Extended quadratic tangent optimization problems

J.F. BONNANS INRIA, B.P. 105, 78153 Rocquencourt, France.

Abstract We associate an *extended quadratic tangent* (EQT) optimization problem with any feasible point of a nonlinear programming problem. The EQT problem has linear homogeneous constraints and its cost is the sum of a linear term and of a finite supremum of quadratic terms. The second-order necessary or sufficient conditions have a natural formulation in terms of the EQT problem. The strong regularity condition may also be formulated using EQT problems to simple perturbations of the original optimization problem. Local properties of sequential quadratic programming algorithms are best understood using this concept. We give a partial extension of the theory to nonisolated solutions. Finally, we discuss the extension of the theory to problems with abstract constraints.

1 INTRODUCTION

This paper gives a new presentation of the second order theory for nonlinear programming problems (a finite dimensional optimization problem with finitely many constraints) and discusses some extensions. The theory of second-order necessary or sufficient conditions is due to Levitin, Miljutin and Osmolovski (1974) (see also Ioffe (1979), Ben-Tal (1980)). The results are expressed in terms of the curvature of the Hessian of the Lagrangian along critical directions. The second order necessary condition tells that, if some qualification condition is satisfied, then this curvature is, for a certain Lagrange multiplier depending on the critical direction, nonnegative. That this curvature is positive is a sufficient condition for optimality. For non-qualified problems, similar results can be expressed in terms of a generalized Lagrangian.

Our presentation of the second-order theory is based on the notion of *extended quadratic tangent optimization problem.* Consider an unconstrained minimization problem

$$\operatorname*{Min}_{x} f(x); \quad x \in \mathbb{R}^n.$$

With $x \in \mathbb{R}^n$ we associate the *tangent quadratic problem*

$$(Q_x) \qquad \operatorname*{Min}_{d} f'(x)d + \frac{1}{2} f''(x)dd; \quad d \in \mathbb{R}^n.$$

The second-order necessary optimality condition may be expressed as follows: 0 is a solution of (Q_x), while the second-order sufficient optimality condition is that 0 is the unique solution of (Q_x). Problem (Q_x) is also a central object in many optimization algorithms.

For constrained problems the situation is not so simple. We nevertheless show that under a certain qualification condition, one may state an extended quadratic problem (in the sense of section 2) that plays the same role. This extended quadratic tangent (EQT) problem has linear homogeneous constraints and its cost is the sum of a linear term and of a finite supremum of quadratic terms. We reformulate in §3 the second-order analysis as follows: a local solution of a nonlinear programming problem (P) is, assuming the constraints to be qualified, such that 0 is a local minimum of the EQT problem; if 0 is an strict local minimum of the EQT problem, then the point is a local minimum of (P). The case of local solutions with unqualified constraints is discussed in §4. In §5 we show that the strong regularity condition of Robinson (1980)) may be formulated by reference to the EQT problem of a simple perturbations of the original optimization problem. We discuss also the directional second-order condition of Shapiro (1988). Then in §6 we express a necessary and sufficient condition for sequential quadratic programming algorithms to converge in the vicinity of a local solution. This condition, whose formulation is based on the quadratic tangent problem at the solution, implies the quadratic convergence of the sequence. In §7 we present an extension of EQT optimization problems to sets of nonisolated solutions. Taking advantage of results in Bonnans and Ioffe (1995,1996), we give a characterization of quadratic growth for convex programs in terms of the EQT optimization problem. Finally, in §8 we discuss an extension of EQT problems to optimization problems with abstract constraints.

2 EXTENDED QUADRATIC PROGRAMMING

We define an *extended quadratic optimization problem* as a problem of the form

$$(EQP) \qquad \operatorname*{Min}_{d \in \mathbb{R}^n} c^t d + Q(d); Ad \leq 0,$$

where $c \in \mathbb{R}^n$, A is a $p \times n$ matrix, the inequality $Ad \leq 0$ is taken componentwise, and

$$Q(d) := \sup\{d^t H d \; ; \; H \in \mathcal{H}\},$$

where $\mathcal{H}$ is a nonempty bounded set of $n \times n$ symmetric matrices. When $\mathcal{H}$ is a singleton, problem (EQP) is quadratic (i.e. it has a quadratic cost and linear constraints). Since $\mathcal{H}$ is bounded, $Q(d) = O(\|d\|^2)$, and an obvious necessary condition for $d = 0$ to be a local minimum of (EQP) is that

$$c^t d \geq 0, \quad \forall d \in \mathbb{R}; \quad Ad \leq 0. \tag{1}$$

This is actually equivalent to existence of some Lagrange multipliers associated with the linear problem $\text{Min}\{c^t d; Ad \leq 0\}$, but we do not need to introduce these multipliers in the subsequent analysis. The cone of critical directions is defined as

$$C := \{d \in \mathbb{R}^n; \quad c^t d \leq 0; \quad Ad \leq 0\}.$$

The theorem below is the cornerstone of the paper. In the case of a quadratic programming problem it reduces to Majthay's result (Majthay (1971)). The statement uses the following concept: A local solution $\bar{x}$ of the optimization problem $\text{Min}\{f(x); x \in X\}$, where X is a subset of a normed space, is said to be a *strict local solution* if

$$\exists \epsilon > 0; \quad f(x) > f(\bar{x}) \quad \text{whenever } x \in X, \ \|x - \bar{x}\| \ \leq \epsilon,$$

and to satisfy the *quadratic growth condition* if

$$\exists \alpha > 0; \quad f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\|^2 + o(\|x - \bar{x}\|^2) \quad \text{whenever } x \in X.$$

The feasible set, set of solutions and optimal value function of an optimization problem (P) are denoted $F(P)$, $S(P)$ and $v(P)$.

