An Optimality Criterion for Global Quadratic Optimization*

ARNOLD NEUMAIER
Institut für Angewandte Mathematik, Universität Freiburg, D-7800 Freiburg, Germany.

(Received: 20 February 1991; accepted: 24 October 1991)

Abstract. In this paper we prove a sufficient condition that a strong local minimizer of a bounded quadratic program is the unique global minimizer. This sufficient condition can be verified computationally by solving a linear and a convex quadratic program and can be used as a quality test for local minimizers found by standard indefinite quadratic programming routines.

Key words. Global optimization, quadratic program, lower bound, branch and bound.

1. Introduction

In this paper we discuss the problem

$$\begin{aligned}&\text{global min} \quad h(x) := \gamma + c^T x + \frac{1}{2} x^T H x \\ &\text{subject to} \quad x \in C := \{x \in \mathbb{R}^n \mid \underline{b} \leq Ax \leq \bar{b}\}\,, \end{aligned} \tag{1}$$

where $H \in \mathbb{R}^{n \times n}$ is symmetric, $c \in \mathbb{R}$, $\gamma \in \mathbb{R}$, $A \in \mathbb{R}^{m \times n}$, and $\underline{b}, \bar{b} \in \mathbb{R}^m$. For simplicity of exposition we only discuss the case of finite $\underline{b}, \bar{b}$. It is not difficult to extend the method to the case of one-sided bounds where one or several entries of $\underline{b}$ or $\bar{b}$ are $\pm\infty$. As long as C remains bounded, this only involves trivial changes.

When H is positive semidefinite and C is nonempty then (1) is a convex problem and standard quadratic programming methods yield a local optimizer which is global by convexity. On the other hand, if H is indefinite then, often, many local optima exist, and the selection of the global optimum has a combinatorial aspect. In the extreme case where H is negative definite, every vertex of C is a local optimizer, and the combinatorial structure is obvious. This shows that global quadratic optimization is a computationally hard problem (see also Murty & Kabadi [4]). Nevertheless, problems with several hundred variables have been successfully solved when the number of negative eigenvalues of H were not too large, see Pardalos & Rosen [5]. Other references for global constrained optimization are Fujii *et al.* [1], Hansen & Sengupta [2], and Horst & Tuy [3].

In the present paper we show that there is a constructive measure of nonconvexity such that for "nearly convex" problems, a local optimizer can be recog-

*Part of this work was done while the author was at the University of Wisconsin-Madison.

nized to be a global optimizer. We expect the result to be of practical significance (see Section 4); however, in this paper, only the theoretical aspects will be analyzed.

2. Auxiliary Results

Basic to our results is the following bound for a positive semidefinite quadratic form on the cone of nonnegative vectors.

PROPOSITION 1. *Let $G \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite, and let $u, v \in \mathbb{R}^n$ be nonnegative vectors satisfying $u_i v_i \geq G_{ii}$ for $i = 1, \ldots, n$. Then*

$$0 \leq x \in \mathbb{R}^n \Rightarrow x^T G x \leq (u^T x)(v^T x) . \tag{2}$$

Proof. By continuity it suffices to treat the case $u > 0$. For fixed $\varepsilon > 0$ we consider the optimization problem

$$\begin{aligned} &\text{global min} \quad \varepsilon v^T x - x^T G x \\ &\text{subject to} \quad x \geq 0 , \quad u^T x = \varepsilon . \end{aligned} \tag{3}$$

(3) is bounded and concave, hence attains its global minimum at a vertex. But the vertices are $x = u_i^{-1} \varepsilon e^{(i)}$, and the corresponding objective function values are $u_i^{-2}(u_i v_i - G_{ii})\varepsilon^2 \geq 0$. Hence the global minimum of (3) is nonnegative, and since $\varepsilon > 0$ was arbitrary, (2) follows. □

Although not needed in the sequel, let us mention the following special case, obtained by minimizing the upper bound in (2) subject to the constraints $u_i v_i \geq G_{ii}$ for $i = 1, \ldots, n$.

