DAMTP 2008/NA06

On the convergence of a wide range of trust region methods for unconstrained optimization[1]

M.J.D. Powell

Abstract: We consider trust region methods for seeking the unconstrained minimum of an objective function $F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, when the gradient $\underline{\nabla} F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, is available. The methods are iterative with $\underline{x}_1$ being given. The new vector of variables $\underline{x}_{k+1}$ is derived from a quadratic approximation to F that interpolates $F(\underline{x}_k)$ and $\underline{\nabla} F(\underline{x}_k)$, where k is the iteration number. The second derivative matrix of the quadratic approximation, B_k say, can be indefinite, because the approximation is employed only if the vector of variables $\underline{x}$ satisfies $\|\underline{x} - \underline{x}_k\| \leq \Delta_k$, where Δ_k is a "trust region radius" that is adjusted automatically. Thus the approximation is useful if $\|\underline{\nabla} F(\underline{x}_k)\|$ is sufficiently large and if $\|B_k\|$ and Δ_k are sufficiently small. It is proved under mild assumptions that the condition $\|\underline{\nabla} F(\underline{x}_{k+1})\| \leq \varepsilon$ is achieved after a finite number of iterations, where ε is any given positive constant, and then it is usual to end the calculation. The assumptions include a Lipschitz condition on $\underline{\nabla} F$ and also F has to be bounded below. The termination property is established in a single theorem that applies to a wide range of trust region methods that force the sequence $F(\underline{x}_k)$, $k = 1, 2, 3, \ldots$, to decrease monotonically. Any choice of each symmetric matrix B_k is allowed, provided that $\|B_k\|$ is bounded above by a constant multiple of k.

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge CB3 0WA,
England.

May, 2008.

[1] Dedicated to Ron Mitchell with special thanks for many kindnesses.

1. Introduction

Let the least value of a differentiable function $F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, be required, where $F(\underline{x})$ and the first derivative $\underline{\nabla} F(\underline{x})$ can be calculated for any vector of variables $\underline{x}$. We consider trust region methods for the unconstrained minimization of F that employ quadratic models of the form

$$F(\underline{x}_k + \underline{s}) \approx Q_k(\underline{s}) = F(\underline{x}_k) + \underline{s}^T \underline{\nabla} F(\underline{x}_k) + \tfrac{1}{2} \underline{s}^T B_k \underline{s}, \qquad \underline{s} \in \mathcal{R}^n. \tag{1.1}$$

They are iterative, k being reserved for the iteration number, and $Q_k(\underline{s})$, $\underline{s} \in \mathcal{R}^n$, being an estimate of $F(\underline{x}_k + \underline{s})$ that is useful during the k-th iteration. The vector $\underline{x}_k$ and the $n \times n$ symmetric matrix B_k are available at the start of the iteration, with a positive number Δ_k called the "trust region radius". We assume that the sequence of iterations is terminated if $\|\underline{\nabla} F(\underline{x}_k)\|$ is sufficiently small, which is the condition

$$\|\underline{\nabla} F(\underline{x}_k)\| \leq \varepsilon, \tag{1.2}$$

where ε is a prescribed nonnegative constant. Otherwise, consideration of the current quadratic model provides a trial step $\underline{s}_k \in \mathcal{R}^n$ that has the properties

$$\|\underline{s}_k\| \leq \Delta_k \quad \text{and} \quad Q_k(\underline{s}_k) < F(\underline{x}_k). \tag{1.3}$$

Then the function value $F(\underline{x}_k + \underline{s}_k)$ and the ratio

$$\rho_k = \frac{\text{Actual reduction}}{\text{Predicted reduction}} = \frac{F(\underline{x}_k) - F(\underline{x}_k + \underline{s}_k)}{F(\underline{x}_k) - Q_k(\underline{s}_k)} \tag{1.4}$$

are calculated. Guided by these numbers, the vector $\underline{x}_{k+1}$ and the new trust region radius Δ_{k+1} are selected for the next iteration. The matrix B_{k+1} is chosen too; in several trust region procedures it satisfies the equation

$$B_{k+1} \underline{s}_k = \underline{\nabla} F(\underline{x}_k + \underline{s}_k) - \underline{\nabla} F(\underline{x}_k). \tag{1.5}$$

Many algorithms of this kind are highly successful in practice. Furthermore, nearly all versions of trust region methods in the literature are supported by proofs of convergence. We are also going to engage in theoretical analysis, in order to provide a single convergence theorem that applies to a wide range of trust region methods, the range being specified in Section 2 with reasons for the generality. In particular, the restrictions on the choice of $\underline{x}_{k+1}$ are mild, in order to emphasise that $\underline{x}_{k+1} = \underline{x}_k + \underline{s}_k$ is admissible whenever the strict reduction $F(\underline{x}_k + \underline{s}_k) < F(\underline{x}_k)$ occurs. Thus $F(\underline{x}_k)$ can be the least calculated value of the objective function so far for every iteration number k, which is a strong preference of the author, although most published trust region methods set $\underline{x}_{k+1} = \underline{x}_k + \underline{s}_k$ only if the ratio (1.4) is sufficiently large.

Therefore, assuming exact arithmetic and some usual properties of F, we prove only that the test (1.2) provides termination eventually if the value of ε is positive. In other words, if ε were zero and if the test (1.2) failed for every k, then "lim inf" of the sequence $\|\underline{\nabla} F(\underline{x}_k)\|$, $k=1,2,3,\ldots$, would be zero. On the other hand, by allowing $\underline{x}_{k+1} \neq \underline{x}_k$ only if ρ_k is sufficiently large, several trust region procedures achieve the limit $\|\underline{\nabla} F(\underline{x}_k)\| \rightarrow 0$ as $k \rightarrow \infty$. The required properties of the objective function $F(\underline{x})$ are that it is bounded below and that its first derivative satisfies the Lipschitz condition

$$\|\underline{\nabla} F(\underline{x}) - \underline{\nabla} F(\underline{y})\| \;\leq\; \Lambda \, \|\underline{x} - \underline{y}\|, \qquad \underline{x},\, \underline{y} \in \mathcal{L}, \tag{1.6}$$

where Λ is a positive constant, and where $\mathcal{L} \subset \mathcal{R}^n$ is the level set $\{\underline{x} : F(\underline{x}) \leq F(\underline{x}_1)\}$, the point $\underline{x}_1$ of the first iteration being given.

Our ubiquitous convergence theorem is proved in Section 3, and its relevance to practical computation is discussed in Section 4. Chapter 4 of Nocedal and Wright (1999) is recommended to readers who seek an introduction to implementations and convergence properties of trust region methods. There is much interesting material on developments of the basic methods in the book by Conn, Gould and Toint (2000) and in Chapter 13 of Sun and Yuan (2006).

2. The range of trust region methods

Our conditions on the choices of $\underline{s}_k$, $\underline{x}_{k+1}$, Δ_{k+1} and B_{k+1}, made during the k-th iteration, are stated and explained in this section. The data for the first iteration, namely $\underline{x}_1$, Δ_1 and B_1, have to be supplied by the user of the trust region method with the value of the constant ε for the termination test (1.2). We assume from now on that ε is positive. Every implementation of a trust region method requires the values of several other parameters to be prescribed too, but, instead of being specific, we try to embrace all of them in the conditions below.