THEOREM 1 *(i) Problem (EQP) has a local solution at 0 iff (1) holds and $Q(d) \geq 0$ for all critical direction d.*

(ii) The point 0 is a strict local solution of (EQP) satisfying the quadratic growth condition iff (1) holds and $Q(d) > 0$ for all non zero critical direction d.

Proof. (i) **"Only if" part**. If 0 is a local solution of (EQP) and d is a feasible direction, then for all $\sigma > 0$

$$0 \leq c^t(\sigma d) + Q(\sigma d) = \sigma c^t d + \sigma^2 Q(d).$$

Dividing by $\sigma \downarrow 0$, we get $c^t d \geq 0$, whence (1) holds. If $c^t d = 0$ then d is a critical direction, and dividing by σ^2 the above relation, we get $Q(d) \geq 0$, as had to be proved.

"If part" Let us proceed by contradiction. If 0 is not a local minimum of (EQP), then there exists sequences $d^k \in \mathbb{R}^n$, $\|d^k\| = 1$, $\sigma_k \downarrow 0$, such that

$$c^t(\sigma_k d^k) + Q(\sigma_k d^k) < 0.$$

Note that $c^t d^k \geq 0$, by (1) and feasibility of direction d^k. We have then

$$Q(d^k) < -c^t d^k / \sigma_k \leq 0. \tag{2}$$

By Hoffmann's lemma (Hoffman (1952)) there exists a critical direction $\hat{d}^k$ such that

$$\hat{d}^k = d^k + O(c^t d^k).$$

As $\mathcal{H}$ is bounded, the function Q is Lipschitz continuous. Using $Q(\hat{d}^k) \geq 0$ and the above equality, we get

$$Q(d^k) = Q(\hat{d}^k) + O(d^k - \hat{d}^k) \geq O(d^k - \hat{d}^k) = O(c^t d^k).$$

Combining with (2), we deduce that $c^t d^k/\sigma_k \leq O(c^t d^k)$, a contradiction.

(ii) **"Only if" part** Let 0 be a strict local solution of (EQP). By (i) we know that (1) holds. Let d be a critical direction. Then $d \in F(EQP)$ and $c^t d = 0$, therefore $Q(d) > 0$ as was to be proved.

"If part" Set $\alpha := \inf\{Q(d); d \in C, \|d\| = 1\}$. The continuous mapping Q attains its minimum on the compact set $\{d \in C, \|d\| = 1\}$; therefore $\alpha > 0$. We claim that $\bar{x}$ satisfies the quadratic groth condition with this value of the parameter α: it is enough to notice that $Q_\alpha(d) := Q(d) - \alpha\|d\|^2$ has the same set of critical directions C than Q and satisfies $Q_\alpha(d) \geq 0$, for all $d \in C$. Therefore, by point (i), $\bar{x}$ is a local minimum of $c^t d + Q_\alpha(d)$, which is equivalent to the desired conclusion. □

3 APPLICATION TO NONLINEAR PROGRAMMING

We consider the nonlinear programming problem

$$(P) \qquad \text{Min}\, f(x); \quad g(x) \prec 0$$

where f and g are smooth mappings from $\mathbb{R}^n$ to $\mathbb{R}$ and $\mathbb{R}^p$, respectively, and $\prec$ stands for a finite number of equalities and inequalities, i.e. for $z \in \mathbb{R}^p$, $z \prec 0$ iff

$$z_i \leq 0 \text{ , for all } i \in I := \{1, \cdots, q\}, \text{ and } z_j = 0 \text{ for all } j \in J := \{q+1, \cdots, p\}.$$

The set of *active* inequality constraints is denoted

$$I(x) := \{i \in I; g_i(x) = 0\}.$$

For future references, let us denote the *critical cone* as

$$C(x) := \{d \in \mathbb{R}^n;\ f'(x)d \leq 0;\ g_i'(x)d \leq 0, i \in I(x);\ g_J'(x)d = 0\}.$$

We start with the case of linear constraints, i.e. $g(x) = Ax + b$, with A being a $p \times n$ matrix and $b \in \mathbb{R}^p$. Define the *tangent quadratic problem* to the linearly constrained problem (P) at x as

$$(Q_x^a) \qquad \underset{d}{\text{Min}}\, f'(x)d + \frac{1}{2} f''(x)dd \ ; \ g_i'(x)d \leq 0, \ i \in I(x) \ ; \ g_j'(x)d = 0, \ j \in J.$$

THEOREM 2 *Let x be a feasible point of the linearly constrained problem (P). Then:*

(i) If x is a local solution of (P), then $d = 0$ is a local solution of (Q_x^a).

(ii) The point 0 is a strict local solution of (Q_x^a) iff x is a local solution of (P) satisfying the quadratic growth condition.

Proof. (*i*) Let x be a local solution of (P), and let $d \in F(Q_x^a)$. Then $x(\sigma) := x + \sigma d$ is feasible for small enough σ. Since

$$f(x(\sigma)) = f(x) + \sigma f'(x)d + \frac{\sigma}{2} f''(x)dd + o(\sigma^2),$$

we deduce that $f'(x)d \geq 0$, and $f''(x)dd \geq 0$ whenever $f'(x)d = 0$. Therefore, by theorem 1(i), 0 is a local solution of (Q_x^a).

