

*Short Communication*

Nicholas I. M. Gould · Philippe L. Toint

## A note on the convergence of barrier algorithms to second-order necessary points

Received October 21, 1997 / Revised version received May 20, 1998  
Published online November 24, 1998

**Abstract.** It has long been known that barrier algorithms for constrained optimization can produce a sequence of iterates converging to a critical point satisfying weak second-order necessary optimality conditions, when their inner iterations ensures that second-order necessary conditions hold at each barrier minimizer. We show that, despite this, strong second-order necessary conditions may fail to be attained at the limit, even if the barrier minimizers satisfy second-order sufficient optimality conditions.

**Key words.** constrained optimization – barrier algorithms – second-order points

---

### 1. Introduction

We consider the constrained optimization problem

$$\text{minimize } f(x) \tag{1}$$

subject to

$$c_i(x) \geq 0 \text{ for all } i \in \mathcal{I}, \tag{2}$$

where  $f$  and the  $c_i$  map  $\mathbf{R}^n$  into  $\mathbf{R}$  and  $\mathcal{I}$  is a finite set of indices. We assume that  $f(x)$  and the  $c_i(x)$  are twice continuously differentiable on an open set containing

$$\mathcal{F} = \{x \in \mathbf{R}^n \mid c_i(x) \geq 0 \text{ for all } i \in \mathcal{I}\}.$$

Our principal interest is in identifying nonlinear programming methods which, under reasonable assumptions, are capable of ensuring convergence to points at which second-order necessary optimality conditions are satisfied. When the problem is unconstrained, it is well known that a number of optimization techniques (principally trust-region-, but also linesearch-, based, see Moré [14], Shultz, Schnabel and Byrd [17], McCormick [13], and Moré and Sorensen [15]) are capable of guaranteeing convergence to second-order points.

---

N.I.M. Gould: Rutherford Appleton Laboratory, Chilton, Oxfordshire OX11 0QX, England, e-mail: N.Gould@rl.ac.uk (Current reports available by anonymous ftp from the directory “pub/reports” on joyous-gard.cc.rl.ac.uk)

Ph.L. Toint: Department of Mathematics, Facultés Universitaires ND de la Paix, 61, rue de Bruxelles, B-5000 Namur, Belgium, e-mail: pht@math.fundp.ac.be (Current reports available by anonymous ftp from the directory “pub/reports” on thales.math.fundp.ac.be)

Let  $\ell(x, y)$  be the Lagrangian function

$$\ell(x, y) = f(x) - \sum_{i \in \mathcal{I}} y_i c_i(x). \quad (3)$$

Under suitable constraint qualifications (see Gould and Tolle [11], Mangasarian [12], and the papers quoted therein), it is well known that a (local) solution  $x_*$  of (1)–(2), together with an associated set of Lagrange multipliers  $y_*$ , satisfies the first-order (Karush-Kuhn-Tucker) necessary conditions

$$\nabla_x \ell(x_*, y_*) = 0 \quad (4)$$

$$c_i(x_*) \geq 0 \text{ and } (y_*)_i \geq 0 \text{ for all } i \in \mathcal{I} \quad (5)$$

$$\text{and } c_i(x_*)(y_*)_i = 0 \text{ for all } i \in \mathcal{I}, \quad (6)$$

as well as the *strong* second-order necessary condition

$$s^T \nabla_{xx} \ell(x_*, y) s \geq 0 \text{ for all } s \in \mathcal{N}_+, \quad (7)$$

where

$$\mathcal{N}_+ = \left\{ s \in \mathbf{R}^n \mid \begin{array}{l} s^T \nabla_x c_i(x_*) = 0 \text{ for all } i \in \{j \in \mathcal{A}(x_*) \mid (y_*)_j > 0\} \text{ and} \\ s^T \nabla_x c_i(x_*) \geq 0 \text{ for all } i \in \{j \in \mathcal{A}(x_*) \mid (y_*)_j = 0\} \end{array} \right\}, \quad (8)$$

