A note on the second-order convergence of optimization algorithms using barrier functions

by N. I. M. Gould[1] and Ph.L. Toint[2]

Report 97/15 October 7, 1997

[1] Rutherford Appleton Laboratory,
Chilton, Oxfordshire OX11 0QX, England
Email: N.Gould@rl.ac.uk
Current reports available by anonymous ftp from the directory
"pub/reports" on joyous-gard.cc.rl.ac.uk

[2] Department of Mathematics, Facultés Universitaires ND de la Paix, 61, rue de Bruxelles,
B-5000 Namur, Belgium, EU
Email: pht@math.fundp.ac.be
Current reports available by anonymous ftp from the directory
"pub/reports" on thales.math.fundp.ac.be

A note on the second-order convergence of optimization algorithms using barrier functions

Nicholas I. M. Gould and Philippe L. Toint

October 7, 1997

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.

1 Introduction

We consider the constrained optimization problem

$$\text{minimize } f(x) \tag{1.1}$$

subject to

$$c_i(x) \geq 0 \ \text{ for all } \ i \in \mathcal{I}, \tag{1.2}$$

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, 1983, Shultz, Schnabel and Byrd, 1985, McCormick, 1977, and Moré and Sorensen, 1979) are capable of guaranteeing convergence to second-order points. The difficulty when constraints are present is that the second-order conditions are not expressible in a computationally convenient form. Indeed, even establishing that the conditions are satisfied is, in general, an NP-hard problem (see Murty and Kabadi, 1987, and Vavasis, 1992).

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

$$\ell(x,y) = f(x) - \sum_{i \in \mathcal{I}} y_i c_i(x). \tag{1.3}$$

Under suitable constraint qualifications (see Gould and Tolle, 1972, Mangasarian, 1979, and the papers quoted therein), it is well known that a (local) solution x_* of (1.1)–(1.2), together with an associated set of Lagrange multipliers y_*, satisfies the first-order (Karush-Kuhn-Tucker) necessary conditions

$$\begin{aligned} \nabla_x \ell(x_*, y_*) &= 0 & (1.4)\\ c_i(x_*) \geq 0 \text{ and } (y_*)_i &\geq 0 \text{ for all } i \in \mathcal{I} & (1.5)\\ \text{and } c_i(x_*)(y_*)_i &= 0 \text{ for all } i \in \mathcal{I}, & (1.6) \end{aligned}$$

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

$$s^T \nabla_{xx} \ell(x_*, y) s \geq 0 \text{ for all } s \in \mathcal{N}_+, \qquad (1.7)$$

where

$$\mathcal{N}_+ = \left\{ s \in \mathbf{R}^n \;\middle|\; \begin{array}{ll} s^T \nabla_x c_i(x_*) = 0 & \text{for all } i \in \{j \in \mathcal{A}(x_*) \mid (y_*)_j > 0\} \text{ and} \\ s^T \nabla_x c_i(x_*) \geq 0 & \text{for all } i \in \{j \in \mathcal{A}(x_*) \mid (y_*)_j = 0\} \end{array} \right\}, \qquad (1.8)$$

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 (1981 Section 9.3). Significantly weaker conditions are given by, for instance, Fiacco and McCormick (1968 Section 2.2) and Gill, Murray and Wright (1981 Section 3.4), which are equivalent to requiring that the solution at the constrained minima under consideration is strictly complementary, that is

$$\{ i \in \mathcal{A}(x_*) \mid (y_*)_i = 0 \} \;=\; \emptyset, \qquad (1.9)$$

and thus that

$$\mathcal{N}_+ = \mathcal{N} \stackrel{\text{def}}{=} \left\{ s \in \mathbf{R}^n \;\middle|\; s^T \nabla_x c_i(x_*) = 0 \text{ for all } \; i \in \mathcal{A}(x_*) \right\}. \qquad (1.10)$$

While such an assumption is realistic for linear programming,—all linear programs have such solutions (see, Wright, 1997, 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 (1.9) is that the second-order optimality conditions reduce to checking that the Hessian of the Lagrangian is positive (semi-) definite on the manifold (1.10) rather than in the cone (1.8). We shall call the requirement that

$$s^T \nabla_{xx} \ell(x_*, y) s \geq 0 \text{ for all } s \in \mathcal{N} \qquad (1.11)$$

a *weak* second-order necessary condition. That (1.11) is weaker than (1.7) 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 (1.4)–(1.6) and the strong condition are together both necessary and sufficient for local optimality of general quadratic programs (see Contesse, 1980, and Borwein, 1982).