The step $\underline{s}_k$ is allowed to be any vector in $\mathcal{R}^n$ that satisfies the inequalities

$$\|\underline{s}_k\| \;\leq\; \Delta_k \quad \text{and} \quad Q_k(\underline{s}_k) \;\leq\; Q_k(\underline{\widehat{s}}_k), \tag{2.1}$$

where $\underline{\widehat{s}}_k$ is the "Cauchy step" of the trust region subproblem of the k-th iteration, which means that $\underline{\widehat{s}}_k$ is the multiple of $\underline{\nabla} F(\underline{x}_k)$ that minimizes $Q_k(\underline{\widehat{s}}_k)$ subject to $\|\underline{\widehat{s}}_k\| \leq \Delta_k$. The conditions (2.1) on $\underline{s}_k$ are usual, because they are suitable for proofs of convergence and because $\underline{\widehat{s}}_k$ is easy to calculate. Each iteration of many trust region algorithms applies its own iterative procedure to generate $\underline{s}_k$, beginning with $\underline{s}_k = \underline{\widehat{s}}_k$, and then adjusting $\underline{s}_k$ in a way that reduces $Q_k(\underline{s}_k)$. Therefore it seems superfluous to replace the second part of expression (2.1) by the weaker condition

$$Q_k(\underline{s}_k) \;\leq\; (1-\theta)\, F(\underline{x}_k) + \theta\, Q_k(\underline{\widehat{s}}_k), \tag{2.2}$$

where θ is any constant from the interval $0 < \theta \leq 1$. This extra generality could have been included in the theory of Section 3, but we assume $\theta = 1$.

We also give up some generality by letting all vector and matrix norms be Euclidean. The use of other norms for the trust region constraint $\|\underline{s}_k\| \leq \Delta_k$ is addressed briefly in Section 4.

The conditions (2.1) provide a well-known upper bound on $Q_k(\underline{s}_k)$ that is important to proofs of convergence. We derive it by letting $\underline{g}_k$ denote $\underline{\nabla} F(\underline{x}_k)$ and by considering the function $\phi(\sigma) = Q_k(-\sigma \underline{g}_k)$, $\sigma \in \mathcal{R}$. The Cauchy step is the vector $\widehat{\underline{s}}_k = -\widehat{\sigma} \underline{g}_k$, where $\widehat{\sigma}$ is the value of σ that minimizes $\phi(\sigma)$ subject to $0 < \sigma \leq \Delta_k / \|\underline{g}_k\|$, so, if $\|\widehat{\underline{s}}_k\| < \Delta_k$ occurs, then $\phi'(\widehat{\sigma})$ is zero. Now the definition (1.1) of Q_k shows that ϕ is quadratic with the first derivative

$$\phi'(\sigma) \;=\; -\|\underline{g}_k\|^2 + \sigma\, \underline{g}_k^T B_k\, \underline{g}_k \;\leq\; (-1 + \sigma\, \|B_k\|)\, \|\underline{g}_k\|^2, \qquad \sigma \in \mathcal{R}. \tag{2.3}$$

Therefore $\|\widehat{\underline{s}}_k\| < \Delta_k$ would imply $\widehat{\sigma} \geq 1/\|B_k\|$, which is the same as $\|\widehat{\underline{s}}_k\| \geq \|\underline{g}_k\| / \|B_k\|$. Thus we deduce the bounds

$$\min\,[\,\Delta_k,\; \|\underline{g}_k\| \,/\, \|B_k\|\,] \;\leq\; \|\widehat{\underline{s}}_k\| \;\leq\; \Delta_k. \tag{2.4}$$

Further, because $\phi(\sigma)$, $\sigma \in \mathcal{R}$, is quadratic and because $\phi'(\widehat{\sigma})$ is nonpositive, the definitions of Q_k, ϕ and $\widehat{\sigma}$ provide the relation

$$\begin{aligned} Q_k(\widehat{\underline{s}}_k) \;&=\; \phi(\widehat{\sigma}) \;=\; F(\underline{x}_k) + \tfrac{1}{2}\, \widehat{\sigma}\, [\, \phi'(0) + \phi'(\widehat{\sigma})\,] \;\leq\; F(\underline{x}_k) + \tfrac{1}{2}\, \widehat{\sigma}\, \phi'(0) \\ &=\; F(\underline{x}_k) - \tfrac{1}{2}\, \widehat{\sigma}\, \|\underline{g}_k\|^2 \;=\; F(\underline{x}_k) - \tfrac{1}{2}\, \|\underline{g}_k\|\, \|\widehat{\underline{s}}_k\|. \end{aligned} \tag{2.5}$$

It follows from expressions (2.1) and (2.4) that $Q_k(\underline{s}_k)$ is bounded above by the inequality

$$Q_k(\underline{s}_k) \;\leq\; Q_k(\widehat{\underline{s}}_k) \;\leq\; F(\underline{x}_k) - \tfrac{1}{2}\, \|\underline{g}_k\|\, \min\,[\,\Delta_k,\; \|\underline{g}_k\| \,/\, \|B_k\|\,], \tag{2.6}$$

which is equivalent to the statement of Lemma 4.5 of Nocedal and Wright (1999).

We now turn to the conditions on $\underline{x}_{k+1}$ and Δ_{k+1}. An important feature of the trust region methods under consideration is that, if the ratio (1.4) is not "sufficiently large", then Δ_{k+1} is always less than Δ_k. In this case, the value of Δ_{k+1} in the analysis of the next section has to be between $\alpha\, \|\widehat{\underline{s}}_k\|$ and $\beta \Delta_k$, where α and β are any prescribed constants that satisfy $0 < \alpha \leq \beta < 1$. The lower bound responds to the possibility that $\Delta_{k+1} = \beta \Delta_k$ may be unsuitable if $\|\widehat{\underline{s}}_k\|$ is much smaller than Δ_k. The term "sufficiently large" means that the inequality

$$\rho_k \;=\; \frac{F(\underline{x}_k) - F(\underline{x}_k + \underline{s}_k)}{F(\underline{x}_k) - Q_k(\underline{s}_k)} \;\geq\; \rho_* \tag{2.7}$$

is achieved, where ρ_* is another prescribed constant from the open interval $(0, 1)$. When $\rho_k \geq \rho_*$ occurs, it is usual for Δ_{k+1} to be at least Δ_k, but the analysis is made a little more general by assuming the weaker condition $\Delta_{k+1} \geq \|\widehat{\underline{s}}_k\|$. Moreover, in all implementations of trust region methods known to the author, Δ_{k+1} is at most Δ_* or $\gamma \Delta_k$ for some prescribed constants $\Delta_* > 0$ and $\gamma > 1$. All of these possibilities are included by allowing Δ_{k+1} to take any value that satisfies the constraints

$$\left.\begin{array}{rcccll} \alpha\, \|\widehat{\underline{s}}_k\| & \leq & \Delta_{k+1} & \leq & \beta\, \Delta_k & \text{if} \quad \rho_k < \rho_* \\ \|\widehat{\underline{s}}_k\| & \leq & \Delta_{k+1} & \leq & \max\,[\,\Delta_*,\; \gamma\, \Delta_k\,] & \text{if} \quad \rho_k \geq \rho_* \end{array}\right\}, \tag{2.8}$$

where the constants α, β, γ, ρ_* and Δ_* have been introduced already. In the "basic algorithm" on page 116 of Conn, Gould and Toint (2000), however, there is no upper bound on Δ_{k+1} in the case $\rho_k \geq \rho_*$.

