Journal of Global Optimization **13:** 349–367, 1998.

Conditions for Global Optimality 2

JEAN-BAPTISTE HIRIART-URRUTY
Université Paul Sabatier, Toulouse, France

(Accepted in final form 30 June 1998)

Abstract. In this paper bearing the same title as our earlier survey-paper [11] we pursue the goal of characterizing the global solutions of an optimization problem, i.e. getting at necessary and sufficient conditions for a feasible point to be a global minimizer (or maximizer) of the objective function. We emphasize nonconvex optimization problems presenting some specific structures like 'convex-anticonvex' ones or quadratic ones.

Key words: Convex minimization, d.c. Optimization, Global optimality conditions, Quadratic optimization, Reverse convex optimization

1. Introduction

Consider the optimization problem which consists in minimizing the objective function f over the constraint set S. A global minimizer $\bar{x}$ of f on S is a feasible point ($\bar{x} \in S$) such that $f(\bar{x}) \leqslant f(x)$ for all $x \in S$. Our aim in the present paper, like in a previous one under the same title [11], is to derive *characterizations* of such $\bar{x}$. We review results in that direction discovered since the period of preparation of [11], emphasizing those concerning nonconvex optimization problems which present some specific structures in their formulations. The choice of results we describe reflects unavoidably our interests in this particular area of optimization.

A natural step towards characterizations of global minimizers in an optimization problem is to complete (if possible !) classical conditions for local optimality (like the Karush–Kuhn–Tucker' ones) with some 'global conditions': for example in the problem of minimizing a concave quadratic function over a polyhedron [1, Theorem 2] or in the chemical and phase equilibrium problem [15, Section 3], [18, Section 3]. In some problems local minimizers turn out to be global ones [8]. Here we however focus our attention on more general problems, following the scheme developed in [11]. The paper is divided into three parts. In Section 2 we review some recent results on higher order optimality conditions in unconstrained differentiable optimization. Although these conditions still deal with local optimization, we believe they can help to detect solutions of the given optimization problem among those satisfying usual first and second-order conditions for optimality. Section 3 is along the lines of Section 3 in [11]. We consider there optimization problems presenting what we call some 'convex–anticonvex' structure: convexity is present twice, but once in the wrong (or reverse) way. The three typical classes of problems in that respect are:

($\mathcal{P}_1$) Maximize a convex function over a convex set (*convex maximization*);

($\mathcal{P}_2$) Minimize (or maximize) a difference of convex functions (*diff-convex* or *d.c. optimization*);

($\mathcal{P}_3$) Minimize a convex function subject to $g(x) \geqslant 0$, where g is a convex function (*reverse convex minimization*).

Conditions for global optimality in such problems fall into two classes: those using the subdifferential of the objective function at the candidate point $\bar{x}$ and enlargements of this subdifferential called $\varepsilon-$subdifferentials, and those appealing to the subdifferential of the objective function at all points at the same level as $\bar{x}$. Section 3 is entirely devoted to such types of conditions for global optimality in problems ($\mathcal{P}_1$), ($\mathcal{P}_2$) or ($\mathcal{P}_3$), by completing those derived in [11, Section 3] and setting some unsolved questions. In Section 4 we move into a particular world, the one where 'everything is quadratic'. When all the data of the optimization problem are quadratic (but not necessarily convex quadratic), all the informations concerning them are contained in the first and second-order differentials at any point, so that we can expect to obtain peculiar (global) optimality conditions. This is true to a certain extent, but even in this very specific area of optimization, there are more unsolved problems than well-understood situations.

Throughout we assume for simplicity that the underlying space is an Euclidean one, say $\mathbb{R}^n$ equipped with the standard inner product denoted by $\langle .,. \rangle$ and the associated norm denoted by $\|.\|$.

2. Higher order optimality conditions in differentiable optimization

Let $\mathcal{O}$ be an open subset of $\mathbb{R}^n$ and $f : \mathcal{O} \to \mathbb{R}$. We consider a point $\bar{x} \in \mathcal{O}$ at which f admits differentials of any order. We denote by $D^p f(\bar{x})$ the p-th differential of f at $\bar{x}$. Higher order optimality conditions consist in necessary and/or sufficient conditions for the optimality of $\bar{x}$ appealing to the $D^p f(\bar{x})$ with p greater than the usual $p = 2$.

Let us begin by recalling some basic facts concerning the symmetric multilinear forms, especially $D^p f(\bar{x})$.

Given a symmetric multilinear (p-linear) form $P : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{p\text{-times}} \to \mathbb{R}$, the knowledge of P on $\underbrace{H \times \cdots \times H}_{p\text{-times}}$, where H is a subspace of $\mathbb{R}^n$, amounts to the knowledge of P on the "diagonal part" of $H \times \cdots \times H$, i.e. on $\Delta(H) := \{(d, \ldots, d) \mid d \in H\}$. Indeed there are at least two ways of recovering P on $H \times$

$\cdots \times H$ from P on $\Delta(H)$: for all $(d^1, \dots, d^p) \in H \times \cdots \times H$,

$$P\left(d^1, \dots, d^p\right) = \frac{1}{2^p p!} \sum_{(\varepsilon_1, \dots, \varepsilon_p) \in \{-1,+1\}^p} \varepsilon_1 \cdots \varepsilon_p \widetilde{P}(\varepsilon_1 d^1 + \cdots + \varepsilon_p d^p), \tag{2.1}$$

$$P\left(d^1, \dots, d^p\right) = \frac{(-1)^p}{p!} \sum_{k=1}^{p} (-1)^k \sum_{1 \le i_1 < \cdots < i_k \le p} \widetilde{P}(d^{i_1} + \cdots + d^{i_k}), \tag{2.2}$$

where $\widetilde{P}(u)$ stands for $P(u, \dots, u)$.

For the symmetric bilinear forms $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the above polarization formulas give the classical ones below:

$$B(u, v) = \frac{1}{4}[B(u+v, u+v) - B(u-v, u-v)],$$

$$B(u, v) = \frac{1}{2}[B(u+v, u+v) - B(u, u) - B(v, v)].$$

The set of symmetric $p-$linear forms $P\colon \mathbb{R}^n \times \cdots \mathbb{R}^n \to \mathbb{R}$ is a vector space of dimension $\begin{pmatrix} p+n-1 \\ n-1 \end{pmatrix}$.

Concerning the symmetric p-linear form $D^p f(\bar{x})\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$, its expression on the diagonal part $\Delta(\mathbb{R}^n)$ of $\mathbb{R}^n$ is given via all the partial derivatives of order p of f at $\bar{x}$: for all $d = (d_1, \dots, d_n) \in \mathbb{R}^n$,

$$D^p f(\bar{x})\,(d, \dots, d) = \sum_{\substack{(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n \\ \alpha_1 + \cdots + \alpha_n = p}} \frac{p!}{\alpha_1! \cdots \alpha_n!} \frac{\partial^p f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(\bar{x})\, d_1^{\alpha_1} \cdots d_n^{\alpha_n} \quad . \tag{2.3}$$

Only expressions of this type are used for example in Taylor–Young expansions of $f(\bar{x} + td)$ from $\bar{x}$.