COROLLARY 2. *Let $G \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite. Then*

$$0 \leq x \in \mathbb{R}^n \Rightarrow x^T G x \leq \Big(\sum_{i=1}^{n} \sqrt{G_{ii}}\, x_i \Big)^2 . \qquad \square$$

Now let x be a local minimizer of (1), let

$$g = c + Hx \tag{4}$$

denote the gradient of the objective function of x, and let $y \in \mathbb{R}^m$ be the vector of Lagrange multipliers. Then the first order optimality conditions give

$$g^T = y^T A , \tag{5}$$

$$\inf\{ y_i(\underline{b} - Ax)_i ,\ y_i(\bar{b} - Ax)_i \} = 0 \quad (i = 1, \ldots, m) . \tag{6}$$

Let $B = \{ i \mid y_i \neq 0 \}$, let y_B be the vector y restricted to the rows indexed by B, and let A_B be the matrix A restricted to the rows indexed by B. We now make the first nondegeneracy assumption that A_B has rank $|B|$ and

$$\underline{b}_i < (Ax)_i < \bar{b}_i \Longleftrightarrow i \notin B . \tag{6a}$$

CASE 1. $|B| = n$ (i.e., the optimum is in a vertex of C). In particular, this is always the case if H is negative semidefinite, or if $A = I$ (simple bounds) and $H_{ii} \leq 0$ for $i = 1, \ldots, n$. In this case A_B is nonsingular. By means of a LDL^T factorization or a modified Cholesky factorization of $H' = A_B^{-T} H A_B^{-1} + \gamma I$ we can write $H' = NN^T - G$ with a positive semidefinite matrix G; then the matrix $H + A_B^T G A_B$ is positive definite. (If H is positive semidefinite we can of course choose $G = 0$.)

CASE 2. $|B| < n$. In this case, by permuting the columns of A (and the rows of x) if necessary, we may assume that $A_B = (A_1 A_2)$ with a nonsingular $|B| \times |B|$ – matrix A_1. Then the columns of the matrix

$$Z = \begin{pmatrix} -A_1^{-1} A_2 \\ I \end{pmatrix}$$

form a basis of the null space A_B, and with

$$P = \begin{pmatrix} -A_1^{-1} \\ 0 \end{pmatrix}$$

we have $A_B Z = 0$, $A_B P = I$, and $(Z\,P)$ is a nonsingular $n \times n$-matrix. (A similar decomposition can be obtained from a QR-factorization of A_B.) In view of (6a), the second order optimality conditions imply that the *reduced Hessian* $Z^T H Z$ is positive semidefinite at the local optimizer. We now make the second nondegeneracy assumption that (in case 2) the reduced Hessian is nonsingular, and hence positive definite. In this case, we have a Cholesky factorization $Z^T H Z = LL^T$ with nonsingular L. Now define

$$M = (P^T H Z) L^{-T} ,$$

$$H' = P^T H P - MM^T .$$

By means of a LDL^T decomposition or a modified Cholesky decomposition of H' we can write

$$H' = NN^T - G$$

with a positive semidefinite symmetric matrix G; moreover, we can choose $G = 0$ when H' is positive semidefinite. Now

$$\begin{aligned} (Z\,P)^T (H + A_B^T G A_B)(Z\,P) &= \begin{pmatrix} Z^T \\ P^T \end{pmatrix} H (Z\,P) + \begin{pmatrix} 0 \\ I \end{pmatrix} G (0\,I) \\ &= \begin{pmatrix} Z^T H Z & Z^T H P \\ P^T H Z & P^T H P + G \end{pmatrix} = \begin{pmatrix} LL^T & LM^T \\ ML^T & MM^T + NN^T \end{pmatrix} \\ &= \begin{pmatrix} L & 0 \\ M & N \end{pmatrix} \begin{pmatrix} L^T & M^T \\ 0 & N^T \end{pmatrix} \end{aligned}$$

is positive semidefinite, and since $(Z\,P)$ is nonsingular, $H + A_B^T G A_B$ is also positive semidefinite. Thus we have proved constructively:

PROPOSITION 3. *Let x be a strong local minimizer of* (1), *i.e.*, (6a) *holds*, A_B *has rank* $|B|$ *and either* $|B| = n$ *or the reduced Hessian at x is positive definite. Then there is a symmetric* $|B| \times |B|$*-matrix G such that both G and* $H + A_B^T G A_B$ *are positive semidefinite.* □

REMARKS. 1. We can choose $G = 0$ precisely when H is positive semidefinite, i.e., when the problem (1) is convex.

2. Verifying local optimality in the presence of degeneracy or a singular reduced Hessian is NP-hard [6], and verifying global optimality in this situation seems very involved. On the other hand, strong local minimizers can be recognized easily by verifying the assumptions of the proposition, and we can hope for detecting global optimality in this case, too.

