JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 75, No. 3, DECEMBER 1992

Role of Copositivity in Optimality Criteria for Nonconvex Optimization Problems

G. DANNINGER[1]

Communicated by G. Leitmann

Abstract. Second-order necessary and sufficient conditions for local optimality in constrained optimization problems are discussed. For global optimality, a criterion recently developed by Hiriart-Urruty and Lemarechal is thoroughly examined in the case of concave quadratic problems and reformulated into copositivity conditions.

Key Words. Copositive matrices, convex maximization problems, concave minimization, global optimality conditions.

1. Introduction

Nonconvex quadratic minimization problems over a polyhedron in n-dimensional Euclidian space $\mathbb{R}^n$ arise in different fields of application from combinatorial optimization to database problems and VLSI design. The solution of such problems is, from the perspective of worst-case complexity, NP-hard. Even checking whether a given feasible point is a local solution is also NP-hard (Refs. 1–2). For general nonconvex problems, a local criterion for global optimality seems to be impossible. However, for the special case of concave functions and especially for quadratic functions, there exists a criterion for global optimality of a feasible point that may be seen as a local one.

The paper is organized as follows. First, we will take a short look on local necessary and sufficient second-order conditions for optimization under constraints in Section 2. Section 3 is devoted to global optimality conditions, and Section 4 to their translation into copositivity conditions. In this last

[1]Assistant Professor, Department of Statistics and Computer Sciences, University of Vienna, Vienna, Austria.

section, we will also see the analogy between global and local optimality conditions in the special case under consideration.

2. Second-Order Conditions for Optimality Under Constraints

Consider the problem to minimize $f(x)$, subject to $x \in M$, with

$$M := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ 1 \leq i \leq m\},$$

where $f: \mathbb{R}^n \to \mathbb{R}$ and $g_i: \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq m$, are twice Fréchet differentiable. The first-order approximation of M around an admissible point $x \in M$ is the tangent cone of M at x,

$$\Gamma := \{u \in \mathbb{R}^n : Dg_i(x)u \leq 0,\ i \in I(x)\},$$

with

$$I(x) := \{i \in \{1, \ldots, n\} : g_i(x) = 0\}$$

denoting the set of binding constraints and D the derivative.

The intuitive idea that it suffices to study the local behavior of f around x on Γ rather than on M to obtain optimality conditions can be made precise in the following way: if $\bar{x}$ is a Karush–Kuhn–Tucker point with Lagrange multipliers $\lambda_i \geq 0$, $1 \leq i \leq m$, we put $\lambda = [\lambda_1, \ldots, \lambda_m]^T \in \mathbb{R}^m$, where T signifies transposition of a column vector, and denote by

$$h_\lambda(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

the Lagrange function at $\bar{x}$. Then, $h_\lambda(\bar{x}) = f(\bar{x})$ and $Dh_\lambda(\bar{x}) = o$ holds, i.e., $\bar{x}$ is a critical point of h_λ. Now, the natural generalization of the elementary fact that a critical point with positive-definite Hessian (matrix of the second-order derivatives) yields a strict local minimum in an unconstrained problem, is that every Karush–Kuhn–Tucker point yields a strict local minimum under constraints, provided the Hessian H_λ of the Lagrangian h_λ is strictly Γ-copositive. Here and in the sequel, a symmetric $n \times n$ matrix H is said to be Γ-copositive iff

$$v^T H v \geq 0, \qquad \text{for all } v \in \Gamma,$$

and strictly Γ-copositive iff

$$v^T H v > 0, \qquad \text{for all } v \in \Gamma \setminus \{o\}.$$

Theorem 2.1. If $\bar{x} \in M$ is a Karush–Kuhn–Tucker point and $H_\lambda = D^2 h_\lambda(\bar{x})$ is strictly Γ-copositive, then $\bar{x}$ is a strict local solution to the problem.

Proof. For a concise proof, see Ref. 3. □

By analogy, one might try to generalize also the necessary condition of positive semidefiniteness of the Hessian in the unconstrained case to Γ-copositivity of H_λ in the constrained case. But the following example shows that this is not possible:

Example 2.1. Consider the problem to minimize $f(x)=\log x$, subject to $x\geq 1$. Then, at the (global) solution $\bar{x}=1$, we have $\Gamma=[0,\infty)$ and $H_\lambda=-1$, which is certainly not Γ-copositive.

This means that we have to modify the above conditions in some way, which is possible as we will show now. The Karush–Kuhn–Tucker conditions imply

$$Df(\bar{x})v\geq 0, \qquad \text{for all } v\in\Gamma.$$

If a direction $v\in\Gamma$ yields a strict inequality above, then the sign of the second-order term $v^T H_\lambda v$ is irrelevant, because in this case one works directly with the Taylor expansion of f,

$$f(x)=f(\bar{x})+Df(\bar{x})v+o(\|v\|)>f(\bar{x}), \qquad \text{if } v=x-\bar{x} \text{ is small},$$

where by $\|x\|=\sqrt{x^T x}$ we denote the Euclidian norm of a vector $x\in\mathbb{R}^n$. Therefore, we may and do shrink the tangential cone Γ to the subcone

$$\Gamma^*=\{v\in\Gamma: Df(\bar{x})v=o\}.$$

Given a set of Lagrange multipliers λ, we put

$$M_\lambda^*=\{x\in M: f(x)=h_\lambda(x)\}=\{x\in M: g_i(x)=0, \text{ if } \lambda_i>0\}.$$

Since

$$\Gamma^*=\{v\in\Gamma: Dg_i(\bar{x})v=0, \text{ if } \lambda_i>0\},$$

this cone can be viewed as the tangential cone of M_λ^*. Contrasting with Γ^*, the set M_λ^* depends not only on the local shape of M and f, but also on λ (see Example 2.2 below). In Refs. 4–5, the last identity is used to define Γ^*, from which it might not immediately be evident that Γ^* does not depend on λ.

As we shall see now, under a constraint qualification on M_λ^*, any local solution $\bar{x}$ yields a Hessian H_λ of h_λ that is Γ^*-copositive [observe that, in the above example, $\Gamma^*=\{0\}$ holds]. Furthermore, as noted already in Ref. 1, also the reverse direction holds: strict Γ^*-copositivity of H_λ suffices to guarantee that $\bar{x}$ is a local solution to the problem.

Theorem 2.2. Let $\bar{x} \in M$ be a Karush–Kuhn–Tucker point with the Lagrange multipliers $\lambda = [\lambda_1, \ldots, \lambda_m]^T$, and denote by $H_\lambda = D^2 h_\lambda(x)$ the Hessian of the Lagrangian

$$h_\lambda(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

(a) Suppose that $\bar{x}$ is a local minimizer of f on M which satisfies Abadie's constraint qualification w.r.t. M_λ^*, i.e., assume that every $u \in \Gamma^*$ is the starting tangent vector of a trajectory in M_λ^* starting at $\bar{x}$ (see, e.g., Ref. 6). Then, H_λ is Γ^*-copositive.

(b) Conversely, if H_λ is strictly Γ^*-copositive, then $\bar{x}$ is a local solution to the problem.