If $\rho_k \geq \rho_*$ holds, it is important for the conditions on $\underline{x}_{k+1}$ to exclude $\underline{x}_{k+1} = \underline{x}_k$, in order to prevent the possibility of setting both $\underline{x}_{k+1} = \underline{x}_k$ and $\Delta_{k+1} = \Delta_k$ for all sufficiently large k. In this case, we expect every trust region algorithm to pick $\underline{x}_{k+1} = \underline{x}_k + \underline{s}_k$, and we also expect $\underline{x}_{k+1} = \underline{x}_k$ to be chosen whenever ρ_k is nonpositive. Otherwise, when ρ_k is in the open interval $(0, \rho_*)$, we expect $\underline{x}_{k+1}$ to be set to $\underline{x}_k$ or $\underline{x}_k + \underline{s}_k$. Our conditions on $\underline{x}_{k+1}$ admit all of these possibilities and many more. Specifically, $\underline{x}_{k+1}$ is allowed to be any vector in $\mathcal{R}^n$ that has the property

$$\left.\begin{array}{llll} F(\underline{x}_{k+1}) \leq F(\underline{x}_k) & \text{if} & \rho_k < \rho_* \\ F(\underline{x}_{k+1}) \leq F(\underline{x}_k + \underline{s}_k) & \text{if} & \rho_k \geq \rho_* \end{array}\right\}. \tag{2.9}$$

It follows that the proof of the termination theorem in Section 3 is valid without the usual constraint $\|\underline{x}_{k+1} - \underline{x}_k\| \leq \Delta_k$. The conditions (2.9) imply that the sequence $F(\underline{x}_k)$, $k = 1, 2, 3, \ldots$, decreases monotonically, and that $F(\underline{x}_{k+1})$ is substantially less than $F(\underline{x}_k)$ on the iterations that achieve $\rho_k \geq \rho_*$.

Particular attention is going to be given in Section 3 to the accuracy of the approximation (1.1) when $\|\underline{s}\|$ is relatively small and $\|\underline{\nabla} F(\underline{x}_k)\| \geq \varepsilon$ holds. Then the term $\underline{s}^T \underline{\nabla} F(\underline{x}_k)$ is more important than the term $\frac{1}{2} \underline{s}^T B_k \underline{s}$, provided that $\|B_k\|$ is not too large, which suggests the kind of condition that B_{k+1} is required to satisfy for each k. Specifically, it is sufficient if there exist nonnegative numbers λ and μ, independent of k, such that the sequence of matrices B_k has the property

$$\|B_k\| \leq \lambda + \mu k, \qquad k = 1, 2, 3, \ldots . \tag{2.10}$$

This constraint occurs also in the study of trust region methods by Powell (1984), but the conditions there on $\underline{s}_k$, $\underline{x}_{k+1}$ and Δ_{k+1} are much stronger than the requirements of our theoretical analysis.

An advantage of allowing $\|B_k\|$ to grow linearly with k, instead of assuming a constant upper bound on $\|B_k\|$, is that, under an extension of the Lipschitz condition (1.6), each new matrix B_{k+1} can be defined by the "symmetric Broyden" formula. The extension is that the set $\mathcal{L}$ of inequality (1.6) is enlarged if necessary so that it includes the line segment between $\underline{x}_k$ and $\underline{x}_k + \underline{s}_k$ for every iteration number k. Thus the difference $\underline{\nabla} F(\underline{x}_k + \underline{s}_k) - \underline{\nabla} F(\underline{x}_k) = \underline{t}_k$, say, has the property

$$\|\underline{t}_k\| \leq \Lambda \, \|\underline{s}_k\|, \qquad k = 1, 2, 3, \ldots . \tag{2.11}$$

That updating formula defines B_{k+1} to be the symmetric matrix that minimizes the Frobenius norm of $B_{k+1} - B_k$ subject to equation (1.5), which is a quadratic programming problem. It has the solution

$$\begin{aligned} B_{k+1} &= B_k + \frac{(\underline{t}_k - B_k \underline{s}_k) \, \underline{s}_k^T + \underline{s}_k \, (\underline{t}_k - B_k \underline{s}_k)^T}{\|\underline{s}_k\|^2} - (\underline{t}_k - B_k \underline{s}_k)^T \underline{s}_k \, \frac{\underline{s}_k \, \underline{s}_k^T}{\|\underline{s}_k\|^4} \\ &= \left(I - \frac{\underline{s}_k \, \underline{s}_k^T}{\|\underline{s}_k\|^2} \right) B_k \left(I - \frac{\underline{s}_k \, \underline{s}_k^T}{\|\underline{s}_k\|^2} \right) + \frac{\underline{t}_k \, \underline{s}_k^T + \underline{s}_k \, \underline{t}_k^T}{\|\underline{s}_k\|^2} - \underline{t}_k^T \underline{s}_k \, \frac{\underline{s}_k \, \underline{s}_k^T}{\|\underline{s}_k\|^4}. \end{aligned} \tag{2.12}$$

We see that B_k is multiplied by symmetric projection matrices in the last line. Moreover, the Euclidean norms of the matrices $\underline{t}_k \, \underline{s}_k^T$ and $\underline{s}_k \, \underline{s}_k^T$ are $\|\underline{s}_k\| \, \|\underline{t}_k\|$ and $\|\underline{s}_k\|^2$, respectively, and the modulus of the scalar product $\underline{t}_k^T \underline{s}_k$ is bounded above by $\|\underline{s}_k\| \, \|\underline{t}_k\|$. It follows from formula (2.12) and the triangle inequality that B_{k+1} satisfies the condition

$$\|B_{k+1}\| \;\leq\; \|B_k\| + 3 \, \|\underline{s}_k\| \, \|\underline{t}_k\| \, / \, \|\underline{s}_k\|^2 \;\leq\; \|B_k\| + 3\Lambda, \tag{2.13}$$

the last inequality being due to the bound (2.11). Therefore, letting λ and μ be the constants $\|B_1\| - 3\Lambda$ and 3Λ, respectively, the required property (2.10) can be established by induction on k.

There are several other updating formulae for calculating B_{k+1} from B_k and equation (1.5), but in some trust region algorithms for unconstrained minimization the updating is suppressed if $\|B_{k+1} - B_k\|$ would become unacceptably large. In particular, the symmetric rank one formula

$$B_{k+1} \;=\; B_k + \frac{(\underline{t}_k - B_k \, \underline{s}_k) \, (\underline{t}_k - B_k \, \underline{s}_k)^T}{(\underline{t}_k - B_k \, \underline{s}_k)^T \underline{s}_k} \tag{2.14}$$

is useless if its denominator is zero, where we are retaining the notation $\underline{t}_k = \underline{\nabla} F(\underline{x}_k + \underline{s}_k) - \underline{\nabla} F(\underline{x}_k)$. Therefore Conn, Gould and Toint (1991) investigate the technique of applying equation (2.14) on the k-th iteration if and only if the inequality

$$|(\underline{t}_k - B_k \, \underline{s}_k)^T \underline{s}_k| \;\geq\; c_1 \, \|\underline{t}_k - B_k \, \underline{s}_k\| \, \|\underline{s}_k\| \tag{2.15}$$

holds for some small prescribed constant $c_1 > 0$, the alternative for the other iterations being $B_{k+1} = B_k$. Their analysis of this "safeguarded SR1 method", however, assumes a linear independence property of the steps $\underline{s}_k$, $k = 1, 2, 3, \ldots$, that is questionable. Byrd, Khalfan and Schnabel (1996) proved later that, instead of the linear independence property, one can assume that the matrices B_k, $k = 1, 2, 3, \ldots$, stay bounded as $k \to \infty$, but that is questionable too. The theory of the next section makes a little more progress by establishing the termination of the safeguarded SR1 method under the weaker condition (2.10).

3. The proof of termination

It is proved in this section that, under the given conditions on $F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, and on the range of trust region methods, the termination property (1.2) is achieved for some iteration number k. The method of proof depends on the sum of the reductions $F(\underline{x}_k) - F(\underline{x}_{k+1})$ that occur on the iterations that satisfy $\rho_k \geq \rho_*$. This sum is a lower bound on the total reduction in F on all iterations, because the sequence $F(\underline{x}_k)$, $k = 1, 2, 3, \ldots$, decreases monotonically. Therefore the sum is finite if F is bounded below, which is one of our assumptions. We will find that if $k \to \infty$, however, then the sum would diverge. It follows that the sequence of iterations must terminate.

