We observe that I_0 cannot be empty. Otherwise d^0 would satisfy $A \cdot d^0 < 0$, which means that $f'(x; d^0) < 0$ whenever $x \in \operatorname{ri} S$: a contradiction to the optimality of x.
By Hoffmann's lemma [7] there exists $\gamma_0 > 0$ such that

$$\operatorname{dist}(d, C) \leq \gamma_0 \|(A \cdot d)^+\|. \tag{4.1}$$

Step 2. We now analyse $C(x)$ when $x \in \operatorname{rb} S$. Define

$$J(x) := I(x) \backslash I, \ B(x) := (\nabla f_i(x))_{i \in J(x)}.$$

As in the proof of Proposition 2.1, we conclude that $\nabla f_i(x) \cdot (x' - x) < 0$ if $x' \in \operatorname{ri} S$ and $i \in J(x)$. The vector $d := x' - x$ belongs to $T_S(x)$, and hence to $C(x)$. Thus, if $J(x) \neq \emptyset$, then

$$C(x) = \{d : A \cdot d \leq 0, \ B(x) \cdot d \leq 0\}$$

and there exists $d^1 \in C(x)$ such that $B(x) \cdot d^1 < 0$.

By Step 1 and the relation $C(x) \subset C$, there exists $\epsilon > 0$ such that $d^2 := d^1 + \epsilon d^0$ satisfies both $A_1 \cdot d^2 < 0$ and $B(x) \cdot d^2 < 0$. For any $J \subset J(x)$, set $B_J(x) := (\nabla f_i(x))_{i \in J}$ and

$$H_J(x) := \{d : A_0 \cdot d = 0,\ A_1 \cdot d \leq 0,\ B_J(x) \cdot d \leq 0\}.$$

We know that $d^2 \in H_J(x)$ satisfies $A_1 \cdot d^2 < 0$ and $B_J(x) \cdot d^2 < 0$. We may assume that A_0 is a set of independent vectors, for dropping linear dependent ones will not change its kernel. We observe that $B_J(u)$ depends continuously on u.
It follows now from Robinson's stability theorem [13] that there exist a neighborhood U of x and a $\delta > 0$ such that

$$\operatorname{dist}(d, H_J(u)) \ \leq \ \delta(\|A_0 \cdot d\| + \|(A_1 \cdot d)^+\| + \|(B_J(u) \cdot d)^+\|), \tag{4.2}$$

for all $d \in R^n$ and all u in a neighbourhood u of x.

We observe further that $\|A_0 \cdot d\| \leq \gamma' \operatorname{dist}(d, \ker A_0)$ for some $\gamma' > 0$, and (4.7), (4.3) imply together with the obvious inclusion $C \subset \ker A_0$ that, for a certain $\gamma = \gamma(x, J)$:

$$\operatorname{dist}(d, H_J(u)) \ \leq \ \gamma(\|(A \cdot d)^+\| + \|(B_J(u) \cdot d)^+\|), \quad \forall d,\ \forall u \in U. \tag{4.3}$$

Step 3. To conclude the proof we first recall that $f'(x; d) = \max\{\nabla f_i(x) \cdot d\ ;\ i \in I(x)\}$. Assuming that the theorem is wrong, we shall find a sequence of $x^k \in S$ and a sequence of vectors $d^k \in R^n$ such that

$$\operatorname{dist}(d^k, C(x^k)) \ \geq \ k f'(x^k; d^k)^+. \tag{4.4}$$

Extracting if necessary a subsequence,we may assume that x^k converges to a certain $x \in S$ and that $J(x^k)$ is equal to some J. Furthermore, as all $C(x)$ are cones and $f(x; \cdot)$ is homogeneous of degree one, we may likewise assume that $\|d^k\| = 1$ for all k and $d^k \to d$. It follows from (4.3) that there is a $\gamma > 0$ such that for sufficiently big k we have

$$\operatorname{dist}(d^k, H_J(x^k)) \ \leq \ \gamma(\|A \cdot d)^+\| + \|(B_J(x^k) \cdot d^k)^+\|).$$