Proof. See Ref. 7 or Ref. 5, p. 61. □

Note that Abadie's constraint qualifications w.r.t. M_λ^* are implied by the following conditions corresponding to the constraint qualifications of Ref. 8:

(i) the gradients $\{\nabla g_i(\bar{x}) : \lambda_i > 0\}$ are linearly independent;

(ii) there is a direction $v \in \mathbb{R}^n$ satisfying

$$Dg_i(\bar{x})v < 0, \qquad \text{if } \lambda_i = 0 \text{ and } i \in I(\bar{x});$$

$$Dg_i(\bar{x})v = 0, \qquad \text{if } \lambda_i > 0 \text{ and } i \in I(\bar{x}).$$

Example 2.2. Consider the example essentially due to Ref. 5, p. 62: Minimize $f^j(x)$, subject to

$$g_1(x) = -x_1^2 - x_2 - x_3 \leq 0, \qquad g_2(x) = -x_1^2 + x_2 - x_3 \leq 0,$$

$$g_3(x) = -x_3 \leq 0.$$

We investigate three different objective functions,

$$f^+(x) = (1/2)x_1^2 + x_3, \qquad f^-(x) = -(1/2)x_1^2 + x_3,$$

$$f^0(x) = x_1^4 + x_3.$$

Now,

$$Df^j(o) = [0, 0, 1], \qquad \text{for all } j \in \{-, 0, +\},$$

so that, for any choice of $t \in [0, 1/2]$, the point $\bar{x} = o$ satisfies the Karush–Kuhn–Tucker conditions with

$$\lambda = \lambda(t) = [t, t, 1 - 2t]^T$$

for all three objectives f^j,

$$I(o)=\{1,2,3\}$$

and

$$\begin{bmatrix}0\\0\\1\end{bmatrix}+t\begin{bmatrix}0\\-1\\-1\end{bmatrix}+t\begin{bmatrix}0\\1\\-1\end{bmatrix}+(1-2t)\begin{bmatrix}0\\0\\-1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Also, the cone Γ^* is given by the set of all vectors $[v_1, 0, 0]^T$ where $v_1 \in \mathbb{R}$ is arbitrary, while

$$M^*_{\lambda(t)}=\begin{cases}\{x\in\mathbb{R}^3: |x_2|\le x_1^2,\ x_3=0\}, & \text{if } t=0,\\ \{o\}, & \text{if } 0<t\le 1/2,\end{cases}$$

so that Abadie's constraint qualification is satisfied only if $t=0$. Note that condition (i) above is violated for $0<t\le 1/2$, and even for $t=0$ condition (ii) fails to hold: indeed, the relations

$$[0,-1,-1]v<0, \qquad [0,1,-1]v<0, \qquad [0,0,1]v=0,$$

are contradictory for any $v\in\mathbb{R}^3$.

On the feasible set, the point $\bar{x}=o$ is even a global minimizer of the functions f^+ and f^0, but no local minimizer of f^-. Now, the Hessian matrices of the Lagrange functions corresponding to f^+, f^0, f^-, are given by

$$H^+_{\lambda(t)}=\begin{bmatrix}1-2t & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}, \qquad H^-_{\lambda(t)}=\begin{bmatrix}-1-2t & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},$$

$$H^0_{\lambda(t)}=\begin{bmatrix}-2t & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},$$

respectively; hence, $H^+_{\lambda(t)}$ is Γ^*-copositive for all $t\in[0,1/2]$; $H^-_{\lambda(t)}$ is not Γ^*-copositive for all $t\in[0,1/2]$; and $H^0_{\lambda(t)}$ is not Γ^*-copositive unless $t=0$. This example shows that:

(i) even under Abadie's constraint qualification, the converse of Theorem 2.2(b) is not true; the same holds for Theorem 2.2(a) if one replaces $f^0(x)$ by $-x_1^4+x_3$;

(ii) the constraint qualifications on M^*_λ cannot be dispensed with to obtain the implication of Theorem 2.2(a);

(iii) the local optimality of $\bar{x}$ and Γ^*-copositivity of H_λ may be correlated to each other although Abadie's constraint qualifications are violated;

(iv) the Cottle–Mangasarian–Fromovitz conditions are more restrictive than Abadie's constraint qualification.

The next example shows that Abadie's constraint qualification is not hereditary from M to M_λ^*.

Example 2.3. Let $m=n=2$ and

$$g_1(x_1, x_2) = \exp(-x_1) + x_1 - x_2 - 1,$$

$$g_2(x_1, x_2) = \exp(x_1) - x_1 - x_2 - 1,$$

which have at $\bar{x}=0$ the same derivative

$$Dg_1(o) = Dg_2(o) = [0, -1].$$

Hence, for

$$M = \{x \in \mathbb{R}^2 : g_i(x) \leq 0,\ i=1,2\},$$

even the Cottle–Mangasarian–Fromovitz constraint qualifications at the point $\bar{x}=o$ are satisfied: indeed, we have $I(o)=\{1,2\}$ and, for $v=[0,1]^T$, we obtain

$$Dg_i(\bar{v})v < 0, \qquad \text{for all } i.$$

The tangential cone is

$$\Gamma = \{v \in \mathbb{R}^2 : v_2 \geq 0\}.$$

For any objective function f with derivative $Df(o)=[0,1]$, the point $\bar{x}=o$ satisfies the Karush–Kuhn–Tucker conditions, any admissible set λ of Lagrange multipliers fulfilling $\lambda_1+\lambda_2=1$. Now,

$$\Gamma^* = \{v \in \mathbb{R}^2 : v_2 = 0\}.$$

If both $\lambda_1>0$ and $\lambda_2>0$, then $M_\lambda^* = \{o\}$, and Abadie's condition is obviously violated. If, however, $\lambda_1=1$ and $\lambda_2=0$, then

$$\begin{aligned} M_\lambda^* &= \{x \in M : g_1(x)=0\} \\ &= \{x \in \mathbb{R}^2 : x_2 \geq \exp(x_1) - x_1 - 1 \text{ and } x_2 = \exp(-x_1) + x_1 - 1\} \\ &= \{x \in \mathbb{R}^2 : \sinh x_1 \leq x_1 \text{ and } x_2 = \exp(-x_1) + x_1 - 1\} \\ &= \{x \in \mathbb{R}^2 : x_1 \leq 0 \text{ and } x_2 = \exp(-x_1) + x_1 - 1\}, \end{aligned}$$

which also violates Abadie's condition, because $v=[1, 0]^T \in \Gamma^*$ cannot be a starting tangent vector of any trajectory in $M_\lambda^\sharp$ starting in $\bar{x}=o$. Similarly also for $\lambda=[0, 1]^T$, Abadie's constraint qualification is not met by

$$M_\lambda^\sharp=\{x \in \mathbb{R}^2: x_2=\exp(x_1)-x_1-1 \text{ and } x_1 \geq 0\}.$$

3. Global Optimality Conditions for Concave Quadratic Problems

In Ref. 9, Hiriart-Urruty studied necessary and sufficient conditions for global optimality of minimization problems with a concave objective function, or equivalently, the problem

$$h(x) \rightarrow \max, \qquad \text{s.t. } x \in C, \tag{1}$$

where h is a convex function and C a convex set. These conditions for global optimality of a point $\bar{x} \in C$ involve essentially approximates of the subdifferential of a convex function h at $\bar{x}$ (see Rockafellar, Ref. 10) and approximates of the normal cone of C at $\bar{x}$ as defined below.