Suppose now we are at a critical (or stationary) point $\bar{x}$ of f, i.e. a point $\bar{x} \in \mathcal{O}$ at which $Df(\bar{x}) = 0$. We set:

$$m := \min\left\{p \ge 2 \mid D^p f(\bar{x}) \ne 0\right\}. \tag{2.4}$$

If there is no $p \ge 2$ such that $D^p f(\bar{x}) \ne 0$, we are caught in a trap and we have no optimality condition to propose .... We therefore assume that there is a $p \ge 2$ for which $D^p f(\bar{x}) \ne 0$, so that the m defined in (2.4) is finite.

The classical higher order optimality conditions take the following form.

THEOREM 2.1

(a) *If $\bar{x}$ is a local minimizer of f, then m is even, say $m = 2k$ with $k \geq 1$, and*

$$D^{2k} f(\bar{x})\,(d, \ldots, d) \geq 0 \quad \textit{for all } d \textit{ in } \mathbb{R}^n. \tag{2.5}$$

(b) *If m is even, $m = 2k$ with $k \geq 1$, and*

$$D^{2k} f(\bar{x})\,(d, \ldots, d) > 0 \quad \textit{for all non-null } d \textit{ in } \mathbb{R}^n, \tag{2.6}$$

then $\bar{x}$ is a strict local minimizer of f.

For $m = 2$ we recognize the usual second-order optimality conditions. The higher order necessary condition (2.5) and sufficient condition (2.6) appeal to the sign of $\widetilde{P}\,(d_1, \ldots, d_n) := D^{2k} f(\bar{x})\,(d, \ldots, d)$, a polynomial function of the n variables $d_1, \ldots, d_n$ which is homogeneous of degree $2k$. Unfortunately there is no easy way (like the spectral theory for $2k = 2$) to test the positivity of $\widetilde{P}$ on $\mathbb{R}^n$. Consequently, apart from some particular situations in polynomial optimization, the optimality conditions of Theorem 2.1 are rarely used in Optimization.

EXAMPLE 2.2 *Consider the particular case where $n = 1$. Given an objective function $f : I \to \mathbb{R}$ defined on the open interval I, a critical point $\bar{x} \in I$ of f, we assume there is a $p \geqslant 2$ such that $f^{(p)}(\bar{x}) \neq 0$. Setting m as the smallest integer p for which $f^{(p)}(\bar{x}) \neq 0$, the sufficient condition in Theorem 2.1 says that if m is even, say $m = 2k$ with $k \geqslant 1$,*

$$\left(f^{(2k)}(\bar{x}) > 0\right) \Rightarrow (\bar{x} \textit{ is a strict local minimizer of } f)\,.$$

We can complement this sufficient condition as follows: if $\bar{x}$ is the only critical point of f on I, then

$$\left(f^{(2k)}(\bar{x}) > 0\right) \Rightarrow (\bar{x} \textit{ is a strict global minimizer of } f \textit{ on } I)\,. \tag{2.7}$$

But this is no more true for $n \geq 2$: we may have $f : \mathbb{R}^2 \to \mathbb{R}$ with only one critical point $\bar{x}$ at which $D^2 f(\bar{x})(d, d) > 0$ for all non-null d in $\mathbb{R}^2$ (hence $\bar{x}$ is a strict local minimizer of f) and still f is unbounded from below ($f(x_1, x_2) = 2x_1^3 + 3e^{2x_2} - 6x_1 e^{x_2}$ is such an example [12, p. 53].

As mentioned earlier, from $m = 2$, Theorem 2.1 says nothing more than what is known from the usual second-order optimality conditions: if $\bar{x}$ is a critical point of f for which $D^2 f(\bar{x}) \neq 0$, then

$$(\bar{x} \text{ is a local minimizer of } f) \quad \Rightarrow \quad \begin{array}{l} (D^2 f(\bar{x})(d, d) \geqslant 0 \text{ for all } d \in \mathbb{R}^n, \\ \text{and} \\ D^2 f(\bar{x})(d, d) > 0 \text{ for some} \\ \ \text{non-null } d \in \mathbb{R}^n); \end{array}$$

$$\begin{pmatrix} D^2 f(\bar{x})(d, d) > 0 \text{ for all} \\ \text{non-null } d \in \mathbb{R}^n \end{pmatrix} \quad \Rightarrow \quad (\bar{x} \text{ is a strict local minimizer of } f)\,.$$

The situation where the non-null quadratic form $D^2 f(\bar{x})$ is positive semidefinite but not positive definite is of a particular importance in Optimization: we are in a *state of uncertainty about the minimality* of $\bar{x}$, the necessary condition for minimality is satisfied but the sufficient condition for (strict) minimality is not satisfied. We therefore examine further this case. Expressed in terms of eigenvalues, the uncertainty case we are considering is as follows: $D^2 f(\bar{x})$ is positive semidefinite and singular, but not zero. We put

$$H := \operatorname{Ker} D^2 f(\bar{x}). \tag{2.8}$$

Under the assumptions we have just made, H is neither reduced to $\{0\}$ nor the whole $\mathbb{R}^n$.

By making use of a fourth-order Taylor–Young expansion of $f(\bar{x} + tu)$ from $\bar{x}$ we easily get at the following third and fourth-order necessary condition for minimality:

$$D^3 f(\bar{x})(u, u, u) = 0 \quad \text{for all } u \in H, \tag{2.9}$$

$$D^4 f(\bar{x})(u, u, u, u) \geq 0 \quad \text{for all } u \in H. \tag{2.10}$$

These conditions, however, are not very informative. There is another third- and fourth-order necessary condition (or sufficient condition) for minimality, sharper than (2.9), (2.10), which is due to Dedieu [2] and improved by Dedieu and Janin [3].

THEOREM 2.3 ([3]). *Let $\bar{x}$ be a critical point of f at which we are in a state of uncertainty about its minimality (i.e. $D^2 f(\bar{x})$ is positive semidefinite and singular but not zero). Then:*

(a)

$$\left\{\begin{array}{ll} \textit{If } \bar{x} \textit{ is a local minimizer of } f, \textit{ we necessarily have:} & \\ D^3 f(\bar{x})(u, u, u) = 0 \textit{ for all } u \in H & (2.11) \\ \textit{and} & \\ D^4 f(\bar{x})(u, u, u, u).D^2 f(\bar{x})(v, v) - 3\left[D^3 f(\bar{x})(u, u, v)\right]^2 \geq 0 & (2.12) \\ \textit{for all } u \in H \textit{ and } v \in H^\perp & \end{array}\right.$$

(b) If (2.11) *holds true and if the strict inequality holds in* (2.12) *for all* $0 \neq u \in H$ *and* $0 \neq v \in H^\perp$ *(recall that $H \neq \mathbb{R}^n$), then $\bar{x}$ is a strict local minimizer of f.*

The proof of these results is based upon fourth-order Taylor–Young expansions of f from $\bar{x}$ along ‘parabolic paths’ $\bar{x} + tu + t^2 v$ instead of ‘linear paths’.

