The Conjugate Gradient Method and Trust Regions in Large Scale Optimization

Author(s): Trond Steihaug

THE CONJUGATE GRADIENT METHOD AND TRUST REGIONS IN LARGE SCALE OPTIMIZATION*

TROND STEIHAUG†

Abstract. Algorithms based on trust regions have been shown to be robust methods for unconstrained optimization problems. All existing methods, either based on the dogleg strategy or Hebden–Moré iterations, require solution of system of linear equations. In large scale optimization this may be prohibitively expensive. It is shown in this paper that an approximate solution of the trust region problem may be found by the preconditioned conjugate gradient method. This may be regarded as a generalized dogleg technique where we asymptotically take the inexact quasi-Newton step. We also show that we have the same convergence properties as existing methods based on the dogleg strategy using an approximate Hessian.

Key words. unconstrained optimization, locally constrained steps, negative curvature

1. Introduction. The unconstrained minimization of a smooth function in many variables is an important problem in mathematical programming. These problems are usually referred to as large scale unconstrained optimization problems and they occur frequently, for example, in structural design and in finite element methods for nonlinear partial differential equations.

Since the function is smooth, the local minima occur at stationary points, i.e., zeros of the gradient. Effective algorithms are usually based on Newton's method or some variation like the quasi-Newton methods for finding a zero of the gradient. To enlarge the region of convergence, the methods need to be modified. There are two main approaches to achieve global convergence. The most familiar modification is the line search approach where a direction p is computed followed by a search for an approximate local minimizer along the line defined by the direction p. An alternate, appealing approach is based on the observation that quasi-Newton methods model the function by a quadratic approximation around the current iterate. The quadratic is accurate only in a neighborhood of the current iterate and the new iterate is now chosen to be an approximate minimizer of the quadratic constrained to be in the region where we trust the approximation.

Let

$$(1.1) \qquad \varphi(p) = g(x)^T p + \tfrac{1}{2} p^T B p,$$

where B is a real symmetric $n \times n$ matrix and g is the gradient of a smooth function $f: \mathbb{R}^n \to \mathbb{R}$, $g(x) \neq 0$.

The trust region problem is to find $p(\delta)$ so that

$$(1.2) \qquad \varphi(p(\delta)) = \min \{\varphi(p) : \|p\|_C \leq \delta\},$$

where

$$(1.3) \qquad \|p\|_C^2 = p^T C p$$

and C is a symmetric and positive definite matrix. We note that a solution of (1.2) always exists, but the solution is not necessarily unique.

Assume that (1.2) has a unique solution for $0 < \delta \leqq \Delta$, and let $p(\delta)$ be the solution curve.

There are two different approaches to solve (1.2). In the first case we find $p(\tilde{\delta})$ so that

$$\frac{|\tilde{\delta} - \delta|}{\delta} \leqq \varepsilon.$$

The step $p(\tilde{\delta})$ may be found by an iterative method. It can be shown that the optimal solution p^* satisfies

(1.4) $$(B + \alpha C)p^* = -g$$

for some $\alpha \geqq 0$, where $B + \alpha C$ is symmetric and positive semidefinite [6]. Hebden [7] and Moré [9] have given iterative methods for (1.2) based on using different values of α in (1.4). However, in large scale optimization solving the linear system (1.4) for different values of α may be expensive. The problem of solving (1.4) for different values of α is a well-known problem and occurs frequently in the solution of stiff ordinary differential equations and is often referred to as the pencil problem [5].

The other approach approximates the curve $p(\delta)$ by a piecewise linear approximation $\tilde{p}(\tau)$. This is usually referred to as a dogleg [10], [11] or double dogleg [3] technique depending on the number of points defining the approximate curve. The approximate curve $\tilde{p}(\tau)$ is so that $\varphi(\tilde{p}(\tau))$ is monotonically decreasing and $\|\tilde{p}(\tau)\|_C$ is monotonically increasing for $0 \leqq \tau \leqq T$. An approximate solution to the trust region problem (1.2) is found by solving

(1.5) $$\min \{\varphi(\tilde{p}(\tau)) : \|\tilde{p}(\tau)\|_C \leqq \delta, 0 \leqq \tau \leqq T\}.$$