Definition 3.1. The ϵ-subgradient of a function $h: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ [where $\bar{\mathbb{R}}= \mathbb{R} \cup \{+\infty\}$, h not identically equal to $+\infty$] in $\bar{x}$ is defined to be the set

$$\partial_\epsilon h(\bar{x})=\{y \in \mathbb{R}^n: h(x) \geq h(\bar{x})+y^T(x-\bar{x})-\epsilon, \text{ for all } x \in \mathbb{R}^n\}.$$

Definition 3.2. The set of the ϵ-normal directions to C (nonempty, closed, convex) at $\bar{x}$ is defined to be

$$N_\epsilon(C; \bar{x})=\{y \in \mathbb{R}^n: y^T(x-\bar{x}) \leq \epsilon, \text{ for all } x \in C\}.$$

Furthermore, let us define an approximate directional derivative in the following way.

Definition 3.3. The ϵ-directional derivative of the function h in the point $\bar{x}$ into the direction d is defined to be

$$h'_\epsilon(\bar{x}, d):=\inf_{\lambda>0}([h(\bar{x}+\lambda d)-h(\bar{x})+\epsilon]/\lambda).$$

Remark 3.1. The set $\partial_\epsilon h(\bar{x})$ reduces to the usual subdifferential $\partial h(\bar{x})$ for $\epsilon=0$ and also the set $N_\epsilon(C; \bar{x})$ reduces to the usual cone of normal directions to C at $\bar{x}$ for $\epsilon=0$, but in general $N_\epsilon(C; \bar{x})$ is no longer a cone if $\epsilon>0$.

Remark 3.2. In the infinite-dimensional case, one has to assume lower semicontinuity of h; see Hiriart-Urruty, Ref. 9.

Remark 3.3. Characterizing the convex set $C \neq \emptyset$ by its indicator function

$$\Psi_C(x) = \begin{cases} 0, & \text{if } x \in C, \\ +\infty, & \text{if } x \notin C, \end{cases}$$

one notes that

$$\partial_\epsilon \Psi_C(\bar{x}) = N_\epsilon(C; \bar{x}).$$

Indeed, consider (a) and (b) below:

(a) If $y \in N_\epsilon(C; \bar{x})$, then

$$y^T(x - \bar{x}) \leq \epsilon, \qquad \text{for all } x \in C,$$

and it follows that

$$y^T(x - \bar{x}) + \Psi_C(\bar{x}) \leq \epsilon + \Psi_C(x), \qquad \text{for all } x \in \mathbb{R}^n.$$

Recall that $\Psi_C(x) = 0$, $x \in C$, and $\Psi_C(x) = +\infty$, $x \notin C$, while $\Psi_C(\bar{x}) = 0$, since $\bar{x} \in C$. Thus, $y \in \partial_\epsilon \Psi_C(\bar{x})$.

(b) If $y \in \partial_\epsilon \Psi_C(\bar{x})$, then

$$y^T(x - \bar{x}) + \Psi_C(\bar{x}) \leq \epsilon + \Psi_C(x), \qquad \text{for all } x \in \mathbb{R}^n.$$

For $x \in C$, we have $\Psi_C(x) = 0$, so that

$$y^T(x - \bar{x}) \leq \epsilon, \qquad \text{for all } x \in C,$$

and therefore $y \in N_\epsilon(C; \bar{x})$.

Remark 3.4. The ϵ-directional derivative of Ψ_C (see Definition 3.3) is given by

$$\begin{aligned}(\Psi_C)'_\epsilon(\bar{x}, d) &= \inf_{\lambda > 0} \{ [\Psi_C(\bar{x} + \lambda d) - \Psi_C(\bar{x}) + \epsilon]/\lambda \} \\ &= \inf\{\epsilon/\lambda \colon \lambda > 0 \text{ such that } \bar{x} + \lambda d \in C\},\end{aligned}$$

where we set $\inf \emptyset = +\infty$, from the definition of Ψ_C.

We now formulate the following theorem for general convex objective function:

Theorem 3.1. Let h be a convex function, and let C be a convex set. For the problem (1), a feasible point $\bar{x} \in C$ is a global maximum of h on C iff

$$\partial_\epsilon h(\bar{x}) \subset N_\epsilon(C; \bar{x}), \qquad \text{for all } \epsilon \geq 0. \tag{2}$$

Proof. A proof can be found in Ref. 9. We will show below a direct proof for the special case of a quadratic objective function. □

However, relation (2) is in general not very helpful. We will show how to translate it into a calculable criterion in Section 4. The idea in Ref. 11 is to exploit in the quadratic case the special structure of h and C in order to simplify relation (2). So we will now study in more detail the quadratic concave problem, or equivalently the problem

$$(1/2)x^TQx + c^Tx \to \max, \qquad \text{s.t. } Ax \le b; \tag{3}$$

here, Q is a symmetric, positive-semidefinite $n \times n$ matrix; $c \in \mathbb{R}^n$; A is an $m \times n$ matrix; and $b \in \mathbb{R}^m$. For this special quadratic case, we will prove Theorem 3.1 directly and reformulate the condition (2) into copositivity conditions.

Remark 3.5. Since both $S(\epsilon) = \partial_\epsilon h(\bar{x})$ and $N(\epsilon) = N_\epsilon(C; \bar{x})$ are convex sets, the inclusion (2) holds iff

$$\sigma_{S(\epsilon)}(d) \le \sigma_{N(\epsilon)}(d), \qquad \text{for all directions } d \in \mathbb{R}^n, \tag{4}$$

where for a set $Y \subset \mathbb{R}^n$ we denote by

$$\sigma_Y(d) = \sup\{d^Ty : y \in Y\}$$

the support functional of Y. This follows by standard separation arguments.

For the proof of Theorem 3.1 (quadratic case), we will first describe the set of ϵ-normal directions and the ϵ-subgradient explicitly. To this end, decompose the polyhedron C in the following way:

$$C = \text{co}(x_1, \ldots, x_k) + \text{pos}(r_1, \ldots, r_l),$$

where we denote by $\text{co}(x_1, \ldots, x_k)$ the convex hull of the extremal points $x_1, \ldots, x_k$ and by $\text{pos}(r_1, \ldots, r_l)$ the cone generated by the extremal rays $r_1, \ldots, r_l$ (see Ref. 12) and use the following lemma.

Lemma 3.1. If C is a convex polyhedron, the set of ϵ-normal directions to C at $\bar{x}$ can be written in the form

$$\begin{aligned} N(\epsilon) &= \{y \in \mathbb{R}^n : (x_i - \bar{x})^T y \le \epsilon,\ 1 \le i \le k;\ r_j^T y \le 0,\ 1 \le j \le l\} \\ &= \{y \in \mathbb{R}^n : B^T y \le e\}, \end{aligned}$$

with $B = [x_1 - \bar{x}, \ldots, x_k - \bar{x}; r_1, \ldots, r_l]$ an $n \times (k+l)$ matrix and $e = [\epsilon, \ldots, \epsilon; 0, \ldots, 0]^T$ a $(k+l)$-vector.

Proof.

(a) If for some $y \in \mathbb{R}^n$ we have

$$(x_i - \bar{x})^T y \leq \epsilon, \ 1 \leq i \leq k \quad \text{and} \quad r_j^T y \leq 0, \ 1 \leq j \leq 1,$$

it follows that, for all $x \in C$, we have

$$\begin{aligned}(x-\bar{x})^T y &= \left[\sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{l} \mu_j r_j - \bar{x}\right]^T y \\ &= \sum_{i=1}^{k} \lambda_i (x_i - \bar{x})^T y + \sum_{j=1}^{l} \mu_j r_j^T y \leq \epsilon,\end{aligned}$$

and so $y \in N(\epsilon) = N_\epsilon(C, \bar{x})$.