By choosing a non-null v in $H^\perp$ (thus verifying $D^2 f(\bar{x})(v, v) > 0$), we see how (2.12) implies (2.10). Note also that (2.12) is homogeneous of degree six so that we can restrict ourselves to u and v with norm one.

EXAMPLE 2.4 *Conditions* (2.11) *and* (2.12) *are easier to check once one has carried out a linear change of variables through an orthogonal transformation diagonalizing the Hessian matrix representing* $D^2 f(\bar{x})$. *Suppose for example* $n = 2$, $\bar{x}$ *is a critical point of* f *at which* $D^2 f(\bar{x}) = \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix}$, *with* $\lambda = \frac{\partial^2 f}{\partial x_2^2}(\bar{x}) > 0$. *Then* (2.11) *and* (2.12) *amount to*

$$\frac{\partial^3 f}{\partial x_1^3}(\bar{x}) = 0 \quad \text{and} \quad \frac{\partial^2 f}{\partial x_2^2}(\bar{x}) \frac{\partial^4 f}{\partial x_1^4}(\bar{x}) - 3\left(\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\bar{x})\right)^2 \geq 0. \tag{2.13}$$

To go further with this example, suppose that

$$f : x = (x_1, x_2) \in \mathbb{R}^2 \mapsto f(x) := (x_2 - \varphi(x_1))\,(x_2 - \psi(x_1))\,, \tag{2.14}$$

where φ *and* ψ *are functions of the real variable satisfying*

$$\begin{cases} \varphi(0) = \psi(0) = 0, & \varphi'(0) = \psi'(0) = 0 \\ \varphi''(0) \neq \psi''(0). \end{cases} \tag{2.15}$$

Then, calculating $D^2 f(\bar{x}),\ D^3 f(\bar{x})$ *and* $D^4 f(\bar{x})$ *at the critical point* $\bar{x} = (0, 0)$ *of* f *leads to*

$$D^2 f(\bar{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \frac{\partial^3 f}{\partial x_1^3}(\bar{x}) = 0,$$

while

$$\frac{\partial^2 f}{\partial x_2^2}(\bar{x}). \frac{\partial^4 f}{\partial x_1^4}(\bar{x}) - 3\left(\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\bar{x})\right)^2 = -3\left[\varphi''(0) - \psi''(0)\right]^2 < 0.$$

Thus, according to the first part of Theorem 2.3, $\bar{x}$ *is not a local minimizer of* f *(which could not have been decided by using the second-order optimality conditions). The famous counterexample by Peano is the* f *in* (2.14) *with* $\varphi(x_1) = x_1^2$ *and* $\psi(x_1) = 3x_1^2$.

We summarize below the process for deciding whether the critical point $\bar{x}$ is a local minimizer of f or not by calling higher order conditions: with the integer m defined in (2.4):

- If m is odd, $\bar{x}$ cannot be a local minimizer.
- If m is even, we distinguish two cases:

$m \geq 4$	$m = 2$
By using Theorem 2.1, decide Yes or No or remain uncertain.	• By using Theorem 2.1, decide Yes or No or remain uncertain. • If uncertain, by using Theorem 2.3, decide Yes or No or remain uncertain.

At this stage several questions arise:

– Are there fifth and sixth-order (or, more generally, $(2p)$-th and $(2p+1)$-th-order) optimality conditions in the same vein as those in Theorem 2.3, which would allow to sharpen what is known for $m \geq 4$ and therefore reduce the range of uncertainty?
– When $m = 2$ could we even sharpen optimality conditions of Theorem 2.3 in order to reduce more the range of uncertainty?

Even by answering these questions we are far from conditions for global optimality. Another question we would pose in that direction is: given a polynomial (or, more generally, an analytic) objective function f, how to decide in view of the coefficients a_i of f (which fully determine f) whether a critical point of f is a (global) minimizer of f or not?

3. Global optimality in 'convex-anticonvex' optimization

3.1. GLOBAL MINIMIZER VS. LOCAL MINIMIZER

We begin with a general result which explains the philosophy we are going to develop essentially: given a point $\bar{x}$ which is candidate for being a global minimizer of f, look at the other points x at the same level as $\bar{x}$ for f (i.e. satisfying $f(x) = f(\bar{x})$).

THEOREM 3.1 *Let $S \subset \mathbb{R}^n$ be arcwise connected and let $f: S \to \mathbb{R}$ be continuous on S. Then the following assertions concerning $\bar{x} \in S$ are equivalent:*

(i) $\bar{x}$ is a global minimizer of f on S.
(ii) Every x in S at the same level as $\bar{x}$ for f is a local minimizer of f on S.

Proof. Only [(ii)⇒(i)] needs a proof. We proceed by contradiction: assuming (ii), we suppose that $\bar{x}$ is not a global minimizer of f on S. Let therefore $u \in S$ be such that $f(u) < f(\bar{x})$. Since S is arcwise connected, u and $\bar{x}$ can be joined by a continuous path of S: there exists a continuous $\sigma: [0,1] \to S$ with $\sigma(0) = u$ and $\sigma(1) = \bar{x}$. Define $\varphi: t \in [0,1] \mapsto \varphi(t) := f(\sigma(t))$ and set $S_\varphi := \{t \in [0,1]: \varphi(t) = f(\bar{x})\}$. Then S_φ is nonempty (because $1 \in S_\varphi$), closed (since φ is continuous), and bounded from below. We set t_0 as the largest lower bound of S_φ; then $t_0 \in S_\varphi$ and $0 < t_0 \leq 1$. The point $\sigma(t_0)$ is (in S) at the same level as $\bar{x}$ for f, hence it is by assumption a local minimizer of f; consequently t_0 is a local minimizer of φ: there exists $\eta > 0$ such that

$$\begin{pmatrix} |t - t_0| \leq \eta \\ t \in [0,1] \end{pmatrix} \Rightarrow (\varphi(t) \geq \varphi(t_0)) .$$

Since t_0 is the lower bound of S_φ, we deduce from above:

$$\begin{pmatrix} t_0 - \eta < t < t_0 \\ 0 < t \end{pmatrix} \Rightarrow (\varphi(t) > \varphi(t_0)) .$$

Let us choose such a t, call it t_1. Then $\varphi(1) = \varphi(t_0) \in [\varphi(0), \varphi(t_1)]$ so that there exists $t_2 \in [0, t_1]$ such that $\varphi(t_2) = \varphi(1)$ $(= f(\bar{x}))$. Thus $t_2 \in]0, t_1] \cap S_\varphi$, and since $t_1 < t_0$ this contradicts the definition of t_0 as the lower bound of S_φ. □

As in all characterization theorems, Theorem 3.1 can be read in its negative form: let $\bar{x}$ be a local minimizer of f on S; then $\bar{x}$ is not a global minimizer of f on S if and only if there exists $x \in S$, $x \neq \bar{x}$, at the same level as $\bar{x}$ for f which is not a local minimizer of f on S. This kind of result could explain why some global minimization procedures like the so-called ‘tunneling method’ [16] are expected to lead to a global minimizer: move in a first phase to a local minimizer $\bar{x}_1$ of f, find in a second phase a point $\bar{x}_2 \neq \bar{x}_1$ at the same level as $\bar{x}_1$ for f, then apply again a local minimization process from $\bar{x}_2$.