and

$$\mathcal{A}(x) = \{i \in \mathcal{I} \mid c_i(x) = 0\}$$

is the active set at  $x$ . The second-order necessary conditions given here are those given by Fletcher [8] Section 9.3). The difficulty with these conditions is that they are not expressible in a computationally convenient form. Indeed, even establishing that the strong second-order conditions are satisfied is, in general, an NP-hard problem (see Murty and Kabadi [16], and Vavasis [18]). Significantly weaker conditions are given by, for instance, Fiacco and McCormick [7] Section 2.2) and Gill, Murray and Wright [10] Section 3.4), which are equivalent to requiring that the solution at the constrained minima under consideration is strictly complementary, that is

$$\{i \in \mathcal{A}(x_*) \mid (y_*)_i = 0\} = \emptyset, \quad (9)$$

and thus that

$$\mathcal{N}_+ = \mathcal{N} \stackrel{\text{def}}{=} \left\{ s \in \mathbf{R}^n \mid s^T \nabla_x c_i(x_*) = 0 \text{ for all } i \in \mathcal{A}(x_*) \right\}. \quad (10)$$

While such an assumption is realistic for linear programming,—all linear programs have such solutions (see, Wright [20], page 28), and many interior-point methods find one—it frequently does not hold for nonlinear programs. On the other hand, the advantage of requiring (9) is that the second-order optimality conditions reduce to checking that the Hessian of the Lagrangian is positive (semi-) definite on the manifold (10) rather than in the cone (8). We shall call the requirement that

$$s^T \nabla_{xx} \ell(x_*, y) s \geq 0 \text{ for all } s \in \mathcal{N} \quad (11)$$

a *weak* second-order necessary condition. That (11) is weaker than (7) is clear once one realizes that the weak condition is satisfied by the *maximizer* of the quadratic programming problem

$$\min_{\substack{x \in \mathbf{R}^n \\ x \geq 0}} -\|x\|_2^2,$$

while (4)–(6) and the strong condition are together both necessary and sufficient for local optimality of general quadratic programs (see Contesse [5], and Borwein [4]).

A number of algorithms for solving (1)–(2) have been shown to converge to points at which the weak second-order necessary conditions hold (see, for example, Gay [9], Bannert [2], Bonnans and Launay [3], Facchinei and Lucidi [6], and Vicente [19]). In particular, Auslender [1] has shown that that, if one traces the trajectory of points at which second-order necessary conditions hold for the barrier function—such points may be found by applying trust-region or line-search methods to the unconstrained barrier problems—then the limit point will satisfy the weak second-order conditions for (1)–(2). However, to our knowledge, no algorithm has been shown to converge to a point at which the strong conditions hold. In this paper, we ask the natural question as to whether interior-point (or, specifically, barrier) methods might do so. It is our purpose to show that, in general, the limit of this barrier trajectory may fail to satisfy the strong second-order necessary conditions.

## 2. A simple counter-example

We shall consider the logarithmic barrier function

$$b_0(x, \mu) = f(x) - \mu \sum_{i \in \mathcal{I}} \log c_i(x),$$

and the reciprocal barrier functions

$$b_\alpha(x, \mu) = f(x) + \frac{\mu}{\alpha} \sum_{i \in \mathcal{I}} \frac{1}{(c_i(x))^\alpha}, \quad (12)$$

for  $\alpha > 0$ .<sup>1</sup> These functions depend on the barrier parameter  $\mu > 0$ . In a typical barrier method, (approximate) stationary points of the barrier functions are traced as the barrier parameter is reduced to zero, and, under reasonable assumptions, this leads to convergence to a Karush-Kuhn-Tucker point.