(ii) The constraints of (P) being linear, problems (P) and (Q_x^a) have (up to a translation) the same feasible set and second-order expansion of the cost function. Therefore, the quadratic growth condition holds for one of them iff it holds for the other. Consequently, point (ii) is a consequence of theorem $1(ii)$. □

We now turn to nonlinearly constrained problems. Let x be a feasible point of (P). The set of*Lagrange multipliers* associated with x is defined as:

$$\Lambda(x) := \left\{\lambda \in \mathbb{R}^p; \nabla f(x) + g'(x)^t\lambda = 0; \lambda_i \geq 0; g_i(x) \leq 0; \lambda_i g_i(x) = 0,\ \forall i \in I\right\}.$$

We say that the Mangasarian-Fromovitz (1967) qualification condition holds at x if

$$(MF) \qquad \begin{cases} \text{(i)} & \{\nabla_x g_i(x)\},\ i \in J \text{ are linearly independent,} \\ \text{(ii)} & \exists d \in \mathbb{R}^n \ ;\ g'_J(x)d = 0\ ;\ g'_i(x)d < 0,\ i \in I(x). \end{cases}$$

It is known that if $\bar{x}$ is a local minimum of (P), then $\Lambda(\bar{x})$ is non empty and bounded iff (MF) holds at $\bar{x}$ (Gauvin (1979)). The*Lagrangian function* associated with (P) is

$$\mathcal{L}(x, \lambda) := f(x) + \lambda^t g(x).$$

With x we associate the *extended quadratic tangent* (EQT) problem below:

$$(Q_x) \quad \operatorname*{Min}_d f'(x)d + \frac{1}{2} \max_{\lambda \in \Lambda(x)} \mathcal{L}''_{x^2}(x,\lambda)dd\ ; \quad g'_i(x)d \leq 0,\ i \in I(x); \quad g'_J(x)d = 0.$$

We use the convention that the maximum over an empty set is $-\infty$. Consequently, (Q_x) is defined at every feasible point of (P), and has value $-\infty$ if $\Lambda(x)$ is empty. Note that, if $\Lambda(x) \neq \emptyset$ and the constraints are linear, then (Q_x) is identical to (Q_x^a).

THEOREM 3 *let x be a feasible point of (P) satisfying (MF). Then:*

(i) If x is a local solution of (P), then $d = 0$ is a local solution of (Q_x).

(ii) The point $d = 0$ is a strict local solution of (Q_x) if and only if x is local solution of (P) satisfying the quadratic growth condition.

Proof. Point (i) is a consequence of the second-order necessary condition (e.g. [2]) and theorem $1(i)$. The proof of (ii) is similar to the one of the second-order sufficient condition (e.g. [2]). □

Note that under the (MF) hypothesis, if no inequality constraint is active, then (Q_x) is an equality constrained quadratic problem, and local optimality for (Q_x) is equivalent to global optimality.

Let (P_1) and (P_2) be two nonlinear programming optimization problems over $\mathbb{R}^n$. We say that (P_1) and (P_2) are *tangent* at $x \in F(P_1) \cap F(P_2)$ if (P_1) and (P_2) have the same EQT problem at x. If x satisfies (MF) for both (P_1) and (P_2), then if follows from the above theorems that x satisfies the second order necessary (resp. sufficient) second order optimality condition for (P_1) iff it satisfies this condition for (P_2).

4 SINGULAR POINTS

A classical idea for dealing with points that do not necessarily satisfy qualification conditions is the following: with a feasible point $\bar{x}$ for (P), we associate the optimization problem (where $x \in \mathbb{R}^n$ and $z \in \mathbb{R}$):

$$(P^s_{\bar{x}}) \qquad \operatorname*{Min}_{x,z} z; \quad f(x) - f(\bar{x}) - z \leq 0; \quad g_i(x) - z \leq 0, i \in I, \quad g_J(x) = 0.$$

It is easily checked that if $\bar{x}$ is a local solution of (P), then $(\bar{x}, 0)$ is a local solution of $(P^s_{\bar{x}})$. The displacement $(0_{\mathbb{R}^n}, 1)$ belongs to the kernel of equality constraints and satisfies strictly the linearized active inequality constraints. It follows that $(\bar{x}, 0)$ satisfies (MF) for $(P^s_{\bar{x}})$ iff $g'_J(x)$ is onto.

The *singular Lagrangian function* associated with (P) is defined as

$$\mathcal{L}^s(x, \lambda) := \lambda_0 f(x) + \sum_{i \in I \cup J} \lambda_i g_i(x),$$

with $\lambda = (\lambda_0, \cdots, \lambda_p)$. The EQT problem for problem $(P^s_{\bar{x}})$, associated with $(\bar{x}, 0)$, may be written as

$$(Q^s_{\bar{x}}) \qquad \begin{array}{l} \operatorname*{Min}_{d^x, d^z} d^z + \dfrac{1}{2} \max_{\lambda \in \Lambda^g(x)} (\mathcal{L}^s)''_{x^2}(\bar{x}, \lambda) d^x d^x; f'(\bar{x}) d^x \leq d^z; \\ \qquad\qquad g'_i(\bar{x}) d^x \leq d^z, i \in I(\bar{x}); g'_J(\bar{x}) d = 0, \end{array}$$

where $\Lambda^g(x)$ is the set of normalized *generalized Lagrange multipliers* associated with $(\bar{x}, 0)$ (John (1948)):

$$\begin{array}{l} \Lambda^g(\bar{x}) := \{\lambda = (\lambda_0, \cdots, \lambda_p); \;\; \lambda_0 + \displaystyle\sum_{i \in I(x)} \lambda_i = 1; \;\; \lambda_0 \geq 0; \\ \qquad \lambda_i \geq 0, \; \lambda_i g_i(\bar{x}) = 0, \; i \in I \; ; \; \lambda_0 f'(\bar{x}) + \displaystyle\sum_{i=1}^{p} \lambda_i g'(\bar{x}) = 0\}. \end{array}$$