3. It is conceivable that the proposition holds for arbitrary local minimizers. If so, the strongness assumption in the next section could be dropped since the results there only depend on the existence of G and not directly on strong optimality.

4. Note that any optimization problem with a local optimizer at x can be perturbed by suitable arbitrarily small perturbations to produce a problem with a strong local optimizer at x.

3. Main Results

THEOREM 4. *Let x be a strong local minimizer of* (1). *With the notation introduced above, and G chosen as in Proposition* 3, *let a_B be the vector indexed by B such that*

$$a_i = G_{ii}/(2y_i) \quad (i \in B)\,.$$

Then, for all $\tilde{x} \in C$,

$$g^T(\tilde{x} - x) \geq 0\,, \tag{7}$$

$$h(\tilde{x}) - h(x) \geq g^T(\tilde{x} - x)\cdot(1 - a_B^T A_B(\tilde{x} - x))\,. \tag{8}$$

In particular, if the number

$$\alpha := \max\{a_B^T A_B(\tilde{x} - x) \mid \tilde{x} \in C\} \tag{9}$$

satisfies $\alpha \leq 1$ then x is a global minimizer of (1), *and if $\alpha < 1$ then x is the unique global minimizer.*

Proof. Note first that for any $\tilde{x} \in C$ we have

$$h(\tilde{x}) = h(x) + g^T(\tilde{x} - x) + \frac{1}{2}(\tilde{x} - x)^T H(\tilde{x} - x)\,, \tag{10}$$

and $g^T(\tilde{x} - x) = y_B^T A_B(\tilde{x} - x)$ by (5). Writing

$$\Sigma = \mathrm{Diag}(y_B)\,, \quad e = (1, \ldots, 1)^T\,, \quad z = \Sigma A_B(\tilde{x} - x)\,, \tag{11}$$

we can rephrase this as

$$g^T(\tilde{x} - x) = e^T z\,. \tag{12}$$

Now $z_i = y_i(A\tilde{x} - Ax)_i \geq \inf\{y_i(\underline{b} - Ax)_i,\, y_i(\bar{b} - Ax)_i\} = 0$ by (6), hence $z \geq 0$, and $z = 0$ only if $A_B(\tilde{x} - x) = 0$. In particular,

$$g^T(\tilde{x} - x) \geq 0, \text{ with equality iff } A_B(\tilde{x} - x) = 0\,. \tag{13}$$

Since G and $H_0 := H + A_B^T G A_B$ are positive semidefinite we have

$$\begin{aligned} 0 &\leq \frac{1}{2}(\tilde{x} - x)^T H_0(\tilde{x} - x) \\ &= \frac{1}{2}(\tilde{x} - x)^T (H + A_B^T G A_B)(\tilde{x} - x) \\ &\underset{(10)}{=} h(\tilde{x}) - h(x) - g^T(\tilde{x} - x) + \frac{1}{2} z^T \Sigma^{-1} G \Sigma^{-1} z \\ &\underset{(*)}{\leq} h(\tilde{x}) - h(x) - g^T(\tilde{x} - x) + (e^T z)(v^T z)/2 \\ &\underset{(12)}{=} h(\tilde{x}) - h(x) - g^T(\tilde{x} - x)(1 - v^T z/2)\,. \end{aligned} \tag{14}$$

The step (*) follows from Proposition 1, where $v_i = (\Sigma^{-1} G \Sigma^{-1})_{ii} = G_{ii}/y_i^2 = 2a_i/y_i$. Now

$$v_i z_i = 2a_i(A_B(\tilde{x} - x))_i\,,$$

hence $v^T z = 2a_B^T A_B(\tilde{x} - x)$, and the inequality (14) becomes (8). Using (9) we get $v^T z \leq 2\alpha$, and (14) implies

$$0 \leq \frac{1}{2}(\tilde{x} - x)^T H_0(\tilde{x} - x) \leq h(\tilde{x}) - h(x) - (1 - \alpha) g^T(\tilde{x} - x)\,. \tag{15}$$