If $\varphi(\tilde{p}(\tau))$ is strictly decreasing and $\|\tilde{p}(\tau)\|_C$ is strictly increasing then for $\|\tilde{p}(T)\|_C > \delta$ the unique solution of (1.5) is found by solving

(1.6) $$\|\tilde{p}(\tau)\|_C = \delta.$$

If B is a nonsingular matrix, then the dogleg curves use the Cauchy point

$$p^C = -\frac{g^T g}{g^T B g} g$$

and the quasi-Newton point

$$p^N = -B^{-1}g.$$

Powell's [11] dogleg curve is

$$\tilde{p}(\tau) = \begin{cases} \tau p^C, & 0 \leqq \tau \leqq 1, \\ p^C + (\tau - 1)\sigma(p^N - p^C), & 1 < \tau \leqq 2, \end{cases}$$

where $\sigma = \text{sign}\,(-g^T(p^N - p^C))$. If B is positive definite, then the double dogleg curve [3] introduces the extra point γp^N, where $0 < \gamma \leqq 1$, and gives an earlier bias to the quasi-Newton step than the dogleg strategy. The double dogleg curve is

$$\tilde{p}(\tau) = \begin{cases} \tau p^C, & 0 \leqq \tau \leqq 1, \\ p^C + \dfrac{\tau - 1}{\gamma}(\gamma p^N - p^C), & 1 < \tau \leqq 1 + \gamma, \\ (\tau - 1)p^N, & 1 + \gamma < \tau \leqq 2. \end{cases}$$

The approximate solution of (1.2) is now either $\tilde{p}(\tau)$ so that $\|\tilde{p}(\tau)\|_C = \delta$ or $\tilde{p}(2)$ if $\|\tilde{p}(2)\|_C \leqq \delta$. The dogleg curves can be defined when B is nonsingular, but it is not clear how to define the dogleg curves when the matrix B is singular. The dogleg curves require the quasi-Newton step and are thus not suited for large scale optimization. Another major weakness of the original dogleg strategy is that the curve $\tilde{p}(\tau)$ is independent of the norm used to define the trust region. However, the solution of (1.6) depends on the norm.

In this paper we show that the preconditioned conjugate gradient (PCG) method (see for instance [1], [2]) applied to (1.1) can be used to generate a piecewise linear curve that approximates $p(\delta)$. The method is well defined for arbitrary B and the approximate curve depends on the norm chosen. The PCG method is well suited for large scale optimization.

In § 2, we describe the algorithm and show that $\|\tilde{p}(\tau)\|_C$ is strictly increasing and $\varphi(\tilde{p}(\tau))$ is monotonically decreasing. Thus (1.5) has a unique solution.

In § 3, we summarize the convergence properties when the algorithm is imbedded in the class of trust region algorithms of Powell [12] and we show that the solution of (1.5) satisfies the condition for convergence in [12]. In the last section, we show that we may achieve a superlinear rate of convergence.

2. The basic algorithm. We include three different termination rules in the preconditioned conjugate gradient method. We terminate when we have a sufficiently good approximation to the quasi-Newton step, which is the inexact quasi-Newton step [4], [13]. Secondly, we terminate when the norm of the approximation is too large, in which case we take a linear combination of the previous iterate and the current one. Finally, when we encounter a direction of negative curvature then we move to the boundary. This may be regarded as a generalized dogleg scheme. We then extend this strategy and show how to find a new approximate solution when the trust region bound is reduced and how to find a double dogleg curve.

Consider the trust region problem (1.2):

$$\min \{\varphi(p): \|p\|_C \leqq \Delta\},$$

where φ is defined in (1.1) and we use the weighted norm (1.3)

$$\|p\|_C = (p, Cp)^{1/2},$$

where C is symmetric and positive definite and $(\cdot, \cdot)$ is the standard inner-product on $\mathbb{R}^n$.

The algorithm is:

Step 1: Set $p_0 = 0$ and $r_0 = -g$.
        Solve $C\tilde{r}_0 = r_0$.
        Set $d_0 = \tilde{r}_0$
        and set $i = 0$.

Step 2: Compute $\gamma_i = (d_i, Bd_i)$.
        If $\gamma_i > 0$ then Continue with Step 3.
(2.1)   Otherwise compute $\tau > 0$ so that $\|p_i + \tau d_i\|_C = \Delta$,
        set $p = p_i + \tau d_i$ and Terminate.