The example we shall exhibit is a bound-constrained quadratic program of the form

$$\min_{\substack{x \in \mathbf{R}^n \\ x \geq 0}} \frac{1}{2} x^T H x, \quad (13)$$

---

<sup>1</sup> The scaling factor  $\alpha$  in (12) is, perhaps, nonstandard, but may easily be assimilated into the barrier parameter. This allows for a uniform treatment of both barrier functions, as  $\alpha = 0$  corresponds to the logarithmic barrier function.

where  $H$  is a symmetric, indefinite  $n \times n$  matrix. For future reference, when (1)–(2) is of the form (13), the first and second derivatives of the barrier functions above are given by

$$\nabla_x b_\alpha(x, \mu) = \nabla_x f(x) - \mu X^{-(\alpha+1)} e \quad (14)$$

and

$$\nabla_{xx} b_\alpha(x, \mu) = \nabla_{xx} f(x) + \mu(\alpha + 1) X^{-(\alpha+2)},$$

for all  $\alpha \geq 0$ , where  $e$  is the vector of all ones and where  $X = \text{diag}(x_1, \dots, x_n)$ . We also note that

$$\nabla_{xx} \ell(x, y) = H \quad (15)$$

because of (3).

We now choose a sequence  $\{\mu_k\}$  of barrier parameters converging to zero and we define  $H$  to be of the form

$$H = I - (\alpha + \frac{3}{2}) \frac{zz^T}{\|z\|_2^2} \quad (16)$$

where  $I$  is the identity matrix and where we have chosen  $z = e - ne_1$ , the vector  $e_1$  being the first vector of the canonical basis. We then verify that

$$z^T e = e^T e - ne_1^T e = n - n = 0, \quad (17)$$

$$z^T e_1 = e^T e_1 - ne_1^T e_1 = 1 - n \quad (18)$$

and

$$\|z\|_2^2 = e^T e + n^2 e_1^T e_1 - 2ne_1^T e = n + n^2 - 2n = n(n - 1). \quad (19)$$

The definition (16) and (17) together imply that

$$He = e. \quad (20)$$

Now let

$$x_k = \mu_k^{\frac{1}{\alpha+2}} e. \quad (21)$$

We then verify that  $x_k$  is a strict local minimizer of the problem

$$\min_{\substack{x \in \mathbb{R}^n \\ x \geq 0}} b_\alpha(x, \mu_k) \quad (22)$$

that satisfies second-order *sufficient* optimality conditions for this problem. Indeed, the first-order optimality condition holds since

$$\nabla_x b_\alpha(x_k, \mu_k) = Hx_k - \mu_k X_k^{-(\alpha+1)} e = \mu_k^{\frac{1}{\alpha+2}} e - \mu_k^{1-\frac{\alpha+1}{\alpha+2}} e = \mu_k^{\frac{1}{\alpha+2}} (e - e) = 0,$$

where we used (14), (20) and (21), and we have also that

$$\nabla_{xx} b_\alpha(x_k, \mu_k) = H + \mu_k(\alpha + 1)X_k^{-(\alpha+2)} = \frac{1}{2}I + (\alpha + \frac{3}{2}) \left( I - \frac{zz^T}{\|z\|_2^2} \right)$$

is obviously positive definite since the first term of the last right-hand side is positive definite and the last term in brackets is an orthogonal projector, which is therefore positive semidefinite. (Note that  $x_k$  is not the global minimizer of (22) for any  $\alpha$ , as ultimately the negative curvature along  $e_1$  becomes dominant for large  $x$ .) As expected,  $\{x_k\}$  converges to zero, a critical point of problem (13). However, using (16), (18) and (19), we find that

$$\begin{aligned} e_1^T \nabla_{xx} \ell(x, y) e_1 &= e_1^T H e_1 = 1 - (\alpha + \frac{3}{2}) \frac{(e_1^T z)^2}{\|z\|_2^2} \\ &= 1 - (\alpha + \frac{3}{2}) \frac{(n-1)^2}{n(n-1)} = \frac{n - (\alpha + \frac{3}{2})(n-1)}{n}, \end{aligned}$$

which is strictly negative for all values of  $n$  satisfying the inequality

$$n > \frac{\alpha + \frac{3}{2}}{\alpha + \frac{1}{2}}.$$

In particular, it is negative for  $n \geq 4$  if the logarithmic barrier function ( $\alpha = 0$ ) is considered. But  $e_1$  belong to  $\mathcal{N}_+ = \{x \in \mathbf{R}^n \mid x \geq 0\}$ , and thus the strong second-order necessary conditions do not hold at the origin.