(b) On the other hand, if $y \in N(\epsilon)$, then:

(b1) $(x_i - \bar{x})^T y \leq \epsilon$, $1 \leq i \leq k$, since $x_i \in C$.

(b2) If there were j such that $r_j^T y > 0$, we set $z = \lambda r_j + \bar{x}$. Then, $z \in C$ for all $\lambda > 0$. From $y \in N(\epsilon)$, we obtain $\epsilon \geq (z - \bar{x})^T y = \lambda r_j^T y$. But since $r_j^T y > 0$, and λ can be chosen arbitrarily large, this cannot be true, and so $r_j^T y \leq 0$ for all $1 \leq j \leq l$. □

We can now write

$$N(\epsilon) = \{y \in \mathbb{R}^n : B^T y \leq e\}, \tag{5}$$

with $B = [x_1 - \bar{x}, \ldots, x_k - \bar{x}; r_1, \ldots, r_l]$ an $n \times (k+l)$ matrix and $e = [\epsilon, \ldots, \epsilon; 0, \ldots, 0]^T$ a $(k+l)$-vector.

Lemma 3.2. The ϵ-subgradient of $h(x) = (1/2)x^T Q x + c^T x$ at $\bar{x}$, with Q positive semidefinite, is of the form

$$S(\epsilon) = \partial_\epsilon h(\bar{x}) = Q\bar{x} + c + \{u \in \text{Im}\, Q : u^T Q^+ u \leq 2\epsilon\},$$

Q^+ being the Moore–Penrose inverse of Q. This set can also be written as

$$S(\epsilon) = \partial_\epsilon h(\bar{x}) = Q\bar{x} + c + \{u \in \mathbb{R}^n : u^T Q^{-1} u \leq 2\epsilon\},$$

if Q is positive definite.

Proof. First note that, due to Definition 3.1, $d \in \partial_\epsilon h(\bar{x})$ iff

$$h(x) - h(\bar{x}) \geq d^T(x - \bar{x}) - \epsilon, \qquad \text{for all } x \in \mathbb{R}^n,$$

which is equivalent to

$$x^T Q x - \bar{x}^T Q \bar{x} + 2c^T(x - \bar{x}) \geq 2d^T(x - \bar{x}) - 2\epsilon, \qquad \text{for all } x \in \mathbb{R}^n.$$

If we set now

$$d = Q\bar{x} + c + u,$$

we obtain

$$\begin{aligned} d \in \partial_\epsilon h(\bar{x}) &\Leftrightarrow x^T Q x - \bar{x}^T Q \bar{x} + 2c^T(x - \bar{x}) \\ &\geq 2(Q\bar{x} + c + u)^T (x - \bar{x}) - 2\epsilon, \qquad \forall x \in \mathbb{R}^n \\ &\Leftrightarrow x^T Q x - \bar{x}^T Q \bar{x} - 2x^T Q \bar{x} + 2\bar{x}^T Q \bar{x} \\ &\geq 2u^T (x - \bar{x}) - 2\epsilon, \qquad \forall x \in \mathbb{R}^n \\ &\Leftrightarrow (x - \bar{x})^T Q (x - \bar{x}) \geq 2u^T (x - \bar{x}) - 2\epsilon, \qquad \forall x \in \mathbb{R}^n. \end{aligned} \tag{6}$$

(a) If $d = Q\bar{x} + c + u$, with $u \in \operatorname{Im} Q$ (there exists some v such that $u = Qv$) and $u^T Q^+ u \leq 2\epsilon$, we obtain, since Q is positive semidefinite,

$$\begin{aligned} 0 &\leq (x - \bar{x} - v)^T Q (x - \bar{x} - v) \\ &= (x - \bar{x})^T Q (x - \bar{x}) - 2v^T Q (x - \bar{x}) + v^T Q v \\ &= (x - \bar{x})^T Q (x - \bar{x}) - 2u^T (x - \bar{x}) + v^T Q Q^+ Q v \\ &= (x - \bar{x})^T Q (x - \bar{x}) - 2u^T (x - \bar{x}) + u^T Q^+ u, \end{aligned}$$

and because of (6) we see that $d \in \partial_\epsilon h(\bar{x})$.

(b) For the other direction:

(b1) If $d \in \partial_\epsilon h(\bar{x})$, put $u = d - Q\bar{x} - c$ and substitute x by $\bar{x} + Q^+ u$ in (6) to obtain immediately that $u^T Q^+ u \leq 2\epsilon$.

(b2) If we replace x by $\bar{x} + \lambda (I - QQ^+) u$ in (6), where I is the $n \times n$ identity matrix, $\lambda \geq 0$ arbitrary, we have

$$\begin{aligned} d \in \partial_\epsilon h(\bar{x}) &\Rightarrow \lambda^2 u^T (I - QQ^+)^T Q (I - QQ^+) u \\ &\qquad - 2\lambda u^T (I - QQ^+) u + 2\epsilon \geq 0 \\ &\Rightarrow u^T (I - QQ^+) u \leq \epsilon / \lambda \\ &\Rightarrow \|u - QQ^+ u\|^2 \leq \epsilon / \lambda. \end{aligned}$$

Since ϵ is fixed and λ may be chosen arbitrarily large, we obtain $u = QQ^+ u$ and so $u \in \operatorname{Im} Q$. □

In the following we calculate and reinterpret the support functional $\sigma_{N(\epsilon)}(d)$ of the ϵ-normal directions in the case of a polyhedron.

Lemma 3.3. The following result holds:

$$\sigma_{N(\epsilon)}(d) = \inf\{\epsilon/\lambda : \lambda > 0 \text{ such that } \bar{x} + \lambda d \in C\}$$
$$= (\Psi_C)'_\epsilon(\bar{x}, d), \qquad \text{by Remark 3.4.}$$

Proof. If $\sigma_{N(\epsilon)}(d) < \infty$, it takes the solution value of the following linear program:

$$\text{(D)} \quad d^T y \to \max,$$
$$B^T y \leq e,$$

which is dual to

$$\text{(P)} \quad e^T w \to \min,$$
$$Bw = d,$$
$$w \geq o,$$

where B and e are given as in (5). Now, (P) can be reformulated as

$$\text{(P}') \quad \epsilon \sum_{i=1}^{k} u_i \to \min,$$
$$\sum_{i=1}^{k} u_i(x_i - \bar{x}) + \sum_{j=1}^{l} v_j r_j = d,$$
$$u_i \geq o,$$
$$v_j \geq o,$$

where we set $w = \begin{bmatrix} u \\ v \end{bmatrix}$. Three possible cases arise:

Case 1. If $u_i = 0$ for all i and all u that constitute a feasible point w for (P′), then

$$d = \sum_{j=1}^{l} v_j r_j$$

is an unbounded ray of C and

$$\bar{x} + \lambda d \in C, \qquad \text{for all } \lambda > 0.$$

The optimal value of (P′) is zero and hence $\sigma_{N(\epsilon)}(d) = 0$ by the duality theorem.