But as $J = J(x^k)$, we have $I \cup J = I(x^k)$, $H_J(x^k) = C(x^k)$, and (4.3) implies that for some $\gamma_1 > 0$

$$\operatorname{dist}(d^k, C(x^k)) \ \leq \ \gamma_1 \max_{i \in I(x^k)} (\nabla f(x^k) \cdot d^k)^+$$

in contradiction with (4.4). This completes the proof of Theorem 3.4. □

5. Proof of Theorem 2.3

The implication (i) ⇒ (ii) is valid even for the non-convex case, see [2]. For the implication (ii) ⇒ (iii) we note that any $d \in \mathbb{R}^n$ can be decomposed as $d = d_T + d_N$, where d_T (resp. d_N) is the projection onto $T_s(x)$ (resp. $N_s(x)$), and $\|d\|^2 = \|d_T\|^2 + \|d_N\|^2$. It follows that

$$\operatorname{dist}(d, T_S(x)) = \|d_N\| \geq (1 - \varepsilon)\|d\| \quad \text{if} \quad d \in N_S^\varepsilon(x),$$

so that (ii) implies (iii) with the same parameter β. To prove that (iii) $\Rightarrow$ (iv), we only have to note that if $d \in C^{\varepsilon}(x) \cap N_S(x)$, there exists $\hat{d} \in C(x)$ with $\|d - \hat{d}\| \le \varepsilon\|d\|$ so that $\|\hat{d}\| \ge (1-\varepsilon)\|d\|$ and

$$\operatorname{dist}(\hat{d}, N_S(x)) \;\le\; \|d - \hat{d}\| \;\le\; \varepsilon\|d\| \;\le\; \frac{\varepsilon}{1-\varepsilon}\|\hat{d}\|,$$

whence $d \in N_S^{\varepsilon'}(x)$ with $\varepsilon' := \varepsilon/(1-\varepsilon)$. We may take ε so small that

$$|\varphi_x(d) - \varphi_x(\hat{d})| \le \frac{\beta}{8}\|\hat{d}\|^2, \quad \text{if } \|d - \hat{d}\| \le \varepsilon\|\hat{d}\|.$$

We may also assume $\varepsilon < \frac{1}{2}$. Then with (iii), using $\|d\| \ge (1-\varepsilon)\|\hat{d}\| \ge \dfrac{1}{2}\|\hat{d}\|$, we obtain $\varphi_x(\hat{d}) \ge \varphi_x(d) - \dfrac{\beta}{8}\|\hat{d}\|^2 \ge \dfrac{\beta}{8}\|\hat{d}\|^2$.

It remains to prove that (iv) $\Rightarrow$ (i). Assuming that the implication does not hold, we find a sequence $u^k \notin S$ converging to a certain $x \in S$and such that

$$f(u^k) \le f(x^k) + \frac{1}{k}\|u^k - x^k\|^2, \tag{5.1}$$

where x^k is the projection of u^k onto S, i.e. $x^k \in S$ and $\|x^k - u^k\| = \operatorname{dist}(u^k, S)$. Set $d^k := (u^k - x^k)/\|u^k - x^k\|$. Then $\|d^k\| = 1$ and $d^k \in N_S(x^k)$. As f is convex, we have $f(u^k) \ge f(x^k) + \|u^k - x^k\| f'(x^k; d^k)$. So with (5.1)

$$f'(x^k; d^k) \le k^{-1}\|u^k - x^k\| \to 0.$$

Theorem 4.1 now implies that $\varepsilon_k := \operatorname{dist}(x^k, C(x^k)) \to 0$. Using (iv), we get for sufficiently large k

$$\varphi_x(d^k) \ge \frac{\beta}{2} > 0. \tag{5.2}$$

Setting $t_k := \|u^k - x^k\|$ we have

$$\begin{aligned} f(u^k) &\ge \max_{\lambda\in\Lambda} \mathcal{L}(u^k, \lambda), \\ &= f(x^k) + \max_{\lambda\in\Lambda}\Big(\sum_{i=1}^{m} \lambda_i \nabla f_i(x^k)\cdot d^k + \frac{t_k^2}{2}\nabla^2 f_i(x^k)(d^k, d^k)\Big) + r_k(t_k)\big), \\ &= f(x^k) + t_k^2\varphi_x(d^k) + r_k(t_k) \ge f(x^k) + \frac{t_k^2}{4}\beta' + r_k(t_k), \end{aligned}$$

where, as $x^k \to x$ and Λ is bounded, the functions $r_k(t)$ are $o(t^2)$ uniformly in k, so for sufficiently large k we have with (5.2)

$$f(u^k) \ge f(x^k) + \frac{\beta}{8}\|u^k - x^k\|^2,$$

in contradiction with (5.1). The proof is completed.