### 3. Conclusion

We have shown that the strong second-order necessary optimality conditions for inequality constrained problems may not hold at limit points of a sequence of barrier minimizers, even if each of these minimizers satisfies the second-order sufficient conditions for unconstrained minimization. This negative conclusion is valid for a large class of barrier functions, including the popular logarithmic and reciprocal barriers.

This result casts doubts on the possibility of obtaining strong second-order convergence properties for a number of practical interior-point methods for nonlinear programming. However, it also raises the intriguing question of determining if there are barrier functions, outside the class considered here, for which the desired strong second-order convergence properties are satisfied. As active-set methods may also have trouble with our example (because it has zero Lagrange multipliers), the more general question of whether there are effective methods which ensure convergence to strong second-order points remains open.

*Acknowledgements.* The authors are grateful to A. R. Conn for his careful comments on the manuscript and to R. Fletcher and F. Jarre for their interest.

## References

1. Auslender, A. (1979): Penalty methods for computing points that satisfy second order necessary points. *Math. Program.* **17**(2), 229–238
2. Bannert, T. (1994): A trust region algorithm for nonsmooth optimization. *Math. Program.* **67**(2), 247–264
3. Bonnans, J.F., Launay, G. (1995): Sequential quadratic-programming with penalization of the displacement. *SIAM J. Optim.* **5**(4), 792–812
4. Borwein, J.M. (1982): Necessary and sufficient conditions for quadratic minimality. *Numer. Funct. Anal. Optim.* **5**, 127–140
5. Contesse, L.B. (1980): Une caractérisation complète des minima locaux en programmation quadratique. *Numer. Math.* **34**, 315–332
6. Facchinei, F., Lucidi, S. (1996): Convergence to second order stationary points in inequality constrained optimization. Technical Report 32-96. University of Rome “La Sapienza”, Rome
7. Fiacco, A.V., McCormick, G.P. (1990): *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, New York, 1968. Reprinted as Classics in Applied Mathematics 4. SIAM
8. Fletcher, R. (1981): *Practical Methods of Optimization: Constrained Optimization*. Wiley, New York
9. Gay, D.M. (1984): A Trust Region Approach to Linearly Constrained Optimization. In: Griffiths, D.F., ed., *Numerical Analysis: Proceedings Dundee 1983*, pp.72–105, Lect. Notes Math. 1066. Springer, Berlin
10. Gill, P.E., Murray, W., Wright, M.H. (1981): *Practical Optimization*. Academic Press, London, New York
11. Gould, F.J., Tolle, J.W. (1972): Geometry of optimality conditions and constraint qualifications. *Math. Program.* **2**(1), 1–18
12. Mangasarian, O.L. (1979): *Nonlinear Programming*. McGraw-Hill, New York
13. McCormick, G.P. (1977): A modification of Armijo’s step-size rule for negative curvature. *Math. Program.* **13**(1), 111–115
14. Moré, J.J. (1983): Recent Developments in Algorithms and Software for Trust Region Methods. In: Bachem, A., Grötschel, M., Korte, B., eds., *Mathematical Programming, The State of the Art*, pp. 258–287. Springer, Berlin
15. Moré, J.J., Sorensen, D.C. (1979): On the use of directions of negative curvature in a modified Newton method. *Math. Program.* **16**(1), 1–20
16. Murty, K.G., Kabadi, S.N. (1987): Some NP-complete problems in quadratic and nonlinear programming. *Math. Program.* **39**(2), 117–129
17. Shultz, G.A., Schnabel, R.B., Byrd, R.H. (1985): A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties. *SIAM J. Numer. Anal.* **22**(1), 47–67
18. Vavasis, S.A. (1992): *Nonlinear Optimization: Complexity Issues*. Number 8 in International Series of Monographs on Computer Science. Oxford University Press, Oxford, Great Britain
19. Vincente, L. (1995): *Trust-Region Interior-Point Algorithms for a Class of Nonlinear Programming Problems*. PhD thesis. Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA. Report TR96-05
20. Wright, S.J. (1997): *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, USA