Note that an equivalent formulation of $(Q^s_{\bar{x}})$ is:

$$\operatorname*{Min}_{d \in \mathbb{R}^n} \max[f'(\bar{x})d, g'_i(\bar{x})d, i \in I(\bar{x})] + \frac{1}{2} \max_{\lambda \in \Lambda^g(\bar{x})} (\mathcal{L}^s)''_{x^2}(\bar{x}, \lambda) dd; \; g_J(\bar{x}) d = 0.$$

THEOREM 4 *Let $\bar{x}$ be a feasible point of (P) such that $g'_J(\bar{x})$ is onto. Then:*

(i) If $\bar{x}$ is a local solution of (P), then $(\bar{x}, 0)$ is a local solution of $(P^s_{\bar{x}})$, and 0 is a local solution of $(Q^s_{\bar{x}})$.

(ii) The point 0 is a strict local solution of $(Q^s_{\bar{x}})$ if and only if $(\bar{x}, 0)$ is a local solution of $(P^s_{\bar{x}})$ satisfying the quadratic growth condition. If in addition the (MF) hypothesis holds, this is equivalent to the fact that $\bar{x}$ is a local solution of (P) satisfing the quadratic growth condition.

Proof. Point (i) is a consequence of theorem 3(i), while point (ii) is a consequence of theorem 3(ii) and of the fact that, under the (MF) condition, $\bar{x}$ satisfies the quadratic growth condition for (P) iff it satisfies the quadratic growth condition for $(P^s_{\bar{x}})$: see e.g. Bonnans and Ioffe (1995), section 5. □

5 STRONG SECOND-ORDER CONDITIONS

Some strong forms of the second-order sufficient conditions are used in the stability analysis of solutions of perturbed nonlinear programs, see Fiacco (1983), Levitin (1994) and Bonnans and Shapiro (1996). We show that two of them, that maybe are the most important, may be expressed using the concepts presented here.

Robinson's *strong stability* condition (Robinson (1980)) is in fact a general stability condition for the sum of a smooth mapping and of a multivalued operator. It says that the *linearized operator*, that is the sum of the the multivalued operator and of the linearization of the smooth mapping, is locally the inverse of a Lipschitz mapping. This concept is useful in the study of numerical algorithms and for conducting a perturbation analysis. When applied to the first-order optimality system of a nonlinear program, strong stability is known to have a simple characterization (see e.g. Bonnans, Sulem (1995)) for a simple proof and references therein). Let (x, λ) be a solution of the first order optimality system of (P) (i.e., $\lambda \in \Lambda(x)$). Define

$$I_+(x) := \{i \in I; g_i(x) = 0 \text{ and } \lambda_i > 0\}.$$

A characterization of strong regularity is

$$(SR) \qquad \left\{ \begin{array}{ll} \text{(i)} & \{\nabla_x g_i(x)\},\ i \in J \cup I(x) \text{ are linearly independent,} \\ \text{(ii)} & \mathcal{L}''_{x^2}(x,\lambda)dd > 0,\ \forall d \in \mathbb{R}^n \backslash \{0\},\ g'_i(x)d = 0,\ i \in J \cup I_+(x). \end{array} \right.$$

Point (i) is a constraint qualification hypothesis, that implies uniqueness of the Lagrange multiplier, while (ii) is a strengthened form of the second-order sufficient condition (since the set of directions d in (ii) contains the critical cone).

As (x, λ) satisfies the first order optimality system of (P), a statement equivalent to (ii) is: 0 is the unique solution of the equality constrained quadratic problem

$$\operatorname*{Min}_d f'(x)d + \frac{1}{2}\mathcal{L}''_{x^2}(x,\lambda)dd; g'_i(x)d = 0,\ i \in J \cup I_+(x).$$

Taking point $(SR)(i)$ into account, we see that (SR) is equivalent to the fact that 0 is the unique solution, associated with a unique multiplier, of the equality constrained quadratic problems

$$\operatorname*{Min}_d f'(x)d + \frac{1}{2}\mathcal{L}''_{x^2}(x,\lambda)dd; g'_i(x)d = 0,\ i \in K,$$

for any K such that

$$J \cup I_+(x) \subset K \subset J \cup I(x).$$

We now turn to the discussion of a condition that proved to be useful for the study of optimization problems of the form

$$\text{Min } f(x,u); \quad g(x,u) \prec 0$$

where for simplicity we assume that $u \in \mathbb{R}_+$. An associated linearized problem, expressed at the point $(\bar{x}, u = 0)$ is

$$(L^*) \qquad \operatorname*{Min}_d f'(\bar{x},0)(d,1); \quad g'_i(\bar{x},u)(d,1) \leq 0, i \in I(\bar{x});\ g'_j(\bar{x},0)(d,1) = 0, j \in J.$$