Case 2. If $u_i > 0$ for some u constituting a feasible point $w = \begin{bmatrix} u \\ v \end{bmatrix}$, we obtain

$$\sigma_{N(\epsilon)}(d) = \inf\{\epsilon/\lambda : \lambda > 0 \text{ such that } \bar{x} + \lambda d \in C\};$$

indeed, for

$$\lambda = 1 \Big/ \sum_{i=1}^{k} u_i,$$

from the equality constraints of (P′) it follows that

$$\bar{x} + \lambda d = (1/\lambda) \sum_{i=1}^{k} u_i x_i + (1/\lambda) \sum_{j=1}^{l} v_j r_j \in C;$$

see the representation of C for Lemma 3.1.

Case 3. If (P′) or (P) has no feasible point, then (D) is unbounded. In this case, there cannot exist a $\lambda > 0$ such that

$$\bar{x} + \lambda d = \sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{l} \mu_j r_j$$

is in C, where

$$\lambda_i, \mu_i \geq 0, \qquad \sum_{i=1}^{k} \lambda_i = 1,$$

since otherwise $w = [{}^u_v]$, with $u_i = \lambda_i/\lambda$ and $v_j = \mu_j/\lambda$, would be a feasible point.

So, taking the three cases together, we can write

$$\sigma_{N(\epsilon)}(d) = \inf\{\epsilon/\lambda \colon \lambda > 0, \text{ such that } \bar{x} + \lambda d \in C\}. \qquad \square$$

Now we are in the position to prove Theorem 3.1 in the special case of maximizing a convex quadratic function over a convex polyhedron:

Proof of Theorem 3.1. Quadratic Case.

(i) Let us first consider the necessary direction. $\bar{x} \in C$ a global maximum means that $h(\bar{x}) \geq h(x)$ for all $x \in C$. Let now be $y \in \partial_\epsilon h(\bar{x}) = S(\epsilon)$. We have

$$h(x) \geq h(\bar{x}) + y^T(x - \bar{x}) - \epsilon, \qquad \text{for all } x \in \mathbb{R}^n.$$

Since $h(x) - h(\bar{x}) \leq 0$, we immediately obtain

$$y^T(x - \bar{x}) \leq \epsilon, \qquad \text{for all } x \in \mathbb{R}^n,$$

and so $y \in N_\epsilon(C; \bar{x}) = N(\epsilon)$, for all $\epsilon \geq 0$.

(ii) To show the more sophisticated sufficient direction, we take four steps, where $h'_\epsilon(\bar{x}, d)$ is the ϵ-directional derivative of h (see Definition 3.3):

(a) $h(\bar{x} + d) - h(\bar{x}) = \sup_{\epsilon > 0}\{h'_\epsilon(\bar{x}, d) - \epsilon\}$,

for all d such that $\bar{x} + d \in C$;

(b) $h'_\epsilon(\bar{x}, d) \leq \sigma_{S(\epsilon)}(d)$, for all d such that $\bar{x}+d \in C$, $\forall \epsilon \geq 0$;

(c) $\sigma_{S(\epsilon)}(d) \leq \sigma_{N(\epsilon)}(d)$ holds because of (4) for all $\epsilon \geq 0$ and all d such that $\bar{x}+d \in C$;

(d) $\sup_{\epsilon>0}\{\sigma_{N(\epsilon)}(d) - \epsilon\} \leq 0$, for all d such that $\bar{x}+d \in C$.

Combining these four steps, we obtain the sufficient part: if $S(\epsilon) \subset N(\epsilon)$, then

$$h(\bar{x}+d) - h(\bar{x}) \leq 0$$

holds for all d such that $\bar{x}+d \in C$.

It remains to prove the assertions (a), (b), (d).

Assertion (a). Let

$$\phi(\lambda) = h(\bar{x}+\lambda d) - h(\bar{x}) = (1/2)\lambda^2 d^T Q d + \lambda (Q\bar{x}+c)^T d,$$

so that $\phi(0)=0$, ϕ is convex,

$$\phi'(\lambda) = \lambda d^T Q d + (Q\bar{x}+c)^T d, \qquad \phi''(\lambda) = d^T Q d,$$

and

$$h'_\epsilon(\bar{x}, d) = \inf_{\lambda>0}([\phi(\lambda)+\epsilon]/\lambda).$$

Now, our first assertion is that

$$\mu\phi(1/\mu) = \sup_{\epsilon>0}\left\{\left(\inf_{\lambda>0}[\phi(\lambda)+\epsilon]/\lambda\right) - \epsilon\mu\right\}, \qquad \text{for all } \mu>0. \tag{7}$$

Indeed, since

$$\inf_{\lambda>0}[\phi(\lambda)+\epsilon]/\lambda - \epsilon\mu \leq ([\phi(1/\mu)+\epsilon]/(1/\mu)) - \epsilon\mu = \mu\phi(1/\mu),$$

it follows that

$$\sup_{\epsilon>0}\left\{\left(\inf_{\lambda>0}[\phi(\lambda)+\epsilon]/\lambda\right) - \epsilon\mu\right\} \leq \mu\phi(1/\mu).$$

To prove the reverse inequality, we have to show that

$$\sup_{\epsilon>0}\left\{\left(\inf_{\lambda>0}\,[\phi(\lambda)+\epsilon]/\lambda\right) - \epsilon\mu\right\} \geq \mu\phi(1/\mu), \qquad \text{for all } \mu>0,$$

or its equivalent (replace $1/\lambda$ by ν),

$$\sup_{\epsilon>0}\left\{\inf_{\nu>0}(\nu\phi(1/\nu)-\mu\phi(1/\mu)+\epsilon(\nu-\mu))\right\}\geq 0, \qquad \text{for all } \mu>0.$$

Denote by $\rho(\lambda)=\lambda\phi(1/\lambda)$. Then, ρ is convex for $\lambda>0$, and so

$$\rho(\nu)-\rho(\mu)\geq\rho'(\mu)(\nu-\mu), \qquad \text{for all } \nu>0.$$

Hence,

$$\inf_{\nu>0}[\rho(\nu)-\rho(\mu)+\epsilon(\nu-\mu)]\geq\inf_{\nu>0}[(\rho'(\mu)+\epsilon)(\nu-\mu)]. \tag{8}$$

Differentiation of ρ gives

$$\rho'(\mu)=-(1/2\mu^2)d^TQd.$$

Substituting this into (8), we obtain

$$\inf_{\nu>0}[(\rho'(\mu)+\epsilon)(\nu-\mu)]=\begin{cases}-\infty, & \text{if } \epsilon<(1/2\mu^2)d^TQd,\\ 0, & \text{if } \epsilon=(1/2\mu^2)d^TQd,\\ -\mu(\rho'(\mu)+\epsilon), & \text{if } \epsilon>(1/2\mu^2)d^TQd.\end{cases}$$

Therefore,

$$\sup_{\epsilon>0}\left\{\inf_{\nu>0}(\nu\phi(1/\nu)-\mu\phi(1/\mu)+\epsilon(\nu-\mu))\right\}$$

$$\geq\sup_{\epsilon>0}\left\{\inf_{\nu>0}[(\rho'(\mu)+\epsilon)(\nu-\mu)]\right\}=0, \qquad \text{for all } \mu>0.$$