Step 3: Compute $\alpha_i = (r_i, \tilde{r}_i)/\gamma_i$,
        $p_{i+1} = p_i + \alpha_i d_i$.
        If $\|p_{i+1}\|_C < \Delta$ then Continue with Step 4.
(2.2)   Otherwise compute $\tau > 0$ so that $\|p_i + \tau d_i\|_C = \Delta$,
        set $p = p_i + \tau d_i$ and Terminate.

Step 4: Compute $r_{i+1} = r_i - \alpha_i B d_i$.
(2.3)   If $\|r_{i+1}\|_C / \|g\|_C \leq \xi$ then set $p = p_{i+1}$ and Terminate.
         Otherwise Continue with Step 5.

Step 5: Solve $C\tilde{r}_{i+1} = r_{i+1}$.
         Compute $\beta_i = (r_{i+1}, \tilde{r}_{i+1}) / (r_i, \tilde{r}_i)$
         and $d_{i+1} = \tilde{r}_{i+1} + \beta_i d_i$.
         Set $i := i + 1$ and Continue with Step 2.

Consider the case when the algorithm has terminated and p is the final direction. Then

$$(2.4) \qquad p = \begin{cases} p_{i+1} & \text{if (2.3) is used,} \\ p_i + \tau d_i & \text{if } \gamma_i > 0 \text{ and (2.2) is used,} \\ p_i + \tau d_i & \text{if } \gamma_i \leq 0 \text{ and (2.1) is used.} \end{cases}$$

In Theorem 2.1 it is shown that $(p_i, Cd_i) > 0$ hence τ in (2.1) and (2.2) is found by choosing the positive root of the quadratic equation in τ:

$$(2.5) \qquad \tau^2 (d_i, Cd_i) + 2\tau (p_i, Cd_i) = \Delta^2 - (p_i, Cp_i).$$

We now state our main theorem. Before we prove the result, we need some properties of the PCG method.

THEOREM 2.1. *Let* $p_j, j = 0, \cdots, i$, *be the iterates generated by the algorithm. Then* $\varphi(p_j)$ *is strictly decreasing* $j = 0, \cdots, i$, *and*

$$(2.6) \qquad \varphi(p) \leq \varphi(p_i).$$

Further, $\|p_j\|_C$ *is strictly increasing* $j = 0, \cdots, i$ *and*

$$(2.7) \qquad \|p\|_C > \|p_i\|_C.$$

In the proof of the theorem, we need the following lemma from Hestenes and Stiefel [8] and Steihaug [13].

LEMMA 2.2. *Assume* $\gamma_i \neq 0, i = 0, 1, \cdots . k$; *then*

$$(2.8) \qquad (r_i, d_j) = (r_j, \tilde{r}_j), \qquad 0 \leq i \leq j \leq k,$$

$$(2.9) \qquad (d_i, Cd_j) = \frac{(r_j, \tilde{r}_j)(d_i, Cd_i)}{(r_i, \tilde{r}_i)}, \qquad 0 \leq i \leq j \leq k,$$

and

$$(2.10) \qquad \varphi(p_{i+1}) = \varphi(p_i) - \frac{1}{2} \frac{(r_i, \tilde{r}_i)^2}{\gamma_i}, \qquad 0 \leq i < k.$$

Proof of Theorem 2.1. We first show that $\|p_j\|_C, j = 0, 1, \cdots, i$, is strictly increasing and $\|p\|_C > \|p_i\|_C$. From Step 3 we have

$$(2.11) \qquad p_j = \sum_{k=0}^{j-1} \alpha_k d_k, \qquad j = 0, 1, \cdots, i$$

and

$$(2.12) \qquad \alpha_j > 0, \qquad j = 0, 1, \cdots, i-1.$$

Hence

$$(2.13) \qquad (p_j, Cd_j) > 0, \qquad j = 0, 1, \cdots, i$$

using (2.9). Consider

$$(2.14) \qquad (p_{j+1}, Cp_{j+1}) = (p_j, Cp_j) + 2\alpha_j(p_j, Cd_j) + \alpha_j^2(d_j, Cd_j) \geqq (p_j, Cp_j)$$

using (2.13) and (2.12). We have thus shown that $\|p_j\|_C$, $j = 0, 1, \cdots, i$, is strictly increasing. If $p = p_{i+1}$, then (2.7) follows from (2.14). If $p = p_i + \tau d_i$, then

$$(p, Cp) = (p_i, Cp_i) + 2\tau(p_i, Cd_i) + \tau^2(d_i, Cd_i) > (p_i, Cp_i)$$

using (2.13) and for τ in (2.5) $\tau > 0$. Hence we have (2.7).