Our analysis begins with two lemmas. The first of them shows that $\rho_k \geq \rho_*$ holds if Δ_k is sufficiently small. Then the second lemma establishes a lower bound on Δ_k for all k, which, for the iterations that satisfy $\rho_k \geq \rho_*$, leads to a lower bound on $F(\underline{x}_k) - F(\underline{x}_{k+1})$ that is independent of Δ_k. Because the proof of termination is by contradiction, we assume that $\|\underline{\nabla} F(\underline{x}_k)\|$ exceeds ε for every k. We recall that ρ_k is the ratio (1.4), that ρ_* is a constant from $(0,1)$, that Λ is the constant of the Lipschitz condition (1.6), and that we may write $\underline{g}_k$ instead of $\underline{\nabla} F(\underline{x}_k)$.

Lemma 1 If the trust region radius has the property

$$\Delta_k \;\leq\; (1-\rho_*)\, \|\underline{g}_k\| \,/\, (\,\Lambda + \|B_k\|\,), \tag{3.1}$$

then the condition $\rho_k \geq \rho_*$ is achieved.

Proof The assumption (3.1) implies that Δ_k is less than $\|\underline{g}_k\|/\|B_k\|$. Therefore inequality (2.6) takes the form

$$Q_k(\underline{s}_k) \;\leq\; F(\underline{x}_k) - \tfrac{1}{2}\, \|\underline{g}_k\|\, \Delta_k, \tag{3.2}$$

any choice of $\underline{s}_k$ being allowed subject to the conditions (2.1).

We consider the function $\psi(\tau) = F(\underline{x}_k + \tau \underline{s}_k)$, $0 \leq \tau \leq \widehat{\tau}$, where $\widehat{\tau}$ is the greatest number in $[0,1]$ that satisfies $\psi(\tau) \leq \psi(0)$, $0 \leq \tau \leq \widehat{\tau}$. The definition (1.1) of Q_k, with $\|\underline{s}_k\| \leq \Delta_k$ and inequality (3.2), give the bound

$$\begin{aligned} \psi'(0) \;=\; \underline{s}_k^T \underline{g}_k \;&=\; Q_k(\underline{s}_k) - F(\underline{x}_k) - \tfrac{1}{2}\, \underline{s}_k^T B_k\, \underline{s}_k \\ &\leq\; \tfrac{1}{2}\, \Delta_k\, (\, -\|\underline{g}_k\| \;+\; \|B_k\|\, \Delta_k\,). \end{aligned} \tag{3.3}$$

It follows from the remark $\Delta_k < \|\underline{g}_k\|/\|B_k\|$ at the beginning of this proof that $\psi'(0)$ is negative. Therefore $\widehat{\tau}$ is nonzero.

The value of $\widehat{\tau}$ allows $\underline{x}$ and $\underline{y}$ in condition (1.6) to be any two points on the line segment between $\underline{x}_k$ and $\underline{x}_k + \widehat{\tau} \underline{s}_k$. Thus some elementary algebra provides the inequality

$$\begin{aligned} |\, \psi(\widehat{\tau}) - \psi(0) - \widehat{\tau}\, \psi'(0)\,| \;&=\; |\, \textstyle\int_0^{\widehat{\tau}} [\, \psi'(\tau) - \psi'(0)\,]\, d\tau\,| \\ &=\; |\, \textstyle\int_0^{\widehat{\tau}} \underline{s}_k^T\, [\, \underline{\nabla} F(\underline{x}_k + \tau\, \underline{s}_k) - \underline{\nabla} F(\underline{x}_k)\,]\, d\tau\,| \;\leq\; \tfrac{1}{2}\, \Lambda\, \widehat{\tau}^2\, \|\underline{s}_k\|^2. \end{aligned} \tag{3.4}$$

It follows that $F(\underline{x}_k + \widehat{\tau} \underline{s}_k) = \psi(\widehat{\tau})$ has the property

$$\begin{aligned} F(\underline{x}_k + \widehat{\tau}\, \underline{s}_k) \;&\leq\; \psi(0) + \widehat{\tau}\, \psi'(0) + \tfrac{1}{2}\, \Lambda\, \widehat{\tau}^2\, \|\underline{s}_k\|^2 \\ &\leq\; F(\underline{x}_k) + \tfrac{1}{2}\, \widehat{\tau}\, \Delta_k\, (\, -\|\underline{g}_k\| + \|B_k\|\, \Delta_k + \Lambda\, \Delta_k\,), \end{aligned} \tag{3.5}$$

where the last line is derived from $F(\underline{x}_k) = \psi(0)$ and from the bounds (3.3) and $\widehat{\tau}^2 \|\underline{s}_k\|^2 \leq \widehat{\tau} \Delta_k^2$. Therefore the restriction (3.1) on Δ_k yields the strict inequality $F(\underline{x}_k + \widehat{\tau} \underline{s}_k) < F(\underline{x}_k)$, which implies $\widehat{\tau} = 1$.

Finally, we consider the ratio (1.4). Expression (3.4) with $\widehat{\tau} = 1$ and $\|\underline{s}_k\| \leq \Delta_k$ provide the condition

$$|\, F(\underline{x}_k + \underline{s}_k) - F(\underline{x}_k) - \underline{s}_k^T \underline{g}_k\,| \;\leq\; \tfrac{1}{2}\, \Lambda\, \Delta_k^2, \tag{3.6}$$

while the first line of expression (3.3) with $\|\underline{s}_k\| \leq \Delta_k$ give the relation

$$|\, Q_k(\underline{s}_k) - F(\underline{x}_k) - \underline{s}_k^T \underline{g}_k \,| \;\leq\; \tfrac{1}{2}\, \|B_k\|\, \Delta_k^2. \tag{3.7}$$

By combining these last two displays, we deduce the inequality

$$|\, F(\underline{x}_k + \underline{s}_k) - Q_k(\underline{s}_k)\,| \;\leq\; \tfrac{1}{2}\, (\Lambda + \|B_k\|\,)\, \Delta_k^2. \tag{3.8}$$

It follows from the property (3.2) that the condition

$$\frac{F(\underline{x}_k + \underline{s}_k) - Q_k(\underline{s}_k)}{F(\underline{x}_k) - Q_k(\underline{s}_k)} \;\leq\; \frac{(\Lambda + \|B_k\|)\, \Delta_k}{\|\underline{g}_k\|} \tag{3.9}$$

is achieved. Further, because of the definition (1.4), the left hand side of expression (3.9) is $1-\rho_k$. Hence, by rearranging terms and by invoking the assumption (3.1) again, we find the bound

$$\rho_k \;\geq\; 1 - (\Lambda + \|B_k\|\,)\, \Delta_k \,/\, \|\underline{g}_k\| \;\geq\; \rho_*, \tag{3.10}$$

which completes the proof. $\bowtie$

Lemma 2 Let ρ_*, ε and μ be taken from expression (2.7), from the termination condition (1.2) and from the bound (2.10), respectively. There exist positive constants δ and M such that the trust region radii have the lower bounds

$$\Delta_k \;\geq\; \delta \,/\, (M + \mu\, k), \qquad k = 1, 2, 3, \ldots, \tag{3.11}$$

and such that, if $\rho_k \geq \rho_*$ holds, then the k-th iteration achieves a reduction in the objective function that satisfies the condition

$$F(\underline{x}_k) - F(\underline{x}_{k+1}) \;\geq\; \tfrac{1}{2}\, \rho_*\, \delta\, \varepsilon \,/\, (M + \mu\, k). \tag{3.12}$$

Proof It will be shown that it is suitable to pick the values

$$\delta \;=\; \min[\,\Delta_1,\, \alpha\,(1-\rho_*)\,\varepsilon\,] \qquad \text{and} \qquad M \;=\; \max[\,1,\, \Lambda+\lambda\,], \tag{3.13}$$

where Δ_1 is the initial trust region radius that is supplied by the user, and where α, Λ and λ are given in expressions (2.8), (1.6) and (2.10), respectively. It follows from $\delta \leq \Delta_1$ and $M \geq 1$ that inequality (3.11) is satisfied in the case $k=1$. We employ induction to establish the inequality for larger values of k.