The last equation is evident if $d^TQd>0$ [the supremum is attained for $\epsilon=(1/2\mu^2)d^TQd$], while for $d^TQd=0$ [i.e., $\rho'(\mu)=0$] we have

$$\inf_{\nu>0}[(\rho'(\mu)+\epsilon)(\nu-\mu)]=-\mu\epsilon, \qquad \text{for all } \epsilon>0,$$

and thus

$$\sup_{\epsilon>0}\left\{\inf_{\nu>0}[(\rho'(\mu)+\epsilon)(\nu-\mu)]\right\}=0.$$

Finally by

$$\mu\phi(1/\mu)\leq\sup_{\epsilon>0}\left\{\left(\inf_{\lambda>0}[\phi(\lambda)+\epsilon]/\lambda\right)-\epsilon\mu\right\}, \qquad \text{for all } \mu>0,$$

assertion (7) is proved. Putting $\mu = 1$ in (7), it follows that

$$h(\bar{x}+d) - h(\bar{x}) = \sup_{\epsilon>0}\{h'_\epsilon(\bar{x}, d) - \epsilon\}, \qquad \text{for all } d \text{ with } \bar{x}+d \in C.$$

Assertion (b). We have

$$\begin{aligned} h'_\epsilon(\bar{x}, d) &= \inf_{\lambda>0}\{[\phi(\lambda)+\epsilon]/\lambda\} \\ &= \inf_{\lambda>0}\{(1/\lambda)[(1/2)\lambda^2 d^T Q d + \lambda(Q\bar{x}+c)^T d + \epsilon]\} \\ &= \inf_{\lambda>0}\{(1/2)[\lambda d^T Q d + 2(Q\bar{x}+c)^T d + 2\epsilon/\lambda]\}. \end{aligned}$$

Differentiating with respect to λ, we obtain

$$h'_\epsilon(\bar{x}, d) = \sqrt{2\epsilon}\sqrt{d^T Q d} + (Q\bar{x}+c)^T d.$$

Let now

$$\begin{aligned} z^T &= [\sqrt{2\epsilon}/\sqrt{d^T Q d}]d^T Q + (Q\bar{x}+c)^T, \qquad &\text{if } d^T Q d > 0, \\ z^T &= (Q\bar{x}+c)^T, &\text{if } d^T Q d = 0. \end{aligned}$$

Then,

$$z^T d = h'_\epsilon(\bar{x}, d) \quad \text{and} \quad z \in \partial_\epsilon h(\bar{x}),$$

since z is of the form

$$z = Q\bar{x} + c + u,$$

with

$$u = \sqrt{2\epsilon/d^T Q d}\, Qd, \qquad u \in \operatorname{Im} Q,$$

and

$$u^T Q^+ u = (2\epsilon/d^T Q d) d^T Q Q^+ Q d = 2\epsilon, \qquad \text{if } d^T Q d > 0,$$

and

$$u = o \in \operatorname{Im} Q, \text{ with } u^T Q^+ u = o \leq 2\epsilon, \qquad \text{if } d^T Q d = 0.$$

So we can say that

$$h'_\epsilon(\bar{x}, d) = z^T d \leq \sup\{d^T y : y \in \partial_\epsilon h(\bar{x})\} = \sigma_{S(\epsilon)}(d).$$

Assertion (d). Studying $\sigma_{N(\epsilon)}(d)$ in more detail with the help of Lemma 3.3, one notices that, since $\bar{x}+d \in C$, it follows that $\sigma_{N(\epsilon)}(d) \leq \epsilon$ for all $\epsilon > 0$. Indeed, if we take in Lemma 3.3 the point $w = [\begin{smallmatrix} u \\ v \end{smallmatrix}]$, with $u_i = \lambda_i$ and $v_j = \mu_j$, if

$$\bar{x} = \sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{l} \mu_j r_j,$$

with

$$\lambda_i, \mu_i \geq 0, \qquad \sum_{i=1}^{k} \lambda_i = 1,$$

see Lemma 3.1, then w is feasible for (P′) and the value of the objective function is equal to ϵ.

So, we have shown that

$$\sup_{\epsilon > 0} (\sigma_{N(\epsilon)}(d) - \epsilon) \leq 0, \qquad \text{if } \bar{x} + d \in C.$$

Concluding, we can see now easily that

$$h(\bar{x} + d) - h(\bar{x}) = \sup_{\epsilon > 0} \{h'_\epsilon(\bar{x}, d) - \epsilon\} \leq \sup_{\epsilon > 0} \{\sigma_{S(\epsilon)}(d) - \epsilon\}$$
$$\leq \sup_{\epsilon > 0} \{\sigma_{N(\epsilon)}(d) - \epsilon\} \leq 0,$$

and so finally we obtain that

$$h(\bar{x} + d) - h(\bar{x}) \leq 0, \qquad \text{if } \bar{x} + d \in C. \qquad \square$$

4. Using Copositivity for Checking Global Optimality

In the form (2), the relation of Theorem 3.1 is, as mentioned above, not very useful. But by taking a closer look at the special structure in the quadratic case, it will be possible to develop copositivity criteria for global optimality. First, we will study again the support functional of $N(\epsilon)$ and that of $S(\epsilon)$. Let

$$h(x) = (1/2) x^T Q x + c^T x,$$

with Q a symmetric positive-semidefinite matrix and c an n-dimensional vector. Then, the special structure of $S(\epsilon) = \partial_\epsilon h(\bar{x})$ and $N(\epsilon) = N_\epsilon(C; \bar{x})$ is of the form presented in the following two lemmas.

Lemma 4.1. We have

$$\sigma_{S(\epsilon)}(d) = d^T (Q\bar{x} + c) + \sqrt{2\epsilon d^T Q d}. \tag{9}$$

Proof. The value of the support function $\sigma_{S(\epsilon)}(d) = \sup_{y \in S(\epsilon)} y^T d$ is the solution of the following problem:

$$(\mathrm{Q}_\epsilon) \quad d^T y \to \max,$$
$$y \in [Q\bar{x} + c + \{u \in \operatorname{Im} Q : u^T Q^+ u \leq 2\epsilon\}],$$

which is equivalent to

$$(\mathrm{Q}'_\epsilon) \quad (Qd)^T v \to \max,$$
$$(1/2) v^T Q v \leq \epsilon.$$

Let v be a feasible solution of (Q'_ϵ). Then, we always have

$$d^T Q v = (\sqrt{Q}\, d)^T (\sqrt{Q}\, v) \leq \|\sqrt{Q}\, d\|\, \|\sqrt{Q}\, v\|$$
$$= \sqrt{d^T Q d} \cdot \sqrt{v^T Q v} \leq \sqrt{2\epsilon} \cdot \sqrt{d^T Q d};$$

and for

$$\bar{v} = \begin{cases} \sqrt{2\epsilon / d^T Q d}\, d, & \text{if } d^T Q d > 0, \\ o, & \text{if } d^T Q d = 0, \end{cases}$$

which is feasible for (Q'_ϵ), we obtain

$$d^T Q \bar{v} = \sqrt{2\epsilon} \cdot \sqrt{d^T Q d}.$$

Returning to the original problem (Q_ϵ), and remembering that $Q\bar{v} = \bar{u}$, we obtain the solution

$$\bar{y} = Q\bar{x} + c + \bar{u}, \qquad \text{with } \bar{u} = Q\bar{v}.$$

For the support functional, we thus have

$$\sigma_{S(\epsilon)}(d) = d^T \bar{y} = d^T (Q\bar{x} + c) + \sqrt{2\epsilon d^T Q d}.$$ □

Lemma 4.2. We have

$$\sigma_{N(\epsilon)}(d) = \epsilon z(d),$$

with

$$z(d) = \begin{cases} \max[\{0\} \cup \{(Ad)_i / u_i : i \notin I(\bar{x})\}], & \text{if } d \in \Gamma, \\ +\infty, & \text{otherwise}, \end{cases} \tag{10}$$

where we denote by

$$I = I(\bar{x}) := \{i \in \{1, \ldots, n\} : (A\bar{x})_i = b_i\}$$

the set of binding constraints at $\bar{x}$, by

$$u_i := b_i - (A\bar{x})_i > 0,$$

the slack variables at $\bar{x}$ for $i \notin I(\bar{x})$, and by

$$\Gamma = \{d: (Ad)_i \leq 0, i \in I(\bar{x})\},$$

the tangential cone of C in $\bar{x}$, if

$$C = \{x \in \mathbb{R}^n: Ax \leq b\};$$

see (3).