Finally, if $\alpha \leq 1$ then (8) gives $h(\tilde{x}) \geq h(x)$ for all $\tilde{x} \in C$; hence x is a global minimizer. And if $\alpha < 1$ and $\tilde{x}$ is an arbitrary global minimizer then (8) gives $h(\tilde{x}) = h(x)$ and $g^T(\tilde{x} - x) = 0$, so that $A_B(\tilde{x} - x) = 0$ by (14). If $|B| = n$ this implies $\tilde{x} = x$. And if $|B| < n$ then this implies $\tilde{x} - x = Zw$ for some vector w. But now (10) implies $w^T Z^T H Z w = 0$, and since the reduced Hessian is definite, $w = 0$. Therefore $\tilde{x} = x$, so that x is the unique global minimizer. □

COROLLARY 5. *In the above setting, any $\tilde{x} \in C$ with $h(\tilde{x}) \leq h(x)$, $\tilde{x} \neq x$ satisfies the inequality*

$$a_B^T A_B(\tilde{x} - x) \geq 1\,. \tag{16}$$

□

EXAMPLE. In the bound constrained case in 2 dimensions we have w.l.o.g.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \bar{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

We consider two particular cases (not exhausting all possibilities, but fairly general) where the global optimizer is verified.

Case 1: $H_{11} < 0$, $H_{22} < |H_{12}| + |c_2|$. In this case $h(x)$ is concave in x_1 and convex in x_2, and one easily sees that the only candidates for a local minimizer have

$$x_1 = \pm 1\,, \quad x_2 = (-c_2 \mp H_{12})/H_{22}\,,$$

with objective function value

$$h(x) = \gamma \pm c_1 + \frac{1}{2} H_{11} - (c_2 \pm H_{12})^2/2H_{22}\,.$$

Gradient and Lagrange multiplier vector have the common value

$$y_1 = g_1 = p \mp q\,, \quad y_2 = p_2 = 0$$

where

$$p = c_1 - c_2 H_{12}/H_{22}\,, \quad q = H_{12}^2/H_{22} - H_{11}\ (>0)\,.$$

Since $B = \{1\}$, $A_B = (1\,0)$, the reduced Hessian H_{22} is definite. Thus we have a strong local minimizer when $\mp y_1 > 0$, i.e.

$$q \mp p > 0\,.$$

Since the objective function value can be written as $\gamma - q/2 - c_2^2/2H_{22} \pm p$, the global minimizer is obtained for the sign which makes $\mp p$ nonnegative. For the choice $G = (q)$ the matrix $H + A_B^T G A_B$ is easily seen to be positive semidefinite. Hence $a_1 = q/2(p \mp q)$. Now (9) yields

$$\alpha = \max\{a_1(\tilde{x}_1 \mp 1) \mid \tilde{x}_1, \tilde{x}_2 \in [-1, 1]\}$$
$$= |a_1| \mp a_1 = q/(q \mp p)\,.$$

For the global minimizer, $\mp p \geq 0$, hence $\alpha < 1$; i.e., the criterion of the theorem recognizes global optimality.

Case 2: $H_{11} \leq 0$, $H_{22} < 0$. In this case, the global minimizer is at a vertex, and w.l.o.g., we assume it to be at $x_1 = x_2 = -1$. Comparison of the objective function at the vertices shows

$$c_1 \geq H_{12}\,, \quad c_2 \geq H_{12}\,, \quad c_1 + c_2 \geq 0\,.$$

Now $B = \{1, 2\}$, $A_B = A = I$, and $G = \left(\begin{smallmatrix} q & 0 \\ 0 & 0 \end{smallmatrix}\right)$ with the above q makes $H + {A_B}^T G A_B$ positive semidefinite. We find

$$y_1 = g_1 = c_1 - H_{11} - H_{12}(\geq 0)\,, \quad y_2 = c_2 - H_{22} - H_{12}(\geq 0)\,,$$
$$a_1 = q/2y_1 \geq 0\,, \quad a_2 = 0\,.$$

Now

$q \leq -H_{11} \leq y_1$, hence $a_1 \leq 1/2$ and

$$\alpha = \max\{a_1(\tilde{x}_1 + 1) + a_2(\tilde{x}_2 + 1) \mid \tilde{x}_1, \tilde{x}_2 \in [-1, 1]\}$$
$$= \max\{0, 2a_1\} \leq 1\,.$$

Thus the global optimizer is again recognized.