Let the property (3.11) hold for a general iteration number k. If $\rho < \rho_*$, then conditions (2.8) and (2.4) with Lemma 1 provide the bound

$$\begin{aligned} \Delta_{k+1} \;\geq\; \alpha\, \|\widehat{\underline{s}}_k\| \;&\geq\; \alpha\, \min[\,\Delta_k,\, \|\underline{g}_k\| \,/\, \|B_k\|\,] \\ &\geq\; \alpha\,(1-\rho_*)\, \|\underline{g}_k\| \,/\, (\Lambda + \|B_k\|). \end{aligned} \tag{3.14}$$

Hence the assumptions $\|\underline{g}_k\| \geq \varepsilon$ and $\|B_k\| \leq \lambda + \mu k$ with the values (3.13) give the inequality

$$\Delta_{k+1} \;\geq\; \alpha\,(1-\rho_*)\,\varepsilon\,/\,(\Lambda + \lambda + \mu\,k) \;\geq\; \delta\,/\,(M + \mu\,k). \tag{3.15}$$

Alternatively, if $\rho_k \geq \rho_*$, then conditions (2.8) and (2.4) provide the bound

$$\begin{aligned} \Delta_{k+1} \;&\geq\; \|\underline{\widehat{s}}_k\| \;\geq\; \min\,[\,\Delta_k,\, \|\underline{g}_k\|\,/\,\|B_k\|\,] \\ &\geq\; \min\,[\,\delta\,/\,(M + \mu\,k),\, \varepsilon\,/\,(\lambda + \mu\,k)\,] \;=\; \delta\,/\,(M + \mu\,k), \end{aligned} \tag{3.16}$$

the last line being derived from the inductive hypothesis and the reasons for inequality (3.15). It follows from the last two displays that the assertion (3.11) remains true when k is increased by one, so by induction it is valid for every k.

Let k be the number of any iteration that satisfies $\rho_k \geq \rho_*$. Our proof of the property (3.12) begins with the remark that the bounds (2.6) and (3.11), with $\|\underline{g}_k\| \geq \varepsilon > \delta$ and $\|B_k\| \leq \lambda + \mu k \leq M + \mu k$, provide the condition

$$F(\underline{x}_k) - Q_k(\underline{s}_k) \;\geq\; \tfrac{1}{2}\,\varepsilon\,\min\,[\,\Delta_k,\, \|\underline{g}_k\|\,/\,\|B_k\|\,] \;\geq\; \tfrac{1}{2}\,\delta\,\varepsilon\,/\,(M + \mu\,k). \tag{3.17}$$

Therefore, using the definition (1.4) and $\rho_k \geq \rho_*$, we find the inequality

$$F(\underline{x}_k) - F(\underline{x}_k + \underline{s}_k) \;=\; \rho_k\,[\,F(\underline{x}_k) - Q_k(\underline{s}_k)\,] \;\geq\; \tfrac{1}{2}\,\rho_*\,\delta\,\varepsilon\,/\,(M + \mu\,k). \tag{3.18}$$

We recall from the second line of expression (2.9) that $\underline{x}_{k+1}$ is allowed to be any vector in $\mathcal{R}^n$ that satisfies $F(\underline{x}_{k+1}) \leq F(\underline{x}_k + \underline{s}_k)$. Therefore the required reduction (3.12) in the objective function is achieved. $\bowtie$

Theorem Let the given conditions on $F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, and on the range of trust region methods be satisfied, as specified in Sections 1 and 2. Then the termination property (1.2) is achieved for some iteration number k, where ε is any positive constant.

Proof As mentioned already, we assume that $\|\underline{\nabla} F(\underline{x}_k)\|$ exceeds ε for every k, and we deduce a contradiction. We define the subset $\mathcal{K}$ of the positive integers by putting the iteration number k in $\mathcal{K}$ if and only if $\rho_k \geq \rho_*$ holds. Lemma 2 provides a lower bound on $F(\underline{x}_k) - F(\underline{x}_{k+1})$ for these iteration numbers. The contradiction is that the sum of the terms $\frac{1}{2}\rho_*\delta\varepsilon/(M+\mu k)$, $k \in \mathcal{K}$, is infinite.

The constraints (2.8) include the condition

$$\Delta_{k+1} \;\leq\; \beta\,\Delta_k, \qquad k \notin \mathcal{K}, \tag{3.19}$$

where β is a constant that is strictly less than one. They also include the bound

$$\begin{aligned} \Delta_{k+1} \;&\leq\; \max\,[\,\Delta_*\,/\Delta_k,\, \gamma\,]\,\Delta_k \\ &\leq\; \max\,[\,\Delta_*\,\delta^{\,-1},\, \gamma\,(M + \mu\,k)^{\,-1}\,]\,(M + \mu\,k)\,\Delta_k, \qquad k \in \mathcal{K}, \end{aligned} \tag{3.20}$$

the last part being obtained from the property (3.11). We replace expression (3.20) by the weaker condition

$$\Delta_{k+1} \;\leq\; \eta\,(M+\mu\, k)\,\Delta_k, \qquad k\in\mathcal{K}, \tag{3.21}$$

where η is the constant $\max\,[\Delta_*\,\delta^{-1},\, \gamma\,(M\!+\!\mu)^{-1}]$. Inequalities (3.11), (3.19) and (3.21) imply that $\mathcal{K}$ has an infinite number of elements. We let $q(\ell)$ be the number of elements of $\mathcal{K}$ that are in the interval $[1,\ell]$, where ℓ is any positive integer.

We derive a useful lower bound on $q(\ell)$ from the remark that inequalities (3.11), (3.19) and (3.21) also imply the relation

$$\delta\,/\,(M+\mu+\mu\,\ell) \;\leq\; \Delta_{\ell+1} \;\leq\; \beta^{\,\ell-q(\ell)}\,\{\eta\,(M+\mu\,\ell)\}^{q(\ell)}\,\Delta_1, \tag{3.22}$$

which we write in the form

$$\left(\frac{\eta\,(M\ell^{-1}+\mu)}{\beta}\,\ell\right)^{q(\ell)} \;\geq\; \frac{\delta}{(M+\mu+\mu\,\ell)\,\Delta_1}\,\left(\frac{1}{\beta}\right)^{\ell}. \tag{3.23}$$

Let L be a fixed positive integer that supplies the conditions

$$\left.\begin{array}{rcl} \log\{\eta\,(M\ell^{-1}+\mu)\,/\,\beta\} & \leq & 0.05\,\log\ell \\ \log\{(M+\mu+\mu\,\ell)\,\Delta_1\,/\,\delta\} & \leq & 0.05\,\ell\,\log(1/\beta) \end{array}\right\}, \qquad \ell\geq L. \tag{3.24}$$

By taking logarithms of both sides of expression (3.23), we find the inequality

$$1.05\,q(\ell)\,\log\ell \;\geq\; 0.95\,\ell\,\log(1/\beta), \qquad \ell\geq L, \tag{3.25}$$

which leads to the lower bound

$$q(\ell) \;\geq\; 0.9\,\log(\beta^{-1})\,\ell\,/\,\log\ell, \qquad \ell\geq L. \tag{3.26}$$

We see that $q(\ell)$ diverges as $\ell\to\infty$, which confirms that the number of elements of $\mathcal{K}$ is infinite.