3.2. CONVEX MAXIMIZATION AND D.C. OPTIMIZATION

It is tempting to try to weaken condition (ii) in Theorem 3.1 by substituting ‘Every x in S at the same level as $\bar{x}$ for f satisfies first-order minimality conditions’ for (ii). This is grossly false as simple (unconstrained) minimization problems show. The situation however is not hopeless for minimization problems with more structure. Here enter the so-called ‘convex-anticonvex’ models; by this name we mean minimization problems where convexity is present twice, but once in the wrong (or reverse) way. The first two typical classes of problems in that respect are:

($\mathcal{P}_1$) Maximize a convex function over a convex set (*convex maximization*);
($\mathcal{P}_2$) Minimize (or maximize) a difference of convex functions over a closed convex set (this problem is called *d.c. optimization*).

Consider the first class of problems ($\mathcal{P}_1$): maximize a convex function f over a convex set C. The first-order optimality condition at $x \in C$ for such a problem can be written as follows: ‘$Df(x)$ is normal to C at x’ (assuming f is differentiable at x). This is satisfied at a local maximizer $\bar{x}$ of f on C but, more interesting, its extension to all the x in C which lie at the same level as $\bar{x}$ for f yields a characterization of a global maximizer $\bar{x}$ of f on C. The statement below which improves an earlier result by Strekalovsky (see [11] and references therein) gives such a characterization.

For a convex $f : \mathbb{R}^n \to \mathbb{R}$ we denote by $\partial f(x)$ the subdifferential of f at x; for a convex set $C \subset \mathbb{R}^n$ and $x \in C$ we denote by $N(C, x)$ the normal cone to C at x.

THEOREM 3.2 ([13]). *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $C \subset \mathbb{R}^n$ is closed convex. Consider a point $\bar{x} \in C$ such that $-\infty \leq \inf_C f < f(\bar{x})$. Then $\bar{x}$ is a global* *maximizer of f on C if and only if:*

$$\partial f(x) \subset N(C, x) \quad \textit{for all } x \textit{ in } C \textit{ satisfying } f(x) = f(\bar{x}). \tag{3.1}$$

Another type of conditions for global optimality in such a class of problems $(\mathcal{P}_1)$ was provided earlier by the author [10] by using the so-called ε-subdifferentials $\partial_\varepsilon f(\bar{x})$ of f at $\bar{x}$ and the sets $N_\varepsilon(C, \bar{x})$ of ε-normal directions to C at $\bar{x}$ (see [14, Chapter XI] for properties of such objects): more precisely,

$$\begin{aligned} &\bar{x} \text{ is a global maximizer of } f \text{ on } C \text{ if and only if} \\ &\partial_\varepsilon f(\bar{x}) \subset N_\varepsilon(C, \bar{x}) \quad \text{for all } \varepsilon > 0. \end{aligned} \tag{3.2}$$

A way of linking directly (3.1) and (3.2) has recently been shown in [4, Section 2]. This strategy of using ε-subdifferentials of convex functions and ε-normal directions to convex sets is also followed in [9] for some further nonconvex minimization problems involving the pointwise minimum of sublinear functions.

The two classes of problems $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$ are known to be ‘equivalent’ in the sense that one can easily transform either of them in such a way that it is structured as the other one. It is therefore natural to ask for global optimality conditions for $(\mathcal{P}_2)$ parallel to those obtained for $(\mathcal{P}_1)$ in Theorem 3.2. But here I must confess that what has been derived in that respect is unsatisfactory to me.

Let the objective function f be a d.c. one, i.e. of the following form: $f = g - h$, where $g\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lower-semicontinuous convex function (not identically equal to $+\infty$) and $h\colon \mathbb{R}^n \to \mathbb{R}$ is a convex function. We are tempted to propose the following assertion, which parallels the one displayed in Theorem 3.2 for $(\mathcal{P}_1)$: $\bar{x}$ is a global minimizer of $f = g - h$ on $\mathbb{R}^n$ if and only if

$$\partial h(x) \subset \partial g(x) \quad \text{for all } x \text{ satisfying } f(x) = f(\bar{x}). \tag{3.3}$$

This is false, not only because simple counterexamples exist but also because (3.3) is *symmetrical*, at least when both g and h are differentiable: for differentiable g and h, (3.3) boils down to

$$Dg(x) - Dh(x) = Df(x) = 0 \quad \text{for all } x \text{ satisfying } f(x) = f(\bar{x}). \tag{3.3$'$}$$

Hence, for such functions, conditions for global minimality of $\bar{x}$ and global maximality of $\bar{x}$ would be the same!

A common way to transform $(\mathcal{P}_2)$ into a problem structured like $(\mathcal{P}_1)$ is as follows. Consider

$(\widehat{\mathcal{P}}_1)$ Maximize $\widehat{h}(x, \alpha) := h(x) - \alpha$ subject to $(x, \alpha) \in \widehat{C}$, where $\widehat{C} := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid g(x) - \alpha \leq 0\}$.

Then:

$$(\bar{x} \text{ solves } (\mathcal{P}_2)) \Rightarrow \big((\bar{x}, g(\bar{x})) \text{ solves } (\widehat{\mathcal{P}}_1)\big);$$

$$\big((\bar{x}, \bar{\alpha}) \text{ solves } (\widehat{\mathcal{P}}_1)\big) \Rightarrow (\bar{\alpha} = g(\bar{x}) \text{ and } \bar{x} \text{ solves } (\mathcal{P}_2)).$$

What if we plug the result of Theorem 3.2 into the format $\left(\widehat{\mathcal{P}_1}\right)$ of $(\mathcal{P}_2)$? This has been done by Strekalovsky [25, Section 2]: we provide below a slight improvement of the subsequent statement and a direct proof of it.

THEOREM 3.3 *Under the assumptions above on $f = g - h$, $\bar{x}$ is a global minimizer of f on $\mathbb{R}^n$ if and only if the following holds:*

$$(\mathcal{C}) \begin{cases} \textit{For all } (r, x) \in \mathbb{R} \times \mathbb{R}^n \textit{ satisfying} \\ r - h(x) = g(\bar{x}) - h(\bar{x}) \textit{ and } r \geq g(x), & (3.4) \\ \textit{we must have} \\ g(x') \geq r + \left\langle x^*, x' - x \right\rangle \textit{ for all } x' \in \mathbb{R}^n \textit{ and } x^* \in \partial h(x). & (3.5) \end{cases}$$

Proof.

– *Condition $(\mathcal{C})$ is necessary.*

Let (r, x) satisfy (3.4). Since

$$g(x') - h(x') \geq g(\bar{x}) - h(\bar{x}) \quad \text{for all } x',$$

$$g(x') - h(x') \geq r - h(x) \quad \text{for all } x',$$

that is

$$g(x') \geq r + h(x') - h(x) \quad \text{for all } x'.$$

Now, $h(x') - h(x) \geq \left\langle x^*, x' - x \right\rangle$ whenever $x^* \in \partial h(x)$. Whence (3.5) is deduced.