Its dual problem is

$$(D^*) \qquad \operatorname*{Max}_{\lambda} \mathcal{L}'_u(\bar{x}, \lambda, 0); \quad \lambda \in \Lambda(\bar{x}).$$

The set of solutions of the dual problem is therefore a subset of the set of Lagrange multipliers. It may be characterized through the complementarity conditions, as follows:

$$S(D^*) = \{\lambda \in \Lambda(\bar{x});\ \lambda_i = 0, \forall i \in I^*(\bar{x})\},$$

where

$$I^*(\bar{x}) = \{i \in I(\bar{x});\ \exists d \in S(L); d_i > 0\}.$$

The statement of Shapiro's condition (Shapiro 1988) is

$$\forall d \in C(\bar{x});\ \exists \lambda \in S(D^*);\ \mathcal{L}''_{x^2}(\bar{x}, \lambda) dd > 0.$$

Used in connection with a certain directional qualification hypothesis, this condition allows to check that the variation of the solution of the optimization problem is of the order of the perturbation in the data (Shapiro (1988), Auslender-Cominetti (1990), Bonnans-Ioffe-Shapiro (1992)).

Using theorem 1, we may restate Shapiro's condition as follows: 0 is a strict local solution of the extended quadratic optimization problem

$$\operatorname*{Min}_{d} f'(\bar{x})d + \frac{1}{2} \max_{\lambda \in S(D^*)} \mathcal{L}''_{x^2}(\bar{x}, \lambda) dd; g'_i(\bar{x})d \leq 0, i \in I(\bar{x});\ g'_J(\bar{x})d = 0.$$

It is not clear, however, if the above problem may be interpretated as an EQT problem.

6 RELATION WITH NEWTON'S METHOD FOR CONSTRAINED OPTIMIZATION

A natural extension of Newton's methods (for solving systems of non linear equations) to constrained optimization consists in linearizing the first order optimality system at a candidate point (x, λ) in order to compute a convenient displacement (d_0, μ_0). The linearization may be done in such a way that d_0 and $\lambda + \mu_0$ are solution of the optimality system of the quadratic problem

$$(Q(x,\lambda)) \ \operatorname*{Min}_{d} f(x)'d + \frac{1}{2}\mathcal{L}''_{x^2}(x, \lambda) dd; \quad g_I(x) + g'_I(x)d \leq 0; \quad g_J(x) + g'_J(x)d = 0.$$

Note that when $(x, \lambda) = (\bar{x}, \bar{\lambda})$, with $\bar{x}$ a local solution and $\bar{\lambda}$ the unique associated Lagrange multiplier, then $(Q(x, \lambda))$ coincides, up to the non active inequality constraints, with the quadratic tangent problem. The basic algorithm (without linesearches) is as follows:

Algorithm
Choose $x^0 \in \mathbb{R}^n$ and $\lambda_0 \in \mathbb{R}^p$, such that $\lambda_I \geq 0$; $k \leftarrow 0$.
1) Compute (d^k, λ^{k+1}), solution of the optimality system of $Q(x^k, \lambda^k)$.
2) $x^{k+1} := x^k + d^k, k := k + 1$. Go to 1.

We discuss the convergence of (x^k, λ^k). We say that an algorithm is *locally convergent* if, given a starting point close enough to the solution, convergence to the solution always occurs. Because the above algorithm is based on Taylor expansion of data, and reduces to Newton's method applied to the gradient of the value function for unconstrained problems, we may hope no more than local convergence. Now let us assume that the algorithm starts from $(\bar{x}, \bar{\lambda})$, a local solution of (P) and an associated Lagrange multiplier. If $\bar{\lambda}$ is not the unique Lagrange multiplier associated with $\bar{x}$, then local convergence does not occur (take $x^k = \bar{x}$ and λ^k a non constant sequence of Lagrange multiplier associated with $\bar{x}$). Therefore, uniqueness of the Lagrange multiplier is a *necessary* condition for local convergence.

A desirable property, that is independent from local convergence, is that if the algorithm starts from $(\bar{x}, \bar{\lambda})$, then $(x^k, \lambda^k) = (\bar{x}, \bar{\lambda})$ for all $k > 0$. A necessary condition for this is that 0 is a strict local minimum of $(Q_{\bar{x}})$. (Otherwise, there exists a nonzero critical direction d such that $\mathbb{R}_+ d$ is solution of $(Q_{\bar{x}})$). Therefore, we see that uniqueness of the Lagrange multiplier and condition that 0 is a strict local minimum of $(Q_{\bar{x}})$ are necessary conditions for well posedness of the algorithm. The theorem below shows that they are also sufficient, provided the algorithm computes a displacement of "small" norm. This is a natural restriction, as otherwise the sequence might not converge, even if the starting point satisfies the hypotheses of the theorem. So, for simplicity, we assume the displacement of primal variables to be of minimum norm. It is remarkable that these weak hypotheses imply quadratic convergence.