Let $k(q)$ be the q-th element of $\mathcal{K}$ when its elements are arranged in ascending order, let $\mathcal{Q}_j$, $j=1,2,3,\ldots$, be the set of integers q that are in the interval

$$\frac{0.9\,\log(\beta^{-1})\,2^j}{j\,\log 2} \;\leq\; q \;<\; \frac{0.9\,\log(\beta^{-1})\,2^{j+1}}{(j\!+\!1)\,\log 2}, \tag{3.27}$$

and let J be a fixed positive integer that supplies the properties

$$2^{j+1} \;\geq\; L \qquad \text{and} \qquad |\mathcal{Q}_j| \;\geq\; 2, \qquad j\geq J, \tag{3.28}$$

where L is introduced in the previous paragraph and where $|\mathcal{Q}_j|$ is the number of elements in $\mathcal{Q}_j$. We compare the bound (3.26) with the right hand side of expression (3.27) in the cases (3.28). Condition (3.26) states that the number of elements of $\mathcal{K}$ in the first 2^{j+1} iterations is at least $0.9\log(\beta^{-1})\,2^{j+1}/\{(j\!+\!1)\log 2\}$.

Therefore, when q is in the interval (3.27), then the q-th element of $\mathcal{K}$, namely $k(q)$, satisfies $k(q) < 2^{j+1}$. Thus Lemma 2 with $k(q) \in \mathcal{K}$ give the inequality

$$F(\underline{x}_{k(q)}) - F(\underline{x}_{k(q)+1}) \;>\; \tfrac{1}{2}\,\rho_*\,\delta\,\varepsilon\,/\,(M + \mu\,2^{j+1}). \qquad (3.29)$$

Hence, because the sets $\mathcal{Q}_j$, $j \geq J$, are disjoint, because the positive integers $k(q)$, $q = 1, 2, 3, \ldots$, are all different, and because the constraints (2.9) maintain $F(\underline{x}_{k+1}) \leq F(\underline{x}_k)$ on every iteration, we deduce the condition

$$\begin{aligned}\textstyle\sum_{k=1}^{\infty} \{\, F(\underline{x}_k) - F(\underline{x}_{k+1}) \,\} \;&\geq\; \textstyle\sum_{j=J}^{\infty} \sum_{q\in\mathcal{Q}_j} \{\, F(\underline{x}_{k(q)}) - F(\underline{x}_{k(q)+1}) \,\} \\ &\geq\; \tfrac{1}{2}\,\rho_*\,\delta\,\varepsilon \textstyle\sum_{j=J}^{\infty} |\mathcal{Q}_j| \,/\, (M + \mu\,2^{j+1}). \qquad (3.30)\end{aligned}$$

We derive a lower bound on $|\mathcal{Q}_j|$ from the length of the interval (3.27), but, because the end-points of the interval are not integers in general, the difference between the end-points may exceed $|\mathcal{Q}_j|$ by a perturbation that is less than one. The perturbation is also less than $0.5\,|\mathcal{Q}_j|$, due to the constraint $|\mathcal{Q}_j| \geq 2$, $j \geq J$, which admits the inequality

$$\begin{aligned}1.5\,|\mathcal{Q}_j| \;&>\; 0.9\,\log(\beta^{-1})\,2^{j+1} \,/\, \{(j+1)\,\log 2\} \;-\; 0.9\,\log(\beta^{-1})\,2^{j} \,/\, \{j\,\log 2\} \\ &=\; 0.9\,\log(\beta^{-1})\,2^{j}\,(j-1) \,/\, \{j\,(j+1)\,\log 2\}, \qquad j \geq J. \qquad (3.31)\end{aligned}$$

Hence, because J is at least 2, due to $\mathcal{Q}_1$ being empty, we find the property

$$|\mathcal{Q}_j| \;>\; 0.2\,\log(\beta^{-1})\,2^{j} \,/\, (j\,\log 2), \qquad j \geq J. \qquad (3.32)$$

Let $\widehat{\varepsilon}$ be the positive constant $\frac{1}{10}\rho_*\delta\varepsilon\log(\beta^{-1})/\log 2$. Conditions (3.30) and (3.32) imply the bound

$$\begin{aligned}\textstyle\sum_{k=1}^{\infty} \{\, F(\underline{x}_k) - F(\underline{x}_{k+1}) \,\} \qquad & \widehat{\varepsilon} \textstyle\sum_{j=J}^{\infty} 2^{j} \,/\, \{\, (M + 2^{j+1}\,\mu)\, j \,\} \\ >\;& \widehat{\varepsilon} \textstyle\sum_{j=J}^{\infty} 1 \,/\, \{\, (M + 2\,\mu)\, j \,\}, \qquad (3.33)\end{aligned}$$

the last line being elementary. We see that the sum (3.33) is divergent, in contradiction to the assumption that F is bounded below. Therefore the theorem is true. $\bowtie$

4. Discussion

Let a trust region method that satisfies the conditions of Section 2 be applied to an objective function $F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, that satisfies the conditions of the penultimate paragraph of Section 1, let all arithmetic be exact, and let the parameter ε of the test (1.2) for termination be set to zero. It follows from the theorem of Section 3 that, if termination does not occur, then an infinite subsequence of the gradient norms $\|\underline{\nabla} F(\underline{x}_k)\|$, $k = 1, 2, 3, \ldots$, converges to zero. It is suggested in Section 1, however, that sometimes the limit $\|\underline{\nabla} F(\underline{x}_k)\| \to 0$ as $k \to \infty$ is not achieved. This possibility is demonstrated by the following example.

There are only two variables, and we pick the objective function

$$F(\underline{x}) \;=\; F(x,y) \;=\; u(x)+v(y), \qquad \underline{x}\in\mathcal{R}^2, \tag{4.1}$$

where x and y are the components of $\underline{x}$, and where u and v are the functions

$$\left.\begin{array}{rcl} u(x) &=& x^2/\,2, \qquad x\in\mathcal{R}, \qquad \text{and} \\ v(y) &=& y^2\,(y+5)\,(y-3)\,/\,\{\,4\,(y-1)^2+56\,\}, \qquad y\in\mathcal{R}. \end{array}\right\} \tag{4.2}$$

This choice of v is made because it has the properties

$$v(-5)=v(0)=v(3)=0, \quad v'(-5)=-1, \quad v'(0)=0 \quad \text{and} \quad v'(3)=1. \tag{4.3}$$

We will find that the vectors $\underline{x}_k$, $k=1,2,3,\ldots$, can have the components $x_k=2^{-k}$ and $y_k=0$, 3, or -5 in the cases $k=3j+1$, $3j+2$, or $3j+3$, respectively, for all nonnegative integers j. Thus $u(x_k)=2^{-2k-1}$ and $v(y_k)=0$ hold for every iteration number k, and the objective function takes the values $F(\underline{x}_k)=2^{-2k-1}$, $k=1,2,3,\ldots$, which decrease strictly montonically. Further, the subsequence $\|\underline{\nabla} F(\underline{x}_{3j+1})\|$, $j=0,1,2,\ldots$, of gradient norms tends to zero in agreement with the theorem, but $\|\underline{\nabla} F(\underline{x}_{3j+2})\|$ and $\|\underline{\nabla} F(\underline{x}_{3j+3}\|$ tend to one as $j\rightarrow\infty$.