We now show that $\varphi(p_j)$, $j = 0, 1, \cdots, i$, is strictly decreasing and (2.6) holds. From Step 2 we have $\gamma_j > 0$, $j = 0, 1, \cdots, i - 1$; then, from Lemma 2.2, $\varphi(p_j)$ is strictly decreasing $j = 0, 1, \cdots, i$.

Consider the last iterate p. If $p = p_{i+1}$ then the result follows directly. From (2.5), we have that

$$(r_i, d_i) = (r_i, \tilde{r}_i) = (Bp_i + g, C^{-1}(Bp_i + g)) > 0,$$

using that C is positive definite, hence d_i is a descent direction for $\varphi(p_i)$, i.e.,

$$-(\nabla\varphi(p_i), d_i) = (Bp_i + g, d_i) > 0.$$

If $\gamma_i > 0$, then we have

$$\varphi(p_i) \geqq \varphi(p_i + \tau d_i) \geqq \varphi(p_{i+1}) \quad \text{for } 0 < \tau \leqq \alpha_i.$$

Since $\tau \leqq \alpha_i$, we have the desired result. For $\gamma_i \leqq 0$, then the quadratic term is nonpositive, and we have

$$\varphi(p_i) \geqq \varphi(p_i + \tau d_i) \quad \text{for } 0 \leqq \tau,$$

and the result follows directly.   Q.E.D.

This theorem tells us that if B is positive definite and we terminate using (2.3), then we know that the unconstrained minimizer is outside the region. If we terminate using (2.1), then the matrix B is indefinite and the minimizer is at the boundary.

Since α_j in Step 3 is the minimizer of $\varphi(p_j + \alpha d_j)$, it follows that $\varphi(p_j + \tau(p_{j+1} - p_j))$ is strictly monotonic decreasing for $0 \leqq \tau \leqq 1$. From (2.1) and (2.2) we know that $\varphi(p + \tau(p - p_i))$ is strictly decreasing. Hence for the piecewise linear curve $\tilde{p}(\tau)$ that connects p_j, $j = 0, 1, \cdots, i$, and p we have that $\varphi(\tilde{p}(\tau))$ is strictly monotonic decreasing and $\|\tilde{p}(\tau)\|_C$ is strictly monotonic increasing.

Consider the case when the trust bound is reduced to Δ' and the matrix B is unchanged. Let $j < i$ be the index so that

$$(2.15) \qquad \|p_j\|_C \leqq \Delta' < \|p_{j+1}\|_C \leqq \|p_i\|_C \leqq \Delta,$$

then it would be desirable to combine p_j and p_i and find τ so that

$$\|p_j + \tau(p_i - p_j)\|_C = \Delta'.$$

From (2.11), (2.12) and (2.13) we immediately have that $(p_i - p_j)$ is a descent direction for $\varphi(p_j)$ and $\|p_j + \tau(p_i - p_j)\|_C$ is increasing in τ, $0 \leqq \tau \leqq 1$.

Let $j < i$. It can be shown that p_j and p_i may be used as a pair of directions for a double dogleg scheme, i.e., we can find positive $\eta \leqq 1$ so that

$$\tilde{p}(\tau) = \begin{cases} \tau p_j, & 0 \leqq \tau \leqq 1, \\ p_j + (\tau - 1)(\eta p_i - p_j), & 1 < \tau \leqq 2, \\ (\tau - 1)\eta p_i, & 2 < \tau \leqq 1 + \dfrac{1}{\eta}, \end{cases}$$

and $\varphi(\tilde{p}(\tau))$ is decreasing and $\|\tilde{p}(\tau)\|_C$ is increasing for increasing τ.