THEOREM 5 *Let $\bar{x}$ be a local solution of (P) and $\bar{\lambda}$ be its unique associated Lagrange multiplier, such that $d = 0$ is an isolated local solution of $Q(\bar{x}, \bar{\lambda})$. Assume that (x^0, λ^0) is close enough to $(\bar{x}, \bar{\lambda})$ and (d^k, λ^k) is such that $\|d^k\|$ is of minimum norm among all solutions of the optimality system of $Q(x^k, \lambda^k)$. Then $(x^k, \lambda^k) \to (\bar{x}, \bar{\lambda})$ quadratically.*

Proof. This is just a reformulation of theorem 6.1 of Bonnans (1994), that is stated in terms of the second-order sufficient condition. Therefore, the equivalence between the two theorems is a consequence of theorem 2. □

7 NONISOLATED SOLUTIONS

The standard second-order sufficient condition implies that the considered point is a strict local solution of problem (P). However, there are important classes of problems that have nonisolated solutions. In particular, convex problems, in the case the solution is not unique, do not have isolated solution. We briefly review some recent results concerning the links between an extended notion of quadratic growth and some recent second-order conditions. Then we extend the notion of EQT problem to nonisolated solutions.

Let S be a closed set of feasible points of (P) such that f has over S a constant value denoted $f(S)$. The distance of x to S is defined as

$$\mathrm{d}_S(x) := \min\{\|x - y\|, y \in S\}.$$

A projection of x onto S, denoted $P_S x$, is a point of S where the minimum is attained. We say that S satisfies the *quadratic growth condition* if

$$\exists\epsilon > 0;\ \exists\alpha > 0 \text{ such that } f(x) \geq f(S) + \alpha \mathrm{d}_S(x)^2 \quad \text{if } x \in F(P),\ \mathrm{d}_S(x) \leq \epsilon.$$

Note that this is an extension of the definition given in section 2. Let us discuss some material borrowed from Bonnans-Ioffe (1995,1996). The *contingent cone* to S at x is defined as

$$T_S(x) := \{d \in \mathbb{R}^n;\ \exists\{x^k\} \in S,\ t_k \downarrow 0, (t_k)^{-1}(x^k - x) \to d\}.$$

The *normal cone* to S at x is the polar cone of $T_s(x)$, i.e.

$$N_S(x) := \{d \in \mathbb{R}^n;\ d^t y \leq 0,\ \forall y \in T_s(x)\}.$$

Given $\epsilon > 0$, the *approximate critical cone* at x is

$$C^\epsilon(x) := \{d \in \mathbb{R}^n;\ \mathrm{dist}(d, C(x)) \leq \epsilon\|d\|\}.$$

We say that (P) is a *stable convex problem* if the set of Lagrange multipliers associated with a solution (that is the same at every solution) is non empty and bounded.

THEOREM 6 *Let S be a compact set of points over which f has a constant value. Assume (MF) to hold at each $x \in S$. Then*

(i) A necessary condition for S to satisfy the quadratic growth condition is

$$\exists\alpha > 0, \epsilon > 0;\ \forall x \in S \text{ and } d \in C^\epsilon(x) \cap N_S(x) :\ \max_{\lambda \in \Lambda(x)} \mathcal{L}''_{x^2}(x,\lambda)dd \geq \alpha\|d\|^2.$$

(ii) Assume that the following estimate of distance to critical cones holds:

$$\exists\gamma > 0; \quad \mathrm{dist}(d, C(x)) \leq \gamma\nabla f(x)^t d,\ \forall(x,d) \in S \times \mathbb{R}^n; \\ g_i'(x)d \leq 0, i \in I(x);\ g_J'(x)d = 0.$$

Then the condition of point (i) is necessary and sufficient for quadratic growth.

(iii) Assume that (P) is a stable convex problem. Then the estimate of distance to critical cones of point (ii) holds and, consequently, the condition of point (i) is necessary and sufficient for quadratic growth.

Proof. Point (i) is proved in Bonnans-Ioffe (1995), theorem 3, point (ii) in Bonnans-Ioffe (1995), theorem 1 and proposition 2, and point (iii) in Bonnans-Ioffe (1996), theorems 2.3 and 4.1. □

By *necessary condition for quadratic growth* and *uniform estimate of distance to critical cones* we will refer to the conditions in points (i) and (ii), respectively, of the theorem. Note that in the case of an isolated solution, the estimate of distance to critical cones holds, being a consequence of Hoffmann's lemma (Hoffman (1952)). A simple example of a nonconvex problem where this does not hold is

$$\mathrm{Min}\, x_1 x_2;\ x \in \mathbb{R}^2, \quad 0 \leq x_1 \leq 1, \quad 0 \leq x_2 \leq 1.$$

Several other sufficient second-order conditions for quadratic growth, in the framework of nonconvex programming, may be found in Bonnans-Ioffe (1995). An early reference on this subject, where the set of active constraints is assumed to be constant over S, is Shapiro (1988a).

We now define an EQT problem associated with a set of possible solutions S, as follows. With $x \in S$ we associate the optimal value function of problem (Q_x):

$$q_x(d) := f'(x)d + \frac{1}{2} \max_{\lambda \in \Lambda(x)} \mathcal{L}''_{x^2}(x, \lambda)dd \text{ if } d \in F(Q_x), \ +\infty \text{ otherwise.}$$

Then we define

$$\chi(x) := \inf\{q_{\bar{x}}(x - \bar{x}); \ \|x - \bar{x}\| = \text{dist}(x, S)\}.$$

In the case $S = \{\bar{x}\}$, we have of course $\chi(x) = q_{\bar{x}}(x)$.

Note that if (P) is a stable convex problem, then S is convex and $\chi(x) := q_{\bar{x}}(x - \bar{x})$, where $\bar{x}$ is the unique projection of x onto S. In addition, as the set of Lagrange multipliers is bounded and constant over S, and $F(Q_x)$ contains $F(P) - x$, $\chi(x)$ is a continuous function over $F(P)$.