Proof. Recalling Lemma 3.3, we know that

$$\sigma_{N(\epsilon)}(d) = \epsilon \inf\{1/\lambda: \lambda > 0, \text{ such that } \bar{x} + \lambda d \in C\}.$$

(a) If $d \notin \Gamma$, then we have $\inf \varnothing = +\infty$.

(b) If $d \in \Gamma$, then the infimum must be taken at $\lambda = 0$ if in direction d the polytope is unbounded. If in direction d the polytope is bounded, the infimum is taken for a point $y = \bar{x} + \tilde{\lambda} d$ fulfilling $(Ay)_i = b_i$ for at least one $i \notin I(\bar{x})$. Noting by short calculation that $1/\lambda = (Ad)_i/u_i$, we obtain in that case

$$z(d) = \max[\{0\} \cup \{(Ad)_i/u_i : i \notin I(\bar{x})\}]. \qquad \square$$

Remark 4.1. Solving the linear program

$$(\mathrm{P}_\epsilon) \quad (Ad)^T v \to \max,$$
$$(b - A\bar{x})^T v \leq \epsilon,$$
$$v \geq o,$$

we obtain the same solution as for the problem

$$\sigma_{N(\epsilon)}(d) = \sup_{y \in N(\epsilon)} d^T y$$

above by noticing that the vertices of the polytope

$$S = \{v \in \mathbb{R}^n: (b - A\bar{x})^T v \leq 1, v \geq o\}$$

are the origin and $(1/u_i)e_i$ for all $i \notin I(\bar{x})$, where e_i is the ith vector of the standard basis in $\mathbb{R}^n$. Since this is true for all directions d, we can conclude that, the support functionals being equal, the corresponding convex sets $N(\epsilon)$ and $\{A^T v, v \geq o, (b - A\bar{x})^T v \leq \epsilon\}$ must also be equal (see Remark 3.5). In this way, one can obtain a representation of $N(\epsilon)$ alternative to that in Lemma 3.1.

Now, (4) [equivalent to (2)] can be reformulated in the quadratic case (3) into

$$f_d(\delta) = \delta^2 z(d) - \delta\sqrt{2d^T Qd} - d^T(Q\bar{x} + c) \geq 0, \qquad \text{for all } d \in \mathbb{R}^n,$$

with $\delta = \sqrt{\epsilon}$. Note that always $z(d) \geq 0$, and thus f_d is a convex function. So, instead of (4), we shall for the quadratic case check the inequality

$$f_d(\delta) \geq 0, \qquad \text{for all } \delta \geq 0,$$

where $d \in \mathbb{R}^n$ is fixed, but arbitrary.

According to (10), the relation $f_d(\delta) \geq 0$ is clearly satisfied for all $\delta \geq 0$, if $d \notin \Gamma$, so we only have to investigate directions d belonging to the tangential cone, as one would expect. In case of $z(d) > 0$, the function f_d attains its minimal value at

$$\delta^* = \sqrt{2d^T Qd}/2z(d) > 0,$$

so that we have only to check that

$$f_d(\delta^*) = -[1/2z(d)]d^T Qd - d^T(Q\bar{x} + c) \geq 0,$$

which can be rephrased as

$$-d^T Qd - 2d^T(Q\bar{x} + c)z(d) \geq 0. \tag{11}$$

This relation also has to hold if $z(d) = 0$, since then f_d is affine and thus has to have a nonnegative slope in order to be nonnegative for arbitrarily large δ. If we now denote, for $i \in \{1, \ldots, m\} \backslash I(\bar{x})$,

$$\Gamma_i = \{d \in \Gamma : (Ad)_i \geq 0 \text{ and } u_j(Ad)_i \geq u_i(Ad)_j, \text{ for all } j \in \{1, \ldots, m\} \backslash I(\bar{x})\}, \tag{12}$$

then Γ_i is a polyhedral cone satisfying

$$\Gamma_i = \{d \in \Gamma : z(d) = (Ad)_i/u_i\}.$$

Similarly,

$$\Gamma_0 = \{d \in \Gamma : (Ad)_i \leq 0 \text{ for all } i \in \{1, \ldots, m\} \backslash I(\bar{x})\} \tag{13}$$

is a polyhedral cone with

$$\Gamma_0 = \{d \in \mathbb{R}^n : Ad \leq o\} = \{d \in \Gamma : z(d) = 0\}.$$

Then, condition (11) can further be reformulated into the conditions

$$d^T Q_i d \geq 0, \qquad \text{for all } d \in \Gamma_i \text{ and all } i \in \{0, \ldots, n\} \backslash I(\bar{x}), \tag{14}$$

where the symmetric $n \times n$ matrices Q_i are defined by

$$Q_i = \begin{cases} -Q, & \text{if } i=0, \\ B_i - u_i Q, & \text{otherwise,} \end{cases} \tag{15}$$

and

$$B_i = -a_i(Q\bar{x}+c)^T - (Q\bar{x}+c)(a_i)^T, \tag{16}$$

where $(a_i)^T$ denotes the ith row of A.

Conditions (14) alone do not suffice to ensure validity of (4), and hence global optimality of $\bar{x}$. Indeed, in the case of $z(d)=0$, where f_d is an affine function, not only the slope of f_d has to be nonnegative to ensure

$$f_d(\delta) \geq 0, \qquad \text{for all } \delta \geq 0.$$

In addition, the relation $f_d(0) \geq 0$ has to hold in order to guarantee $f_d(\delta) \geq 0$ also for small values of δ. Now, observe that the condition

$$0 \leq f_d(0) = -d^T(Q\bar{x}+c) = -d^T \nabla h(\bar{x}), \qquad \text{for all } d \in \Gamma, \tag{17}$$

exactly corresponds to the Karush–Kuhn–Tucker conditions. Hence, for a Karush–Kuhn–Tucker point $\bar{x}$, the weaker condition

$$d^T(Q\bar{x}+c) \leq 0, \qquad \text{for all } d \in \Gamma_0, \tag{18}$$

is automatically satisfied. So (14) and (18) together ensure (4), and hence global optimality, but the latter can be ignored if $\bar{x}$ is a Karush–Kuhn–Tucker point. Note that (18) is a boundedness condition: indeed, any direction $d \in \Gamma_0$, with

$$(Q\bar{x}+c)^T d > 0,$$

satisfies

$$\bar{x} + td \in C, \qquad \text{for all } t \geq 0,$$

as well as

$$\begin{aligned} h(\bar{x}+td) - h(\bar{x}) &= (t^2/2) d^T Q d + t d^T (Q\bar{x}+c) \\ &\geq t d^T(Q\bar{x}+c) \to \infty, \quad \text{as } t \to \infty. \end{aligned}$$

There is another unboundedness condition which is independent of the current feasible point $\bar{x}$: if $Q_0 = -Q$ is not Γ_0-copositive, i.e., if there is a direction $d \in \Gamma_0$ with $d^T Q d > 0$, then as above, $\bar{x} + td \in C$, for all $t \geq 0$, as well as

$$h(\bar{x}+td) - h(\bar{x}) = (t^2/2) d^T Q d + t d^T (Q\bar{x}+c) \to \infty, \qquad \text{as } t \to \infty.$$

Let us recapitulate the above arguments in the following theorem.