A number of algorithms for solving (1.1)–(1.2) have been shown to converge to points at which the weak second-order necessary conditions hold (see, for example, Gay, 1984, Bannert, 1994, Bonnans and Launay, 1995, Facchinei and Lucidi, 1996, and Vicente, 1995). In particular, Auslender (1979) 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 (1.1)–(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

$$b_\alpha(x,\mu) = f(x) + \frac{\mu}{\alpha} \sum_{i\in\mathcal{I}} \frac{1}{(c_i(x))^\alpha}, \tag{2.1}$$

for $\alpha > 0$.[(1)] 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

$$\min_{\substack{x\in\mathbf{R}^n \\ x\geq 0}} \tfrac{1}{2} x^T H x, \tag{2.2}$$

where H is a symmetric, indefinite $n \times n$ matrix. For future reference, when (1.1)–(1.2) is of the form (2.2), the first and second derivatives of the barrier functions above are given by

$$\nabla_x b_\alpha(x,\mu) = \nabla_x f(x) - \mu X^{-(\alpha+1)} e \tag{2.3}$$

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 $X = \mathrm{diag}(x_1, \ldots, x_n)$. We also note that

$$\nabla_{xx} \ell(x,y) = H \tag{2.4}$$

because of (1.3).

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

$$H = I - (\alpha + \tfrac{3}{2}) \frac{zz^T}{\|z\|_2^2} \tag{2.5}$$

(1) The scaling factor α in (2.1) is, perhaps, nonstandard, but may easily be assimilated into the barrier parameter. This allows for a uniform treatment of both barrier functions.

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

$$\begin{aligned} z^T e &= e^T e - n e_1^T e = n - n = 0, & (2.6)\\ z^T e_1 &= e^T e_1 - n e_1^T e_1 = 1 - n & (2.7) \end{aligned}$$

and

$$\|z\|_2^2 = e^T e + n^2 e_1^T e_1 - 2n e_1^T e = n + n^2 - 2n = n(n-1). \tag{2.8}$$

The definition (2.5) and (2.6) together imply that

$$He = e. \tag{2.9}$$

Now let

$$x_k = \mu_k^{\frac{1}{\alpha+2}} e. \tag{2.10}$$

We then verify that x_k is a minimizer of the problem

$$\min_{\substack{x \in \mathbf{R}^n \\ x \geq 0}} b_\alpha(x, \mu_k) \tag{2.11}$$

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

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

where we used (2.3), (2.9) and (2.10), and we have also that

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

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. As expected, $\{x_k\}$ converges to zero, the only critical point of problem (2.2). However, using (2.5), (2.7) and (2.8), we find that

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

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

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

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 log- 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 intriging question of determining if there are barrier functions, outside the class considered here, for which the desired strong second-order convergence properties are satisfied. More generally, the question of whether there are effective methods which ensure convergence to stong second-order points remains open.

Acknowledgments

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

References

A. Auslender. Penalty methods for computing points that satisfy second order necessary points. *Mathematical Programming*, **17**(2), 229–238, 1979.

T. Bannert. A trust region algorithm for nonsmooth optimization. *Mathematical Programming*, **67**(2), 247–264, 1994.

J. F. Bonnans and G. Launay. Sequential quadratic-programming with penalization of the displacement. *SIAM Journal on Optimization*, **5**(4), 792–812, 1995.

J. M. Borwein. Necessary and sufficient conditions for quadratic minimality. *Numerical Functional Analysis and Optimization*, **5**, 127–140, 1982.

L. B. Contesse. Une caractérisation complètes des minima locaux en programmation quadratique. *Numerische Mathematik*, **34**, 315–332, 1980.

F. Facchinei and S. Lucidi. Convergence to second order stationary points in inequality constrained optimization. Technical Report 32-96, University of Rome "La Sapienza", Rome, 1996.

A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. J. Wiley and Sons, New York, 1968. Reprinted as *Classics in Applied Mathematics 4*, SIAM, 1990.

R. Fletcher. *Practical Methods of Optimization: Constrained Optimization*. J. Wiley and Sons, New-York, 1981.

D. M. Gay. A trust region approach to linearly constrained optimization. *in* D. F. Griffiths, ed., 'Numerical Analysis: Proceedings Dundee 1983', pp. 72–105, Berlin, 1984. Springer Verlag. Lecture Notes in Mathematics 1066.

P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, London and New York, 1981.

F. J. Gould and J. W. Tolle. Geometry of optimality conditions and constraint qualifications. *Mathematical Programming*, **2**(1), 1–18, 1972.

O. L. Mangasarian. *Nonlinear Programming*. McGraw-Hill, New-York, 1979.

G. P. McCormick. A modification of armijo's step-size rule for negative curvature. *Mathematical Programming*, **13**(1), 111–115, 1977.

J. J. Moré. Recent developments in algorithms and software for trust region methods. *in* A. Bachem, M. Grötschel and B. Korte, eds, 'Mathematical Programming: The State of the Art', pp. 258–287, Berlin, 1983. Springer Verlag.

J. J. Moré and D. C. Sorensen. On the use of directions of negative curvature in a modified Newton method. *Mathematical Programming*, **16**(1), 1–20, 1979.

K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, **39**(2), 117–129, 1987.

G. A. Shultz, R. B. Schnabel, and R. H. Byrd. A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties. *SIAM Journal on Numerical Analysis*, **22**(1), 47–67, 1985.

S. A. Vavasis. *Nonlinear Optimization: Complexity Issues*. Number 8 *in* 'International Series of Monographs on Computer Science'. Oxford University Press, Oxford, Great Britain, 1992.

L. Vicente. *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, 1995. Report TR96-05.

S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, USA, 1997.