THEOREM 7 *Let S be a compact set of points over which f has a constant value. Assume (MF) to hold at each $x \in S$, and the uniform estimate of distance to critical cones to hold. Then S satisfies the quadratic growth for (P) iff the problem*

$$(Q_S) \qquad \text{Min}\, \chi(x); \ \text{dist}(x, S) \leq \epsilon$$

has solution S, and S satisfies the quadratic growth condition for (Q_S).

Note that, by theorem 6, the hypotheses on problem (P) are satisfied if (P) is a stable convex problem with a bounded set of solutions.

Proof. (a) Assume that S satisfies the quadratic growth condition for (P). If S does not satisfy the quadratic growth condition for (Q_S), then there exist $\{x^k\} \subset F(Q_S)$ such that $\text{dist}(x^k, S) \to 0$ and $\chi(x^k) \leq \text{dist}(x^k, S)^2/k$ (as the value of χ over S is 0). Let $\bar{x}^k = P_S x^k$ be such that $\chi(x^k) = q_{\bar{x}^k}(x^k - \bar{x}^k)$. Denote

$$\sigma_k := \|x^k - \bar{x}^k\|, \quad d^k := (x^k - \bar{x}^k)/\sigma_k.$$

Set $\epsilon > 0$. As

$$0 \geq \limsup \chi(x^k)/\sigma_k = q_{x^k}(d^k)/\sigma_k = \nabla f(\bar{x}^k)^t d^k,$$

we have by the uniform estimate of distance to critical cones that $d^k \in C^\epsilon(\bar{x}^k)$ for large enough k. Note that $f'(x^k)d^k \geq 0$ as $d^k \in F(Q_{\bar{x}^k})$ and (MF) condition holds over S. By theorem 6, the second-order necessary condition for quadratic growth implies

$$\chi(x^k) = q_{x^k}(d^k) \geq \frac{1}{2} \max_{\lambda \in \Lambda(x)} \mathcal{L}''_{x^2}(x, \lambda)dd \geq \alpha \|d\|^2,$$

for some $\alpha > 0$ not depending on k, which gives the desired contradiction.

(b) Let S satisfies the quadratic growth condition for (Q_S). Then in particular it satisfies the second-order necessary condition for quadratic growth. By theorem 6, S satisfies the quadratic growth condition for (P). □

8 EXTENSION TO ABSTRACT CONSTRAINTS

We now turn our attention to a general optimization problem in Banach spaces (e.g. Bonnans-Cominetti (1996))

$$(P) \qquad \min_x f(x) : G(x) \in K,$$

where f and G are $\mathcal{C}^2$ mappings from X to $\mathbb{R}$ and Y respectively, X and Y are Banach spaces, and K is a closed convex subset of Y. In order to state a natural extension of the EQT problem, we recall the definition of the first and second order tangent sets to K

$$\begin{aligned} T_K(y) &:= \{h \in Y : \text{ there exists } o(t) \text{ such that } y + th + o(t) \in K\}, \\ T_K^2(y,h) &:= \{k \in Y : \text{ there exists } o(t^2) \text{ such that } y + th + \frac{1}{2}t^2 k + o(t^2) \in K\}. \end{aligned}$$

A natural first-order approximation of the optimization problem at a point $\bar{x}$ is

$$(L) \qquad \operatorname*{Min}_d f'(\bar{x})d; \quad G'(\bar{x})d \in T_K(G(\bar{x})).$$

We assume that $\bar{x}$ satisfies an extension of (MF) condition, due to Robinson (1976)

$$(EMF) \qquad 0 \in \text{int } [G(x_0, 0) + G'_x(x_0, 0)X - K].$$

Under this condition, if $\bar{x}$ is a local solution of (P), then $v(L) = v(D) = 0$ where (D) is the dual of (L), i.e.

$$(D) \qquad \operatorname*{Max}_\lambda 0; \quad \mathcal{L}'_x(\bar{x}, \lambda) = 0; \quad \lambda \in N_K(G(\bar{x})),$$

and $S(D)$ is the set of Lagrange multipliers associated with $\bar{x}$. The analysis of critical directions leads to the problem (Cominetti (1990))

$$(L_d) \qquad \operatorname*{Min}_w f'(\bar{x})w + f''(\bar{x})dd; \quad G'(\bar{x})w + G''(\bar{x})dd \in T_K^2(G(\bar{x}), G'(\bar{x})d).$$

Assuming (EMF), a *second-order necessary condition* is that $v(L_d) \geq 0$ for all critical direction d (Cominetti (1990)). Another consequence of (EMF) is that the value of (L_d) is equal to that of its dual

$$(D_d) \qquad \operatorname*{Max}_\lambda \mathcal{L}''_{x^2}(\bar{x}, \lambda)dd - \sigma(\lambda, T_K^2(G(\bar{x}), G'(\bar{x})d)); \quad \lambda \in S(D).$$

Here σ stands for the support function

$$\sigma(y, Z) := \sup\{\langle y, z \rangle; \ z \in Z\}.$$

That $v(D_d) \geq 0$ appears as a natural extension of the second-order necessary condition in nonlinear programming (as this σ term is zero in that case). Therefore, a natural extension of the EQT problem is

$$(Q_x) \qquad \operatorname*{Min}_d f'(\bar{x})d + \frac{1}{2}v(D_d) \quad \text{if} \quad G'(\bar{x})d \in T_K(G(\bar{x})), \quad +\infty \text{ if not,}$$

and a natural question is whether the second-order necessary condition is equivalent to the fact that 0 is a local solution of (Q_x).