6. Constrained optimization

We now consider an optimization problem of the form

$$\min f_0(x); \quad f_i(x) \le 0, i = 1, \cdots, m. \tag{P}$$

with $f_i(x), i = 0, \cdots, m$ convex and C^2. We say that $x \in \mathbb{R}^n$ is *feasible* if $f_i(x) \le 0, i = 1, \cdots, m$.
Assume that (P) has solutions and let c be the value of $f_0(x)$ over the set of solutions. Changing f_0 into $f_0 - c$ if necessary, we may assume that this optimal value is null. Then x is solution of (P) iff it minimizes the function

$$f(x) := \max_{0 \le i \le m} f_i(x).$$

We now define two concepts of quadratic growth, and study the relations between them. First, we define quadratic growth for (P) as

$$\exists \beta > 0 \ : \ f_0(x) \ge \beta \operatorname{dist}(x, S)^2, \text{ for all feasible } x \text{ in a neighborhood of } S. \tag{QGC}$$

Taking in account a penalization of unfeasibilities, we define generalized quadratic growth for (P) as

$$\exists \beta > 0 \ : \ f(x) \ge \beta \operatorname{dist}(x, S)^2, \text{ for all } x \text{ in a neighborhood of } S. \tag{GQGC}$$

It is obvious that (GQGC) implies (QGC). The converse is not true, as shown by the following example: $\min x^4$; $x^4 \le 0$. The Slater qualification hypothesis is:

$$\exists x^* \in \mathbb{R}^n; f_i(x^*) < 0, i = 1, \cdots, m.$$

If the set of solutions of (P) is nonempty, and the Slater hypothesis holds, we know (see e.g. [8]) that there exists $r > 0$ such that the set of solutions of (P) coincides with the set of points that minimize the so-called exact penalty function

$$\theta_r(x) := f_0(x) + r \sum_{i=1}^{m} \max(0, f_i(x)).$$

Define

$$\hat{f}_i(x) := f_0(x) + r f_i(x), i = 1, \cdots, m.$$
$$\hat{f}_0(x) := f_0(x).$$

Then θ_r can be written as

$$\theta_r(x) := \max_{i=0,\cdots,m} \hat{f}_i(x).$$

Define the collection of active indices

$$\bar{I}(x) := \{i = 1, \cdots, m \ : \ f_i(x) = 0\},$$

the critical cone at x

$$C(x) := \{d \in \mathbb{R}^n \ : \ f_0(x; d) \le 0, f_i'(x; d) \le 0, \ \forall i \in \bar{I}(x)\},$$

the set of Lagrange and generalized multipliers (the components of $z \in \mathbb{R}^{m+1}$ are denoted $z_0, \cdots, z_m$) :

$$\Lambda_0(x) := \{\lambda \in \mathcal{S}^{m+1} \; ; \; \lambda_i = 0, \; i \notin \bar{I}(x) \; ; \; \sum_{i=0}^{m} \lambda_i \nabla f_i(x) = 0\},$$

$$\Lambda_1(x) := \{\lambda \in \mathbb{R}^m_+ \; ; \; \lambda_i = 0, \; i \notin \bar{I}(x) \; ; \; \nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0\},$$

the ordinary and the generalized Lagrangian

$$\mathcal{L}(x, \lambda) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \; ; \quad \hat{\mathcal{L}}(x, \lambda) := \sum_{i=0}^{m} \lambda_i f_i(x),$$

and the associated functions

$$\varphi_x(d) := \max_{\lambda \in \Lambda_1} \nabla^2_x \mathcal{L}(x, \lambda)(d, d) \; ; \quad \hat{\varphi}_x(d) := \max_{\lambda \in \Lambda_0} \nabla^2_x \hat{\mathcal{L}}(x, \lambda)(d, d).$$