– *Condition $(\mathcal{C})$ is sufficient.*

Let x be arbitrarily chosen in $\mathbb{R}^n$; we want to prove that $f(x) \geq f(\bar{x})$. We set $r := h(x) + g(\bar{x}) - h(\bar{x})$ and distinguish two cases: $r \leq g(x)$, $r > g(x)$.
If $r \leq g(x)$, according to the definition above of r, we get

$$g(\bar{x}) - h(\bar{x}) \leq g(x) - h(x).$$

If $r > g(x)$, it follows from (3.5):

$$g(x') \geq h(x) + g(\bar{x}) - h(\bar{x}) + \left\langle x^*, x' - x \right\rangle \tag{3.6}$$

for all $x' \in \mathbb{R}^n$ and $x^* \in \partial h(x)$. It then suffices to let $x' = x$ in (3.6) to obtain the desired inequality. □

Condition $(\mathcal{C})$ bears a resemblence to (3.3): (3.3) is condition $(\mathcal{C})$ restricted to the r equalling $g(x)$. The difference between (3.3) and $(\mathcal{C})$ is the presence of an additional parameter r, and we know it cannot be completely removed. We can however get rid of the r by implicitly including them in the values taken by f, with the aim to present a necessary and sufficient condition for global minimality in d.c. optimization bearing the spirit of (3.3) and which is nicer than $(\mathcal{C})$.

THEOREM 3.4 *Under the assumptions of Theorem 3.3 on* $f = g - h$*:*

(a) If $\bar{x}$ *is a global minimizer of* f *on* $\mathbb{R}^n$*, then* $\partial h(x) \subset \partial g(x)$ *for all* x *satisfying* $f(x) = f(\bar{x})$*;*
(b) If $\bar{x}$ *is not a global minimizer of* f *on* $\mathbb{R}^n$*, there then exists* x *satisfying* $f(x) \le f(\bar{x})$ *for which* $\partial h(x)$ *is not included in* $\partial g(x)$*.*

Proof. Only (b) needs a proof. Let $\bar{x}$ not be a global minimizer of f on $\mathbb{R}^n$; then, according to Theorem 3.3, for some (r, x) satisfying $r - h(x) = g(\bar{x}) - h(\bar{x})$ and $r \ge g(x)$, there exists $x^* \in \partial h(x)$ and $x' \in \mathbb{R}^n$ such that

$$g(x') < r + \langle x^*, x' - x \rangle . \tag{3.7}$$

What about such an x? Firstly, $g(x) - h(x) \le r - h(x) = g(\bar{x}) - h(\bar{x})$, that is $f(x) \le f(\bar{x})$. Secondly, $\partial h(x)$ is not included in $\partial g(x)$. Indeed, in the opposite case, we would have

$$\begin{aligned} g(x') &\ge g(x) + \langle x^*, x' - x \rangle \\ &\ge r + \langle x^*, x' - x \rangle , \end{aligned}$$

which contradicts (3.7). □

3.3. REVERSE CONVEX MINIMIZATION

The third class of ‘convex-anticonvex’ optimization problems we treat of is the following one:

$(\mathcal{P}_3)$ Minimize $f(x)$ subject to $g(x) \ge 0$,

where both f and g are convex functions (these are the so-called *reverse convex minimization* problems). For the feasible candidate point $\bar{x}$ we are considering, we assume throughout this subsection the following:

$$(\mathcal{H}) \begin{cases} \bullet \ \text{The constraint set in } (\mathcal{P}_3) \text{ is nonempty} \\ \quad \text{and is not the whole space.} \\ \bullet \ -\infty \leqslant \inf_{\mathbb{R}^n} f(x) < f(\bar{x}). \end{cases}$$

As consequences of $(\mathcal{H})$, the boundary of the constraint set in $(\mathcal{P}_3)$ is $\{x \in \mathbb{R}^n \mid g(x) = 0\}$ and the candidates $\bar{x}$ for being global minimizers in $(\mathcal{P}_3)$ have to be looked for there.
As for the two previous ‘convex-anticonvex’ optimization problems (section 3.2), there are two types of conditions for global optimality in $(\mathcal{P}_3)$: those using the ε-subdifferential of f and g at $\bar{x}$, and those appealing to the subdifferential of f at all feasible points at the same level as $\bar{x}$.

THEOREM 3.5 *Let $\bar{x}$ satisfy $g(\bar{x}) = 0$. Then $\bar{x}$ is a global minimizer in $(\mathcal{P}_3)$ if and only if:*

$$\partial_\varepsilon g(\bar{x}) \subset \bigcup_{\alpha \geq 0} \partial_\varepsilon(\alpha f)(\bar{x}) \text{ for all } \varepsilon > 0. \tag{3.8}$$

Proof. Let $C := \{x \in \mathbb{R}^n \mid f(x) \leq f(\bar{x})\}$. Since there exists x_0 such that $f(x_0) < f(\bar{x})$ (Slater's assumption for the representation above of C as an inequality constraint), the set $N_\varepsilon(C, \bar{x})$ of ε-normal directions to C at $\bar{x}$ can be expressed via the ε-subdifferentials of f at $\bar{x}$:

$$N_\varepsilon(C, \bar{x}) = \bigcup_{\alpha \geq 0} \partial_\varepsilon(\alpha f)(\bar{x})$$

[14, Corollary 3.6.2 of Chapter XI]. So, what (3.8) says is

$$\partial_\varepsilon g(\bar{x}) \subset N_\varepsilon(C, \bar{x}) \text{ for all } \varepsilon > 0. \tag{3.8$'$}$$

Let us, therefore, prove that $\bar{x}$ satisfying $g(\bar{x}) = 0$ is a global minimizer in $(\mathcal{P}_3)$ if and only if (3.8$'$) holds true.

The global minimality of $\bar{x}$ in $(\mathcal{P}_3)$ is expressed by the inclusion

$$\{x \mid f(x) < f(\bar{x})\} \subset \{x \mid g(x) < g(\bar{x}) = 0\}. \tag{3.9}$$

But, with the assumption $(\mathcal{H})$, (3.9) is equivalent to

$$(C =)\ \{x \mid f(x) \leq f(\bar{x})\} \subset \{x \mid g(x) \leq g(\bar{x}) = 0\} \tag{3.10}$$

[14, Proposition 1.3.3 of Chapter VI]. Now, what (3.10) says is that $\bar{x}$ is a global maximizer of g on C. It then remains to call (3.2). □

REMARK 3.6 *When $\alpha > 0$, $\partial_\varepsilon(\alpha f)(\bar{x}) = \alpha \partial_{\frac{\varepsilon}{\alpha}} f(\bar{x})$. Also, when f is bounded from below, $\inf_{\mathbb{R}^n} f(x) > -\infty$, then 0 (which is the only element in $\partial_\varepsilon(\alpha f)(\bar{x})$ for $\alpha = 0$) is contained in $\alpha \partial_{\frac{\varepsilon}{\alpha}} f(\bar{x})$ for some $\alpha > 0$. Whence, in that case, the right-hand side of* (3.8) *is* $\bigcup_{\alpha > 0} \alpha \partial_{\frac{\varepsilon}{\alpha}} f(\bar{x})$.