We say that the cone of critical directions $S(L)$ has the *approximation property* if, given $d \in F(L)$, there exists $\bar{d} \in S(L)$ with $\|d - \bar{d}\| = O(f'(\bar{x})d)$. By Hoffman (1952), the approximation property holds for a qualified problem with a finite number of equality and inequality constraints.

THEOREM 8 *Assume that the (EMF) condition is satisfied, that the cone of critical directions $S(L)$ has the approximation property, and that the cost function of (Q_x) is Lipschitz continuous over $F(Q_x)$. Then the second-order necessary condition holds iff 0 is a local solution of (Q_x).*

Proof. The proof is similar to the one of theorem 1(i). Indeed, set $Q(q) := \frac{1}{2}v(D_d)$. Then Q is positively homogeneous of order 2 by definition of second order tangent sets, and Lipschitz continuous over $F(Q_x)$ by hypothesis. These are the two properties used in the proof of theorem 1(i). □

We note that the hypothesis that the cost function of (Q_x) is Lipschitz continuous over $F(Q_x)$ is always satisfied for positive definite optimization problems, as follows from section 4 in Shapiro (to appear).

Another approach to tangent quadratic problems is presented in Bonnans (1996), in the context of optimal control problems with polyhedric feasible sets.

References

[1] A. Auslender and R. Cominetti, *First and second order sensitivity analysis of nonlinear programs under directional constraint qualification conditions*, Optimization 21(1990), pp. 351-363.

[2] A. Ben-Tal, *Second-order and related extremality conditions in nonlinear programming*, J. Optimization Theory Applications 21(1980), pp. 143–165.

[3] J.F. Bonnans, *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*, Applied Mathematics Optimization 29(1994), pp. 161–186.

[4] J.F. Bonnans, *Second order analysis for control constrained optimal control problems of semilinear elliptic systems.* Rapport de Recherche INRIA 3014, 1996.

[5] J.F. Bonnans and R. Cominetti, *Perturbed optimization in Banach spaces I: a general theory based on a weak directional constraint qualification.* SIAM J. Control Optimization 34 (1996), 1151–1171.

[6] J.F. Bonnans, R. Cominetti and A. Shapiro, *Second order necessary and sufficient optimality conditions under abstract constraints*, Rapport de Recherche INRIA 2952 (1996).

[7] J.F. Bonnans and A.D. Ioffe, *Quadratic growth and stability in convex programming problems with multiple solutions.* J. Convex Analysis 2(1995) (Special issue dedicated to R.T. Rockafellar), pp. 41–57.

[8] J.F. Bonnans and A.D. Ioffe, *Second-order sufficiency and quadratic growth for non isolated minima.* Mathematics of Operations Research 20 (1996), 801–817.

[9] J.F. Bonnans, A.D. Ioffe and A. Shapiro, *Développement de solutions exactes et approchées en programmation non linéaire.* Comptes Rendus de l'Académie des Sciences de Paris, t. 315, Série I, p. 119–123.

[10] J.F. Bonnans and A. Shapiro, *Optimization Problems with perturbations, A guided tour.* Rapport de Recherche INRIA 2872, 1996.

[11] J.F. Bonnans and A. Sulem, *Pseudopower expansion of solutions of generalized equations and constrained optimization problems*, Mathematical Programming 70(1995), pp. 123–148.

[12] R. Cominetti, *Metric regularity, tangent sets and second order optimality conditions*, Applied Mathematics and Optimization 21(1990), pp. 265-287.

[13] A.V. Fiacco, *Introduction to sensitivity and stability analysis in nonlinear programming*, Academic Press, New York, 1983.

[14] J. Gauvin, *A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming*, Mathematical Programming 12(1977), pp. 136–138.

[15] A. Hoffman, *On approximate solutions of systems of inequalities*, J. Research National Bureau of Standards, Sect. B49(1952), pp. 629–649.

[16] A.D. Ioffe, *Necessary and sufficient conditions for a minimum*, SIAM J. Control and Optimization (1979), pp. 245–288.

[17] F. John, *Extremum problems with inequalities as subsidiary conditions*, in Studies and Essays, R. Courant anniversary volume, Interscience, New York, 1948, pp. 187-204.

[18] E.S. Levitin, *Perturbation theory in mathematical programming and its applications.* J. Wiley, Chichester, 1994.

[19] E.S. Levitin, A.A. Miljutin and N.P. Osmolovski, *On conditions for a local minimum in a problem with constraints.* In "Mathematical economics and functional analysis, B.S. Mitjagin ed., Nauka, Moscow, 1974, pp. 139-202 (In Russian).

[20] A. Majthay, *Optimality conditions for quadratic programming*, Mathematical Programming 1(1971), pp. 359–365.

[21] O.L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. of Mathematical Analysis and Applications 17(1967), pp. 37–47.

[22] S.M. Robinson, *Stability theory for systems of inequalities, part II: differentiable nonlinear systems*, SIAM J. Numerical Analysis 13(1976), pp. 497–513.

[23] S.M. Robinson, *Strongly regular generalized equations*, Mathematics of Operations Research 5(1980), pp. 43–62.

[24] A. Shapiro, *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections*, SIAM J. Control and Opt. 26(1988), pp. 628–645.

[25] A. Shapiro, *Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton*, Applied Mathematics and Optimization 18(1988a), pp. 215-229.

[26] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Mathematical Programming, Series B, to appear.