Theorem 6.1. *Assume that all functions $f_i, i = 0, \cdots, m$ are convex and twice continuously differentiable. Let S be a non empty convex compact set of solutions of (P). Then the following properties are equivalent:*
(i) (GQGC) *holds.*
(ii) *There exists $\beta > 0$ such that*

$$\hat{\varphi}_x(d) \geq \beta \, \mathrm{dist}(d, T_S(x))^2, \; \forall x \in S, \; \forall d \in C(x) \backslash T_S(x);$$

(iii) *There exists $\beta > 0$, such that for all $\varepsilon \in (0, 1)$*

$$\hat{\varphi}_x(d) \geq (1 - \varepsilon)^2 \beta \|d\|^2, \; \forall x \in S, \; \forall d \in C(x) \cap N^\varepsilon_S(x).$$

(iv) *There exist $\beta' > 0$, $\varepsilon > 0$ such that*

$$\hat{\varphi}_x(d) \geq \beta' \|d\|^2, \; \forall x \in S, \; \forall d \in C^\varepsilon(x) \cap N_S(x).$$

If, in addition, the Slater qualification hypothesis holds, the following properties are equivalent:
(v) (QGC) *holds.*
(vi) *There exists $\beta > 0$ such that*

$$\varphi_x(d) \geq \beta \, \mathrm{dist}(d, T_S(x))^2, \; \forall x \in S, \; \forall d \in C(x) \backslash T_S(x);$$

(vii) *There exists $\beta > 0$, such that for all $\varepsilon \in (0, 1)$*

$$\varphi_x(d) \geq (1 - \varepsilon)^2 \beta \|d\|^2, \; \forall x \in S, \; \forall d \in C(x) \cap N^\varepsilon_S(x).$$

(viii) *There exists $\beta' > 0$, $\varepsilon > 0$ such that*

$$\varphi_x(d) \geq \beta' \|d\|^2, \; \forall x \in S, \; \forall d \in C^\varepsilon(x) \cap N_S(x).$$

Proof. The equivalence between conditions (i) to (iv) is an immediate consequence of Theorem 2.3 applied to $f(x)$. Applying now Theorem 2.3 to $\theta_r(x)$, we find that after some simple computations that properties (vi) to (viii) are equivalent to

$$\exists \beta > 0 \; ; \; \theta_r(x) \geq \beta \text{ dist}(x, S)^2, \text{ in a neighborhood of } S. \tag{6.1}$$

To end the proof it suffices to check that, thanks to the Slater hypothesis, that (6.1) is equivalent to (QGC) when r is large enough. That (6.1) implies (QGC) is obvious. We now prove that the converse holds.

Given $x \in \mathbb{R}^n$, $\alpha \in [0, 1]$, we consider

$$x(\alpha) := \alpha x^* + (1 - \alpha)x$$

where x^* is given by the Slater condition. Set $\gamma^* := -\max\limits_{1 \leq i \leq m} f_i(x^*) > 0$, and $\gamma(x) := \max(0, \max\limits_{1 \leq i \leq m} f_i(x))$. By convexity

$$\begin{aligned} f_i(x(\alpha)) &\leq \alpha f_i(x^*) + (1 - \alpha) f_i(x) \\ &\leq -\alpha\gamma^* + (1 - \alpha)\gamma(x) \\ &\leq -\alpha\gamma^* + \gamma(x). \end{aligned}$$

We deduce that for

$$\alpha(x) := \min(1, \gamma(x)/\gamma^*),$$

the point $y(x) := \alpha(x)x^* + (1 - \alpha(x))x$ is feasible, and

$$\|y(x) - x\| = \alpha(x)\|x^* - x\| \leq \frac{\gamma(x)}{\gamma^*}\|x^* - x\|.$$

Let U be a compact convex neighborhood of S containing x^* and L the Lipschitz constant of f_0 on U. Set

$$r := \max_{x \in U} \left\{ \frac{\|x^* - x\|}{\gamma^*}(L + 2\beta \text{ dist}(x, S))\gamma(x)) \right\}.$$

Then

$$\begin{aligned} \theta_r(x) &= f_0(y(x)) + f_0(x) - f_0(y(x)) + r\gamma(x), \\ &\geq \beta \text{ dist}(y(x), S)^2 - L\|y(x) - x\| + r\gamma(x), \\ &\geq \beta \text{ dist}(y(x), S)^2 + \left(r - L\tfrac{\|x^* - x\|}{\gamma^*} \right) \gamma(x). \end{aligned}$$