REMARK 3.7 *The limiting case $\varepsilon \downarrow 0$ in* (3.8) *is*

$$\partial g(\bar{x}) \subset \mathbb{R}^+ \partial f(\bar{x}), \tag{3.11}$$

which is the classical necessary condition satisfied by a local minimizer in $(\mathcal{P}_3)$.

The second type of conditions for global optimality in $(\mathcal{P}_3)$ are 'à la Strekalovsky'; what we present below is a simplified form (statements and proofs) of [24, Theorems 1 and 2].

THEOREM 3.8 *Let $\bar{x}$ satisfy $g(\bar{x}) = 0$.*

(a) If $\bar{x}$ is a global minimizer in $(\mathcal{P}_3)$, then

$$\partial g(x) \subset \mathbb{R}^+ \partial f(x) \text{ for all } x \text{ satisfying } g(x) = 0 \text{ and } f(x) = f(\bar{x}). \tag{3.12}$$

(b) Let C denote $\{x \mid f(x) \leq f(\bar{x})\}$. If $\bar{x}$ is not a global minimizer in $(\mathcal{P}_3)$, there then exists $x \in C$ satisfying $g(x) = 0$ for which $\partial g(x)$ is not included in $N(C, x)$.

Proof. Only (b) needs a proof. As in the proof of Theorem 3.5, the global minimality of $\bar{x}$ in $(\mathcal{P}_3)$ is expressed (under $(\mathcal{H})$) as: $\bar{x}$ is a global maximizer of g on $C = \{x \mid f(x) \leq f(\bar{x})\}$. It then remains to call Theorem 3.2. □

4. Quadratic-quadratic optimization

By quadratic-quadratic optimization we mean optimization problems where all the data functions (objective, equality and/or inequality constraints) are quadratic. In short, we consider the following format:

$$(\mathcal{P}) \begin{cases} \text{Minimize } f(x) := \frac{1}{2}\langle A_0 x, x\rangle + \langle b_0, x\rangle + c_0 \\ \text{subject to} \\ \quad h_i(x) := \frac{1}{2}\langle A_i x, x\rangle + \langle b_i, x\rangle + c_i = 0 \text{ for } i = 1, \dots, m \\ \quad g_j(x) := \frac{1}{2}\left\langle A_j x, x\right\rangle + \left\langle b_j, x\right\rangle + c_j \leq 0 \text{ for } j = m+1, \dots, p, \end{cases}$$

where all the A_i are n-by-n symmetric matrices, the b_i are vectors in $\mathbb{R}^n$, and the c_i are real numbers. The objective function f in $(\mathcal{P})$ is not assumed to be convex (i.e. A_0 is not necessarily positive semidefinite) and, contrary to what is supposed in what is usually called quadratic optimization, the constraint set in $(\mathcal{P})$ is not a convex polyhedron (it may even be disconnected). Many optimization problems are modelled *ab initio* as quadratic-quadratic ones and the model $(\mathcal{P})$ turns out to be fairly general [7, Section 8]; as a consequence there is no hope to derive characterizations of global solutions of $(\mathcal{P})$ in all possible cases. There actually are two well-understood situations:

- When $(\mathcal{P})$ is *convex*, i.e. when: A_0 is positive semidefinite, only inequality constrains occur with A_j positive semidefinite for all j.
- When there is only *one* constraint (equality or inequality) in $(\mathcal{P})$.

The first situation falls in the area of *convex minimization* we let aside here. The second situation, we are going to examine in more details below, developed from the so-called *trust region optimization problems*. We begin by summarizing the latest developments concerning this model (from the point of view of optimality conditions). The trust region optimization problem is formulated as follows:

$$(\mathcal{P}_1)\begin{cases}\text{Minimize } f(x) := \frac{1}{2}\langle A_0 x\rangle + \langle b_0, x\rangle + c_0\\ \text{over } C := \{x \in \mathbb{R}^n \mid \|x\| \leq \delta\}\\ (\text{or over } S := \{x \in \mathbb{R}^n \mid \|x\| = \delta\}).\end{cases}$$

A slightly more general situation is when the constraint set is ellipsoidal instead of a ball or spherical as above, but the gist of the problem is in $(\mathcal{P})$. In pioneering works, D.M. Gay and D.C. Sorensen (1981, 1982) derived characterizations of solutions of $(\mathcal{P}_1)$, i.e. necessary and sufficient conditions for $\bar{x}$ being a global minimizer of f over C. Since then:

- Flippo and Jansen [6] observed that this nonconvex problem is in a sense equivalent to a convex problem of the same type from which known necessary and sufficient conditions for optimality follow.
- The hidden convexity of $(\mathcal{P}_1)$ was confirmed by Pham Dinh Tao and Le Thi Hoai An [21] by showing there is no duality gap between $(\mathcal{P}_1)$ and the associated dual optimization problem.
- Martinez [17] clarified the difference between local and global solutions of $(\mathcal{P}_1)$ by showing there is at most one local-nonglobal minimizer for $(\mathcal{P}_1)$.

For more details on trust region optimization problems $(\mathcal{P}_1)$, peruse Section 1.1 in [22].

Until now the constraint set in $(\mathcal{P}_1)$ is *convex*, which explains why an alteration of the objective function f (by adding $\alpha \|.\|^2$ with α large enough) renders the problem convex. What if the quadratic function defining the constraint set in $(\mathcal{P}_1)$ (in an equality or inequality form) is not convex? We then are faced with a genuinely *nonconvex* optimization problem. Surprisingly enough, the characterizations of the Gay–Sorensen type extend to that case. Such results were proposed by Moré [19] and, for different but related formulations, by Stern and Wolkowicz [23]. Consider:

$$(\mathcal{P}_2)\begin{cases}\text{Minimize } f(x) := \dfrac{1}{2}\langle A_0 x, x\rangle + \langle b_0, x\rangle + c_0\\ \text{subject to}\\ \quad h(x) := \dfrac{1}{2}\langle A_1 x, x\rangle + \langle b_1, x\rangle + c_1 = 0.\end{cases}$$

THEOREM 4.1 [19, p. 195]. *Assume that $A_1 \neq 0$ and that*

$$-\infty \leq \inf_{\mathbb{R}^n} h(x) < 0 < \sup_{\mathbb{R}^n} h(x) \leq +\infty. \tag{4.1}$$

Then $\bar{x}$ satisfying $h(\bar{x}) = 0$ is a global minimizer of problem $(\mathcal{P}_2)$ if and only if there is a multiplier $\bar{\lambda} \in \mathbb{R}$ such that:

(α) $(A_0 + \bar{\lambda} A_1)\bar{x} + b_0 + \bar{\lambda} b_1 = 0$*;*

(β) $A_0 + \bar{\lambda} A_1$ *is positive semidefinite.*

So, just adding condition (β) to the classical Lagrange conditions for optimality (α) gives us a way of characterizing global solutions in ($\mathcal{P}_2$). An instance of problem ($\mathcal{P}_2$) is determining the points in the hypersurface of equation $h(x) = 0$ which are closest to a given point $a \in \mathbb{R}^n$. We illustrate Theorem 4.1 with a very simple example of this kind.