The second derivative matrix B_k of the quadratic model (1.1) is allowed to be any symmetric matrix that has the property (2.10). Therefore we may employ the models

$$Q_k(\underline{s}) \;=\; \left\{\begin{array}{ll} F(\underline{x}_k)+\xi\,x_k+\xi^2 & \text{if} \;\; y_k=0, \\ F(\underline{x}_k)+\xi\,x_k+\xi^2+\eta+\eta^2/\,16 & \text{if} \;\; y_k=3, \\ F(\underline{x}_k)+\xi\,x_k+\xi^2-\eta+\eta^2/\,10 & \text{if} \;\; y_k=-5, \end{array}\right. \tag{4.4}$$

which satisfy $Q_k(0)=F(\underline{x}_k)$ and $\underline{\nabla} Q_k(0)=\underline{\nabla} F(\underline{x}_k)$ as required, ξ and η being the components of $\underline{s}$. Assuming that the trust region radii Δ_k are sufficiently large, which is addressed in the next paragraph, the conditions (2.1) are maintained by choosing each $\underline{s}_k$ so that $Q_k(\underline{s}_k)$ is the least value of $Q_k(\underline{s})$, $\underline{s}\in\mathcal{R}^2$. This construction gives $\underline{s}_k$ the components $\xi_k=-\frac{1}{2}x_k$ and $\eta_k=-8$ or $\eta_k=5$ in the cases $y_k=3$ or $y_k=-5$, respectively, and we keep the example going by selecting $\eta_k=3$ whenever y_k is zero. It follows that every iteration provides the strict reduction

$$F(\underline{x}_k+\underline{s}_k) \;=\; F(\underline{x}_k)-\tfrac{3}{8}\,x_k^2 \;<\; F(\underline{x}_k) \tag{4.5}$$

in the objective function. Therefore the "any decrease" implementation of the conditions (2.9) makes the choice $\underline{x}_{k+1}=\underline{x}_k+\underline{s}_k$ for each k, regardless of the sufficient decrease parameter ρ_*. Thus the sequence $\underline{x}_k$, $k=1,2,3,\ldots$, of the previous paragraph is generated, after letting $\underline{x}_1$ have the components $x_1=0.5$ and $y_1=0$.

The given choices of $\underline{s}_k$ are permitted if the trust region radius is $\Delta_k=4$, $\Delta_k=10$ and $\Delta_k=6$ in the cases $y_k=0$, $y_k=3$ and $y_k=-5$, respectively. We find by easy calculations that the ratio (1.4) takes the values

$$\rho_k \;=\; \frac{F(\underline{x}_k) - F(\underline{x}_k + \underline{s}_k)}{F(\underline{x}_k) - Q_k(\underline{s}_k)} \;=\; \begin{cases} \frac{3}{8}\, x_k^2 \,/\, (\,\frac{1}{4}\, x_k^2) & \text{if} \quad y_k = 0, \\ \frac{3}{8}\, x_k^2 \,/\, (\,4 + \frac{1}{4}\, x_k^2) & \text{if} \quad y_k = 3, \\ \frac{3}{8}\, x_k^2 \,/\, (\,\frac{5}{2} + \frac{1}{4}\, x_k^2) & \text{if} \quad y_k = -5. \end{cases} \tag{4.6}$$

Therefore, letting ρ_* be at least 0.1 as usual and recalling $x_k = 2^{-k}$, the second or first line of expression (2.8) applies when y_k is zero or nonzero, respectively. Hence, by setting $\alpha = 1/2$, $\beta = 2/3$, $\gamma = 5/2$ and $\Delta_1 = 4$, the new trust region radii $\Delta_{k+1} = 10$, $\Delta_{k+1} = 6$ and $\Delta_{k+1} = 4$ become admissible as required in the cases $y_k = 0$, $y_k = 3$ and $y_k = -5$. The description of the example is complete.

Many objective functions $F(\underline{x})$, $\underline{x} \in \mathcal{R}^n$, of unconstrained calculations have unbounded second derivatives, sometimes as $\|\underline{x}\| \rightarrow \infty$, and especially when inequality constraints on the variables are handled by the use of barrier functions. It is important, therefore, that the Lipschitz condition (1.6) has to hold only on the level set $\mathcal{L} = \{\underline{x} : F(\underline{x}) \leq F(\underline{x}_1)\}$. In some other proofs of convergence, however, the Lipschitz condition is required on a larger region of $\mathcal{R}^n$, such as the convex hull of $\mathcal{L}$. For example, see the discussion of Assumption AF3 on page 121 of Conn, Gould and Toint (2000). Our less restrictive condition (1.6) adds some interesting refinements to the proof of Lemma 1 in Section 3. Further, it brings the advantage of allowing F to include an infinite barrier function of a feasible region that is not convex.

It is mentioned after expression (2.8) that Conn, Gould and Toint (2000) do not require an upper bound on Δ_{k+1} in the cases $\rho_k \geq \rho_*$, but their convergence theory depends on the assumption that the matrices B_k, $k = 1, 2, 3, \ldots$, are uniformly bounded, while our condition (2.10) allows $\|B_k\|$ to grow linearly as $k \rightarrow \infty$. We compare these approaches by considering the differences $k(q{+}1) - k(q)$, where $k(q)$ is still the q-th element of the set $\mathcal{K} = \{k : \rho_k \geq \rho_*\}$, this notation being taken from the beginning of the sentence that includes the interval (3.27). If $\Delta_{k(q)+1}$ can be arbitrarily large, and if the first line of expression (2.8) is satisfied by setting $\Delta_{k+1} = \beta\, \Delta_k$ in the cases $\rho_k < \rho_*$, then each $k(q{+}1) - k(q)$ can be arbitrarily large too. Thus condition (2.10) does not restrict the growth of the sequence $\|B_{k(q)}\|$, $q = 1, 2, 3, \ldots$. It follows that the steplengths $\|\underline{x}_{k(q)+1} - \underline{x}_{k(q)}\|$, $q = 1, 2, 3, \ldots$, can tend to zero so rapidly that their sum is finite, even if the constant ε is positive and $\|\underline{g}_k\| \geq \varepsilon$ holds on every iteration. Moreover, the constraints (2.9) allow the choice $\underline{x}_{k+1} = \underline{x}_k$ in all the cases $\rho_k < \rho_*$. Thus it is possible for the sequence $\underline{x}_k$, $k = 1, 2, 3, \ldots$, to converge when $\|\underline{g}_k\|$ is bounded away from zero, which would be a failure of the trust region method. Therefore the upper bound on Δ_{k+1} in the second line of expression (2.8) is important to our analysis.

Another departure from the "basic algorithm" on page 116 of Conn, Gould and Toint (2000) occurs when the parameters η_1 and η_2 of their expression (6.1.5) satisfy $\eta_1 < \eta_2 < 1$. Then, if the ratio (1.4) has the property $\eta_1 \leq \rho_k < \eta_2$, they allow the new trust region radius Δ_{k+1} to take any value from the interval $[\gamma_2 \Delta_k, \Delta_k]$, where γ_2 is another positive parameter that is less than one. On the other hand, when $\|\widehat{\underline{s}}_k\| = \Delta_k$ and $\Delta_{k+1} < \Delta_k$ occur in our range of trust region methods, then our conditions (2.8) do not allow the relative change in the trust region radius to be arbitrarily small, which agrees with most implementations. When comparing our theory with a particular implementation, we pick the ρ_* in expression (2.8) that is most suitable for the choice of Δ_{k+1}. Then, if a sufficient decrease condition is required, the freedom (2.9) in $\underline{x}_{k+1}$ allows the formula

$$\underline{x}_{k+1} = \begin{cases} \underline{x}_k & \text{if} \quad \rho_k < \widehat{\rho}_* \\ \underline{x}_k + \underline{s}_k & \text{if} \quad \rho_k \geq \widehat{\rho}_*, \end{cases} \tag{4.7}$$

where $\widehat{\rho}_*$ is any constant that satisfies $0 < \widehat{\rho}_* \leq \rho_*$.