Theorem 4.1. Let $\bar{x}$ be a feasible point of the quadratic problem (3). Define Q_i and Γ_i according to (12), (13), (15), and (16). Then the following assertions are equivalent:

(a) $\bar{x}$ is a global solution to (3);

(b) $\bar{x}$ is a local solution to (3) and

$$Q_i \text{ is } (\Gamma_i \backslash \Gamma^*)\text{-copositive for all } i \in \{0, \ldots, m\} \backslash I(\bar{x}),$$

$$\Gamma^* = \{d \in \Gamma : Dh(\bar{x})d = o\}$$
$$= \{d \in \mathbb{R}^n : (Ad)_i \le 0, \text{ if } (A\bar{x})_i = b_i, \text{ and } (Q\bar{x} + c)^T d = 0\};$$

(c) $\bar{x}$ is a Karush-Kuhn-Tucker point of (3) and

$$Q_i \text{ is } \Gamma_i\text{-copositive for all } i \in \{0, \ldots, m\} \backslash I(\bar{x});$$

(d) $\bar{x}$ satisfies $d^T(Q\bar{x} + c) \le 0$, for all $d \in \Gamma_0$, rather than for all $d \in \Gamma$ as in (c), and

$$Q_i \text{ is } \Gamma_i\text{-copositive for all } i \in \{0, \ldots, m\} \backslash I(\bar{x}).$$

Proof. See Ref. 7. □

To stress similarity of local and global optimality criteria, one might compare Theorem 4.1(c) with the following, rather inefficient version of Theorem 3.1 in Ref. 7: a Karush-Kuhn-Tucker point $\bar{x}$ is a local solution of (3) iff

$$Q_i \text{ is } (\Gamma_i \cap \Gamma^*)\text{-copositive for all } i \in \{0, \ldots, m\} \backslash I(\bar{x}).$$

In any case, checking global optimality involves at most $m-1$ copositivity problems of the same structure like that arising in the check of local optimality.

Remark 4.2. If one wants to show (d)⇒(c) directly without using (a) or (b), i.e., without using the knowledge that linearity of constraints makes constraint qualifications superfluous, then one may use the observation that, for an $i \in \{1, \ldots, m\} \backslash I(\bar{x})$ and any direction $d \in \Gamma_i \backslash \Gamma_0$, we have

$$(a_i)^T d = (Ad)_i > 0,$$

so that the inequality

$$0 \le d^T Q_i d = -2(a_i)^T d (Q\bar{x} + c)^T d - u_i d^T Q d \le -2(a_i)^T d (Q\bar{x} + c)^T d$$

immediately yields the desired relation $d^T(Q\bar{x} + c) \le 0$, which establishes the Karush-Kuhn-Tucker condition (17).

From a computational point of view, formulations (c) or (d) of Theorem 4.1 seem to be preferable, since several procedures for checking Γ-copositivity are available if Γ is a polyhedral cone. See, e.g., Refs. 13–17, but of course these algorithms suffer from a high worst-case complexity. Nevertheless, the above criteria may be useful in finding global optima in the quadratic case. A corresponding algorithm, which is able to escape from local solutions, has been developed in Ref. 18.

References

1. MURTY, K. G., and KABADI, S. N., *Some NP-Complete Problems in Quadratic and Linear Programming*, Mathematical Programming, Vol. 39, pp. 117–129, 1987.
2. PARDALOS, P. M., and SCHNITGER, G., *Checking Local Optimality in Constrained Quadratic Programming Is NP-Hard*, OR Letters, Vol. 7, pp. 33–35, 1988.
3. BOMZE, I. M., *Copositivity and Optimization*, Proceedings of the 12th SOR Meeting, Athenäum Verlag, Frankfurt, Germany, pp. 27–36, 1989.
4. HESTENES, M. R., *Optimization Theory: The Finite-Dimensional Case*, Wiley, New York, New York, 1975.
5. FLETCHER, R., *Practical Methods of Optimization, Vol. 2: Constrained Optimization*, Wiley, New York, New York, 1981.
6. BAZARAA, M. S., and SHETTY, C. M., *Nonlinear Programming: Theory and Algorithms*, Wiley, New York, New York, 1979.
7. DANNINGER, G., and BOMZE, I. M., *Using Copositivity for Local and Global Optimality Criteria in Smooth Nonconvex Programming Problems*, Technical Report 103, Department of Statistics and Computer Sciences, University of Vienna, Vienna, Austria, 1991.
8. MANGASARIAN, O. N., and FROMOVITZ, S., *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints*, Journal of Mathematical Analysis and Applications, Vol. 17, pp. 37–47, 1967.
9. HIRIART-URRUTY, J. B., *From Convex Optimization to Nonconvex Optimization, Part 1: Necessary and Sufficient Conditions for Global Optimality*, Nonsmooth Optimization and Related Topics, Edited by F. H. Clarke et al., Plenum Press, New York, New York, pp. 219–239, 1989.
10. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
11. HIRIART-URRUTY, J. B., and LEMARECHAL, C., *Testing Necessary and Sufficient Conditions for Global Optimality in the Problem of Maximizing a Convex Quadratic Function over a Convex Polyhedron*, Preliminary Report, Seminar of Numerical Analysis, University Paul Sabatier, Toulouse, France, 1990.
12. WETS, R., *Grundlagen Konvexer Optimierung*, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, Germany, Vol. 137, 1976.

13. DIANANDA, P. H., *On Nonnegative Forms in Real Variables Some or All of Which Are Nonnegative*, Proceedings of the Cambridge Philosophical Society, Vol. 58, pp. 17–25, 1962.
14. COTTLE, R. W., HABETLER, G. J., and LEMKE, C. E., *Quadratic Forms Semidefinite over Convex Cones*, Proceedings of the Princeton Symposium in Mathematical Programming, Edited by H. W. Kuhn, Princeton University Press, Princeton, New Jersey, pp. 551–565, 1970.
15. HADELER, K. P., *On Copositive Matrices*, Linear Algebra and Its Applications, Vol. 49, pp. 79–89, 1983.
16. BOMZE, I. M., *Remarks on the Recursive Structure of Copositivity*, Journal of Information and Optimization Science, Vol. 8, pp. 243–260, 1987.
17. DANNINGER, G., *A Recursive Algorithm for Determining Strict Copositivity of a Symmetric Matrix*, Methods of Operations Research, Vol. 62, pp. 45–52, 1990.
18. BOMZE, I. M., and DANNINGER, G., *A Global Optimization Algorithm for Concave Linear-Quadratic Problems*, SIAM Journal on Optimization, 1992 (to appear).