EXAMPLE 4.2 *Find the closest points of* $a := (1, 0)$ *in the set of* $(x_1, x_2) \in \mathbb{R}^2$ *satisfying* $x_1^2 - rx_2^2 - 4 = 0$ *(where* $r > 0$*). We therefore have to minimize globally* $f(x_1, x_2) := (x_1 - 1)^2 + x_2^2$ *subject to* $x_1^2 - x_2^2 - 4 = 0$. *Two points are candidates since they satisfy the Lagrange conditions* (α)*:*

$(2, 0)$ *with the multiplier* $-1/2$*;*
$(-2, 0)$ *with the multiplier* $-3/2$.

Only the first point $(2, 0)$ *satisfies condition* (β)*; we knew it was the only solution of our problem.*

It is interesting to note that the usual trick which transforms an optimization problem (in $\mathbb{R}^n$) with p inequality constraints into an optimization problem (in $\mathbb{R}^n \times \mathbb{R}^p$) with p equality constraints by adding slack variables y_i *preserves the quadratic* character of the optimization problem. Hence, if we consider:

$$(\mathcal{P}_3) \begin{cases} \text{Minimize } f(x) := \dfrac{1}{2} \langle A_0 x, x\rangle + \langle b_0, x\rangle + c_0 \\ \text{subject to} \\ \quad g(x) := \dfrac{1}{2} \langle A_1 x, x\rangle + \langle b_1, x\rangle + c_1 \leq 0, \end{cases}$$

($\mathcal{P}_3$) is equivalent to that of minimizing

$$(x, y) \in \mathbb{R}^n \times \mathbb{R} \mapsto \widehat{f}(x, y) := f(x)$$

subject to

$$\widehat{h}(x, y) := g(x) + y^2 = 0.$$

Thus, Theorem 4.1 has a counterpart for ($\mathcal{P}_3$).

THEOREM 4.3 [19, p. 199]. *Assume that* $A_1 \neq 0$ *and that*

$$-\infty \leq \inf_{\mathbb{R}^n} g(x) < 0. \tag{4.2}$$

Then $\bar{x}$ *satisfying* $g(\bar{x}) \leq 0$ *is a global minimizer of problem* ($\mathcal{P}_3$) *if and only if there is a multiplier* $\bar{\mu} \geq 0$ *such that:*

$(\alpha)'$ $(A_0 + \bar{\mu} A_1)\bar{x} + b_0 + \bar{\mu} b_1 = 0$*;* $(\alpha)''$ $\bar{\mu} g(\bar{x}) = 0$*;*

$(\beta)'$ $A_0 + \bar{\mu} A_1$ *is positive semidefinite.*

EXAMPLE 4.4 *Find the closest point of* $a := (1, 0)$ *in the set of* $(x_1, x_2) \in \mathbb{R}^2$ *satisfying* $x_1^2 + 4x_2^2 - 4 \geq 0$ *(the constraint set is thus the complement of a convex elliptic set). Four points are candidate since they satisfy the Karush–Kuhn–Tucker conditions for optimality* $(\alpha)'$ *and* $(\alpha)''$*:*

$(2, 0)$ *with the multiplier* $1/2$*;* $(-2, 0)$ *with the multiplier* $3/2$*;*

$\left(\frac{4}{3}, \pm\frac{\sqrt{5}}{3}\right)$ *with the multiplier* 1.

Condition $(\beta)'$ *is not satisfied at* $(\pm 2, 0)$, *but is satisfied at the two other points* $\left(\frac{4}{3}, \pm\frac{\sqrt{5}}{3}\right)$*: these last two points are the solutions of our problem.*

Let us now go back to the general quadratic-quadratic minimization problem $(\mathcal{P})$: what if there are *several* quadratic constraints? The situation is not as nice as when only one quadratic constraint were present; it has been explored by Yuan [26] for the case where $(\mathcal{P})$ has *two convex* quadratic inequality constraints and by Peng and Yuan [20] when $(\mathcal{P})$ has *two* (general) quadratic inequality constraints. Roughly speaking the situation is as follows:

- There is a gap between necessary conditions for global optimality and sufficient conditions for global optimality: the necessary conditions assert that the Hessian of the Lagrangian has *at most one strictly negative eigenvalue*; the sufficient conditions require that the Hessian of the Lagrangian be positive semidefinite.
- Even if they are stronger than the standard second-order necessary conditions for optimality, the necessary conditions we refer to above are not sufficient for optimality.
- Examples and counterexamples show the variety of situations and the fundamental differences with the case where only one quadratic constraint was present.

Here is the kind of results obtained for

$$(\mathcal{P}_4)\begin{cases}\text{Minimize } f(x) := \dfrac{1}{2}\langle A_0x, x\rangle + \langle b_0, x\rangle + c_0\\ \text{subject to}\\ \quad g_1(x) := \dfrac{1}{2}\langle A_1x, x\rangle + \langle b_1, x\rangle + c_1 \le 0\\ \quad \text{and}\\ \quad g_2(x) := \dfrac{1}{2}\langle A_2x, x\rangle + \langle b_2, x\rangle + c_2 \le 0.\end{cases}$$

THEOREM 4.5 [20, p. 589]. *Let $\bar{x}$ be a global minimizer of $(\mathcal{P}_4)$ satisfying $g_1(\bar{x}) = g_2(\bar{x}) = 0$. Assume that the gradients $\nabla g_1(\bar{x})$ and $\nabla g_2(\bar{x})$ are linearly independent. There then exist unique multipliers $\bar{\mu}_1 \ge 0$ and $\bar{\mu}_2 \ge 0$ such that:*

$(\alpha)'''$ $\nabla f(\bar{x}) + \bar{\mu}_1\nabla g_1(\bar{x}) + \bar{\mu}_2\nabla g_2(\bar{x}) = 0$*;*
$(\beta)'''$ $A_0 + \bar{\mu}_1A_1 + \bar{\mu}_2A_2$ *has at most one strictly negative eigenvalue.*

If only one constraint is active at $\bar{x}$, say $g_1(\bar{x}) = 0$ and $g_2(\bar{x}) < 0$, and if $\nabla g_1(\bar{x}) \neq 0$, there still exists $\bar{\mu}_1 \ge 0$ such that: $\nabla f(\bar{x}) + \bar{\mu}_1\nabla g_1(\bar{x}) = 0$ and $A_0 + \bar{\mu}_1A_1$ has at most one strictly negative eigenvalue. But this result directly follows from standard second-order necessary conditions for (local) optimality and does not have the strength of Theorem 4.5.

By just using second order developments (which are *exact* for quadratic functions), the following sufficient condition for global optimality is derived.