I am grateful to Ya-xiang Yuan for studying a draft of this paper, partly because he raised the interesting possibility of replacing $\widehat{\underline{s}}_k$ by $\underline{s}_k$ on the left hand side of both lines of expression (2.8). Thus the steplength that contributes to the constraints on Δ_{k+1} would become the steplength that is most relevant to the definition (2.7) of ρ_k. On the other hand, we want the theorem of Section 3 to apply to a range of trust region methods without unnecessary restrictions. Hence the perfect response to Yuan's suggestion would be to find a proof of the theorem after replacing the constraints (2.8) on Δ_{k+1} by the weaker conditions

$$\left.\begin{array}{lllll} \alpha \min[\,\|\widehat{\underline{s}}_k\|,\, \|\underline{s}_k\|\,] \leq \Delta_{k+1} \leq & \beta\,\Delta_k & \text{if} & \rho_k < \rho_* \\ \min[\,\|\widehat{\underline{s}}_k\|,\, \|\underline{s}_k\|\,] \leq \Delta_{k+1} \leq & \max[\,\Delta_*,\, \gamma\,\Delta_k\,] & \text{if} & \rho_k \geq \rho_* \end{array}\right\}. \tag{4.8}$$

The inequalities (2.1) allow $\underline{s}_k$ to be the shortest vector that satisfies $Q_k(\underline{s}_k) \leq Q_k(\widehat{\underline{s}}_k)$. By considering the KKT conditions of this extreme choice, we find that $\underline{s}_k = \widehat{\underline{s}}_k$ occurs only if $\widehat{\underline{s}}_k$, which is a multiple of $\underline{g}_k$, is also a multiple of $B_k \underline{g}_k$, the alternative being $\|\underline{s}_k\| < \|\widehat{\underline{s}}_k\|$. Therefore we expect the new conditions (4.8) to permit a strict reduction in the trust region radius on every iteration, which may provide some challenging analysis. In practice, however, $\underline{s}_k$ is calculated usually by minimizing the model $Q_k(\underline{s})$ subject to $\|\underline{s}\| \leq \Delta_k$, either to high accuracy or by the truncated conjugate gradient algorithm from the starting point $\underline{s} = 0$. Both of these methods give the property $\|\underline{s}_k\| \geq \|\widehat{\underline{s}}_k\|$ automatically, and then the bounds (2.8) and (4.8) on Δ_{k+1} are the same.

The trust regions of some algorithms have the form $\{\underline{s} : \|\underline{s}\|_{TR} \leq \Delta_k\}$, where the vector norm $\|\cdot\|_{TR}$ is not Euclidean. For example, advantage may be taken of the remark that, if the infinity norm is preferred, then the minimization of $Q_k(\underline{s})$ within the trust region is a quadratic programming problem, even if there are linear constraints on the variables. We address this situation briefly by retaining Euclidean vector norms and by generalizing the condition $\|\underline{s}\| \leq \Delta_k$ for each iteration number k to $\underline{s} \in \mathcal{N}_k$, where $\mathcal{N}_k$ is any subset of $\mathcal{R}^n$ that has the properties

$$\|\underline{s}\| \leq \omega\,\Delta_k \;\Rightarrow\; \underline{s} \in \mathcal{N}_k \qquad \text{and} \qquad \underline{s} \in \mathcal{N}_k \;\Rightarrow\; \|\underline{s}\| \leq \Omega\,\Delta_k, \tag{4.9}$$

ω and Ω being constants. By introducing scaling if necessary, we assume the value $\omega = 1$ without loss of generality. Thus the old trust region $\{\underline{s} : \|\underline{s}\| \leq \Delta_k\}$ is a subset of $\mathcal{N}_k$, implying that the important property (2.6) remains valid when $\widehat{\underline{s}}_k$ is changed to the Cauchy step of $\mathcal{N}_k$, but the right hand side of inequality (2.4) becomes $\Omega\Delta_k$. It follows that the right hand sides of expressions (3.6)–(3.9) are all multiplied by Ω^2, so, in order to establish that $\rho_k \geq \rho_*$ is achieved for sufficiently small Δ_k, we replace assumption (3.1) in the statement of Lemma 1 by the more restrictive condition

$$\Delta_k \;\leq\; (1-\rho_*)\, \|\underline{g}_k\|\, \Omega^{-2} \,/\, (\,\Lambda + \|B_k\|). \tag{4.10}$$

Similar factors in the proof of Lemma 2 are treated by changing the definition (3.13) of δ to $\min[\Delta_1,\, \alpha(1-\rho_*)\,\Omega^{-2}\,\varepsilon]$. These modifications seem to be sufficient to prove not only the lemmas but also the theorem of Section 3, after replacing $\|\underline{s}_k\| \leq \Delta_k$ by $\underline{s}_k \in \mathcal{N}_k$ in the conditions (2.1) on $\underline{s}_k$.

One can try to estimate from the theory of Section 3 a bound on the number of iterations before the termination condition (1.2) holds. We find that all iterations with numbers k in the interval $2^j \leq k < 2^{j+1}$ are guaranteed to provide a total reduction in F that is only of magnitude ε^2/j, which is sufficient to prove the theorem. Let ε and every $\|\underline{g}_k\|$ be of magnitude one, and let it be necessary for the reduction in F to be about 10 in order to achieve termination. Then the largest value of j that is relevant, j_* say, may have to satisfy $\sum_{j=1}^{j_*} j^{-1} = 10$, which implies $j_* \approx e^{10} > 22000$, giving the possibility that 2^{22000} iterations may be required. Although this number is monstrous, our proof of convergence may be welcome, because each B_k can be the least favourable matrix that satisfies the bound (2.10), each $\underline{x}_k$ can be any vector in $\mathcal{R}^n$ that is permitted by the conditions (2.9), and the constraints (2.8) allow much freedom in the adjustment of the trust region radius. One benefit of our work is that there is now no need to search for a counter-example if one wishes to know whether or not the test (1.2) for termination is satisfied eventually. Furthermore, if the convergence of a single efficient trust region method is investigated, then it is likely that our theorem will make a useful start by providing the property that $\|\underline{\nabla} F(\underline{x}_k)\|$ does not stay bounded away from zero as $k \rightarrow \infty$.

References

R.H. Byrd, H.F. Khalfan and R.B. Schnabel (1996), “Analysis of a symmetric rank-one trust region method”, *SIAM J. Optimization*, Vol. 6, pp. 1025–1039.

A.R. Conn, N.I.M. Gould and Ph.L. Toint (1991), “Convergence of quasi-Newton matrices generated by the symmetric rank one update”, *Math. Programming*, Vol. 50, pp. 177–195.

A.R. Conn, N.I.M. Gould and Ph.L. Toint (2000), *Trust-Region Methods*, MPS/SIAM Series on Optimization, SIAM (Philadelphia).

J. Nocedal and S.J. Wright (1999), *Numerical Optimization*, Springer (New York).

M.J.D. Powell (1984), "On the global convergence of trust region algorithms for unconstrained minimization", *Math. Programming*, Vol. 29, pp. 297–303.

W. Sun and Y. Yuan (2006), *Optimization Theory and Methods: Nonlinear Programming*, Springer (New York).