On the other hand,

$$\text{dist}(y(x), S) \geq \text{dist}(x, S) - \|y(x) - x\| \geq \text{dist}(x, S) - \frac{\|x^* - x\|}{\gamma^*}\gamma(x),$$

so that

$$\begin{aligned}\theta_r(x) \geq & \beta \operatorname{dist}(x,S)^2 \\ & + \left[r - L\frac{\|x^* - x\|}{\gamma^*} - 2\beta\frac{\|x^* - x\|}{\gamma^*} \operatorname{dist}(x,S) + \beta\frac{\|x^* - x\|^2}{(\gamma^*)^2}\gamma(x)\right]\gamma(x), \\ \geq & \beta \operatorname{dist}(x,S)^2 + \left[r - \frac{\|x^* - x\|}{\gamma^*}(L + 2\beta \operatorname{dist}(x,S)\right]\gamma(x), \\ \geq & \beta \operatorname{dist}(x,S)^2,\end{aligned}$$

as was to be proved. □

References

[1] A. Auslender, R. Cominetti: First and second order sensitivity analysis of nonlinear programs under directional constraint qualification condition. Optimization 21, 1990, 351–363.

[2] J.F. Bonnans, A.D. Ioffe: Second-order sufficiency and quadratic growth for non isolated minima. Inria Report 1853. Mathematics of Operations Research, 1993, to appear.

[3] J.F. Bonnans, A.D. Ioffe, A. Shapiro: Expansion of exact and approximate solutions in nonlinear programming. In Lecture Notes in Economics and Mathematical Systems 382, W. Oettli and P. Pallaschke eds., Springer Verlag, 1993, 103–117.

[4] J.F. Bonnans, A.D. Ioffe, A. Shapiro: Développement de solutions exactes et approchées en programmation non linéaire. Comptes Rendus Acad. Sci. Paris Ser. I, 1992, 119–123.

[5] J. Gauvin, R. Janin: Directional behavior of optimal solutions in nonlinear mathematical programming. Mathematics of Operation Research 13, 1988, 629–649.

[6] E.G. Gol'stein: Theory of convex programming. Nauka, Moscow, 1970; Transl. Math. Monographs 36, Amer. Mathematics Soc., Providence, R.I., 1972.

[7] A. Hoffmann: On approximate solutions of systems of inequalities, J. Res. National Bureau Standards, Sect. B 49, 1952, 263–265.

[8] A.D. Ioffe: Necessary and sufficient conditions for a local minimum. SIAM J. Control Optimization 17, 1979, 245–288.

[9] A.D. Ioffe: Variational analysis of a composite function : a formula for the lower second order epi-derivative, Mathematics of Analysis Applications 160, 1991, 379–405.

[10] A.D. Ioffe: On sensitivity analysis of nonlinear program in Banach spaces : the approach via composite unconstrained optimization, SIAM J. Optimization, 4, 1994, 1–43.

[11] E.S. Levitin, A.A. Miljutin, N.P. Osmolowski: On conditions for a local minimum in problems with constraints, in Mathematical Economics and Functional Analysis, B.S. Mitiagin ed., Nauka, Moscow, 1974, 139–202 (In Russian).

[12] O.L. Mangasarian: A simple characterization of solution sets of convex programs. Oper. Res. Letters 7, 1988, 21–26

[13] S.M. Robinson: Stability theory for systems of inequalities, part 2, SIAM J. Numerical Analysis, 13, 1976, 487–513.

[14] R.T. Rockafellar: Convex Analysis, Princeton University Press 1970.

[15] A. Shapiro: Sensitivity analysis of nonlinear programs and differentiability properties of metric projections. SIAM J. Control Optimization 26, 1988, 628–645.

[16] A. Shapiro: Perturbation theory of nonlinear programs when the set of solution is not a singleton. J. Applied Mathematics of Optimization 18, 1988, 215–229.

[17] A. Shapiro: Perturbation analysis of optimization problems in Banach spaces. Numerical Functional Analysis and Optimization 13, 1992, 97–116.

[18] D. Klatte: On quantitative stability for non-isolated minima. Control and Cybernetics 23, 1994.

HIER :

Leere Seite
58