THEOREM 4.6 *Let $\bar{x}$ satisfy $g_1(\bar{x}) \le 0$ and $g_2(\bar{x}) \le 0$. Assume there exist multipliers $\bar{\mu}_1 \ge 0$ and $\bar{\mu}_2 \ge 0$ such that:*

$$\nabla f(\bar{x}) + \bar{\mu}_1\nabla g_1(\bar{x}) + \bar{\mu}_2\nabla g_2(\bar{x}) = 0,$$
$$\bar{\mu}_1 g_1(\bar{x}) = \bar{\mu}_2 g_2(\bar{x}) = 0;$$
$$A_0 + \bar{\mu}_1A_1 + \bar{\mu}_2A_2 \text{ is positive semidefinite.}$$

Then $\bar{x}$ is a global minimizer of $(\mathcal{P}_4)$.

The gap between necessary conditions and sufficient conditions for optimality in $(\mathcal{P}_4)$ is illustrated by the next example.

EXAMPLE 4.7 (corrected from [20, p. 591]). *Consider the following problem of the $(\mathcal{P}_4)$-type in $\mathbb{R}^3$:*

$$\begin{cases}\textit{Minimize } (x_1-1)^2 + x_2^2 - 6x_2 - 10x_3^2\\ \textit{subject to}\\ \quad x_1^2 + x_2^2 + x_3^2 - 2 \le 0,\\ \quad (x_1-2)^2 + x_2^2 + x_3^2 - 2 \le 0.\end{cases}$$

The necessary conditions exhibited in Theorem 4.5 are indeed satisfied at $\bar{x} = (1, 1, 0)$ with $\bar{\mu}_1 = \bar{\mu}_2 = 1$; the Hessian of the Lagrangian function is diag

(6,6,-16) *which has one strictly negative eigenvalue. But* $\bar{x}$ *is not a minimizer (not even a local one) since* $d/d\varepsilon\, f(x_\varepsilon) < 0$ *at* $\varepsilon = 0$, *if* $x_\varepsilon := (1, 1-\varepsilon, \sqrt{\varepsilon(2-\varepsilon)})$.

The reason why everything goes nicely when there is only one constraint in the quadratic-quadratic optimization problem ($\mathcal{P}$) and why the same is not expected when there are several constraints can be explained to a great extent via the duality theory (in [19, Section 5] in a hidden form, in [5, Chapter 3]), but we must admit additional efforts are necessary to clear this matter up.

Acknowledgments

I would like to thank the referee for his detailed and constructive report.

References

1. Danninger, G. and Bomze, I. (1993), Using copositivity for global optimality criteria in concave quadratic programming problems, *Mathematical Programming* 62: 575–580.
2. Dedieu, J.-P. (1995), Third- and fourth-order optimality conditions in optimization, *Optimization* 33: 97–104.
3. Dedieu, J.-P. and Janin, R. (1996), On third- and fourth-order optimality conditions in unconstrained programming, Technical report.
4. Dür, M., Horst, R. and Locatelli, M. (1998), Necessary and sufficient global optimality conditions for convex maximization revisited, *Journal of Mathematical Analysis and Applications* 217: 637–649.
5. Ferrier, C. (1997), *Bornes duales de problèmes d'optimisation polynomiaux*, Ph.D. Thesis, Laboratoire Approximation et Optimisation, Université Paul Sabatier, Toulouse.
6. Flippo, O.E. and Jansen, B. (1996), Duality and sensitivity in nonconvex quadratic optimization over an ellipsoid, *European Journal of Operational Research* 94: 167–178.
7. Floudas, C.A. and Visweswaran, V. (1995), Quadratic optimization, in: *Handbook for Global Optimization*, pp. 217–269. Kluwer Academic Publishers, Dordrecht.
8. Giner, E. (1995), Local minimizers of integral functionals are global minimizers, *Proceedings of the American Mathematical Society* 123(3): 755–757.
9. Glover, B.M., Ishizuka, Y., Jeyakumar, V. and Tuan, H.D. (1996), Complete characterization of global optimality for problems involving the pointwise minimum of sublinear functions, *SIAM J. Optimization* 6(2): 362–372.
10. Hiriart-Urruty, J.-B. (1989), From convex optimization to nonconvex optimization. Part 1: Necessary and sufficient conditions for global optimality, in: *Nonsmooth Optimization and Related Topics*, pp. 219–239, Plenum Press, New York.
11. Hiriart-Urruty, J.-B. (1995), Conditions for global optimality, in: *Handbook for Global Optimization*, pp. 1–26. Kluwer Academic Publishers, Dordrecht.
12. Hiriart-Urruty, J.-B. (1996), L'optimisation, in *Que sais-je?*, Presses Universitaires de France.
13. Hiriart-Urruty, J.-B. and Ledyaev, J.S. (1996), A note on the characterization of the global maxima of a (tangentially) convex function over a convex set, *J. of Convex Analysis* 3(1): 55–61.
14. Hiriart-Urruty, J.-B. and Lemaréchal, C. (1993), *Convex Analysis and Minimization Algorithms,* Vols. 305 and 306 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag.
15. Jiang, Y., Smith, W.R. and Chapman, G.R. (1995), Global optimality conditions and their geometric interpretation for the chemical and phase equilibrium problem, *SIAM J. Optimization* 5(4): 813–834.

16. Levy, A.V. and Montalvo, A. (1985), The tunneling algorithm for the global minimization of functions, *SIAM J. Sci. Stat. Comp.* 6(1): 15–29.
17. Martinez, J.M. (1994), Local minimizers of quadratic functions on Euclidean balls and spheres, *SIAM J. Optimization* 4(1): 159–176.
18. McKinnon, K. and Mongeau, M. (1998), A generic global optimization algorithm for the chemical and phase equilibrium problem [Technical Report 94-05, Laboratoire Approximation et Optimization], *J. of Global Optimization* 12(4): 325–351.
19. Moré, J.J. (1993), Generalizations of the trust region problem, *Optimization Methods and Software* 2: 189–209.
20. Peng, J.M. and Yuan, Y. (1997), Optimality conditions for the minimization of a quadratic with two quadratic constraints, *SIAM J. Optimization* 7(3): 579–594.
21. Pham Dinh Tao and Le Thi Hoai An (1996), Difference of convex functions optimization algorithms (dca) for globally minimizing nonconvex quadratic forms on euclidean balls and spheres, *Operations Research Letters* 19: 207–216.
22. Rendl, F. and Wolkowicz, H. (1997), A semidefinite framework for trust region subproblems with applications to large scale minimization, *Mathematical Programming* 77: 273–299.
23. Stern, R.J. and Wolkowicz, H. (1995), Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations, *SIAM J. Optimization* 5(2): 286–313.
24. Strekalovsky, A. (1993), Extremal problems on complements of convex sets, Translated from *Kibernetika i Sistemnyi Analiz* 1: 113–126.
25. Strekalovsky, A. (1996), On d.c.-programming theory, Technical report, Irkutsk State University.
26. Yuan, Y. (1990), On a subproblem of trust region algorithms for constrained optimization, *Mathematical Programming* 47: 53–63.