REMARKS. 1. The number α in (9) can be viewed as a *measure of nonconvexity* since $\alpha = 0$ if $G = 0$, i.e., if H is positive semidefinite. In particular, Theorem 4 proves global optimality of an optimizer of any "nearly convex" quadratic program. It is easy to see from (1) and (9) that

$$\alpha \leq |a_B|^T(\bar{b} - \underline{b})_B = \frac{1}{2} \sum_{i \in B} d_i G_{ii}/|y_i|\,, \quad d_i = \bar{b}_i - \underline{b}_i\,.$$

Thus a local optimizer is global if the multipliers y_i are absolutely large compared with the width d_i of the i-th constraint weighted by the convexity correction G_{ii}. Note that multipliers corresponding to equality constraints $d_i = 0$ may be small (or even zero) without affecting the magnitude of α.

2. Inequality (16) can be regarded as a cutting plane which eliminates a part of the feasible region containing the local optimizer.

We end with showing how to obtain a lower bound for the global minimum $\min\{h(\tilde{x}) \mid \tilde{x} \in C\}$ in case that the sufficient condition of Theorem 4 is not satisfied. The idea is to construct an underestimating convex quadratic objective function.

PROPOSITION 6. *Suppose that*

$$\alpha := \max\{a_B^T A_B(\tilde{x} - x) \mid \tilde{x} \in C\} > 1\,. \tag{17}$$

Let $H_0 = H + A_B^T G A_B$ and let h_0 be the minimum of the convex quadratic program

$$\begin{array}{ll} \textit{minimize} & h(x) + (1 - \alpha) g^T \delta + \frac{1}{2} \delta^T H_0 \delta \\ \textit{subject to} & \underline{b} - Ax \leq A\delta \leq \bar{b} - Ax\,. \end{array} \tag{18}$$

Then $h(\tilde{x}) \geq h_0$ for all $\tilde{x} \in C$.

Proof. Clearly, $\delta = \tilde{x} - x$ satisfies the constraints of (18) for every $\tilde{x} \in C$. The inequality $h(\tilde{x}) \geq h(x) + (1 - \alpha) g^T \delta + \frac{1}{2} \delta^T H_0 \delta \geq h_0$ now follows from (15) and the definition of h_0. □

4. Discussion

The sufficient condition for global optimality given above can be used in two ways:

Either as a supplement to an indefinite quadratic programming method to check whether the local optimizer found is global; in this case the test either yields the information (I) "optimizer global" or (II) "optimizer possibly not global".

Or as a method to provide cutting planes which enhance the search for further local optima in the ambiguous case (II). In this case the test can be incorporated into an algorithm which solves (1) by a sequence of local quadratic optimization problems; branch and bound techniques can be used since the method proposed yields also lower bounds on the objective function.

To use Theorem 4 and Proposition 6 (e.g. in a branch and bound algorithm) one has to solve a linear program to find α and, if $\alpha > 1$, a convex quadratic program to find h_0. The values of α and h_0 are useful indicators for the assessment of "how optimal" the local optimizer $\tilde{x}$ is. It may also happen that the optimizer $\tilde{x}$ of (17) or (18) already has $h(\tilde{x}) < h(x)$; in this case $\tilde{x}$ can be used as the starting point for a new local optimization leading to a better local optimum for (1).

By repeating the construction for the new quadratic program we can systematically reduce the size of the feasible region. Together with the methods of Pardalos & Rosen [5], this appears to be a fruitful approach which will be explored in a separate paper.

References

1. Fujii, Y., K. Ichida and M. Ozasa (1986), Maximization of Multivariable Functions Using Interval Analysis, pp. 17–26 in K. Nickel (ed.), *Interval Mathematics 1985*, Springer Lecture Notes Comp. Sci. 212, Berlin.
2. Hansen E. and S. Sengupta (1980), Global Constrained Optimization Using Interval Analysis, pp. 25–47 in K. Nickel (ed.), *Interval Mathematics 1980*, Academic Press, New York.
3. Horst R. and H. Tuy (1990), *Global Optimization*, Springer, Berlin.
4. Murty, K. G. and S. N. Kabadi (1987), Some NP-Complete Problems in Quadratic and Nonlinear Programming, *Math. Programming* **39**, 117–129.
5. Pardalos, P. M. and J. B. Rosen (1987), *Constrained Global Optimization: Algorithms and Applications*, Springer Lecture Notes Comp. Sci. 268, Berlin.
6. Pardalos, P. M. and G. Schnitger (1988), Checking Local Optimality in Constrained Quadratic Programming is NP-Hard, *Oper. Res. Lett.* **7**, 33–35.