To see this, we first note that $p_i - p_j$ is a strict descent direction for $\varphi(p_j)$, hence there exists positive $\eta_0 \leqq 1$, so that $\eta p_i - p_j$, where

$$\eta_0 \leqq \eta \leqq 1,$$

is a descent direction for $\varphi(p_j)$. The constant η_0 is found by solving a quadratic equation in η_0 in a similar way as in Dennis and Mei [3].

Further, $\|p_j + \tau(p_i - p_j)\|_C$ is strictly increasing in τ, hence there exists positive $\eta_1 \leqq 1$, so that for $\eta p_i - p_j$, where

$$\eta_1 \leqq \eta \leqq 1,$$

then $\|p_j + \tau(\eta p_i - p_j)\|_C$ is strictly increasing in τ. The constant η_1 is found by solving a quadratic equation in η_1.

Now choose any η so that

$$\max\{\eta_0, \eta_1\} \leqq \eta \leqq 1,$$

and we have that $\tilde{p}(\tau)$ is biased toward p_i as the original double dogleg curve is biased toward the quasi-Newton step.

3. Convergence results. In this section, we summarize the convergence results for the unconstrained minimization problem when we use the modified PCG algorithm imbedded in the class of minimization algorithms of Powell [12].

Consider the unconstrained minimization problem

$$\min\{f(x): x \in \mathbb{R}^n\},$$

where $f: \mathbb{R}^n \to \mathbb{R}$ satisfies:

A.1. f is bounded below;

A.2. f is differentiable and the gradient is uniformly continuous.

The framework of the algorithms is given below. For simplicity we assume that $C_k = I$, where C_k is the weight in (1.3). Let $\bar{\Delta}$ be a fixed upper bound on the step lengths.

Step 1: Given $x_0, B_0, \Delta_0 \leqq \bar{\Delta}$ and ξ_0.
Calculate $f(x_0)$ and $g(x_0)$, and set $k = 0$.

Step 2: Use the modified PCG algorithm to find approximate solution p_k with relative error ξ_k in (2.3).

Step 3: Calculate $f(x_k + p_k)$, $g(x_k + p_k)$ and

$$(3.1) \qquad \rho_k = \frac{f(x_k) - f(x_k + p_k)}{-\varphi_k(p_k)}.$$

Step 4: Update x_{k+1}:

$$(3.2) \qquad x_{k+1} = \begin{cases} x_k + p_k & \text{if } \rho_k > \alpha_2, \\ x_k & \text{otherwise.} \end{cases}$$

Update Δ_{k+1}:

$$(3.3) \qquad \|p_k\| \leqq \Delta_{k+1} \leqq \min\{\gamma_3\|p_k\|, \bar{\Delta}\} \quad \text{when } \rho_k \geqq \alpha_1,$$

$$(3.4) \qquad \gamma_4\|p_k\| \leqq \Delta_{k+1} \leqq \gamma_5\|p_k\| \quad \text{when } \rho_k < \alpha_1.$$

Update B_{k+1}:

(3.5)
$$\|B_{k+1}\| \leq \gamma_1 + \gamma_2 \sum_{i=0}^{k} \|p_i\|.$$

Update ξ_{k+1}, $k := k+1$, and continue with Step 2.

The constants γ_1 and γ_2 depend on the update formula of B_{k+1}. Further,

(3.6)
$$0 \leq \alpha_2 < \alpha_1 < 1,$$
$$0 < \gamma_4 \leq \gamma_5 < 1 \leq \gamma_3.$$

If the matrix B_{k+1} is a secant update, then we have to decide whether or not to update when the step is rejected. If the step is rejected and the trust region parameter decreased, then the new direction in general is different from the old one, which is not the case in line searches. Hence we can expect to get a better approximation by updating the matrix even if the step is rejected. However, from § 2 we have that the new step can easily be computed if $B_{k+1} = B_k$ and $\rho_k \leq \alpha_2$. We note that this is the case when B_k only depends on x_k as in Newton's method.

The first theorem is an important modification of the original convergence result of Powell [12]. In his convergence result Powell assumes that the direction p_k satisfies

(3.7)
$$-\varphi_k(p_k) \geq c_1 \|g(x_k)\| \min \left\{ \|p_k\|, \frac{\|g(x_k)\|}{\|B_k\|} \right\}$$

for some $c_1 > 0$, and he shows that

$$\min \{\varphi_k(\tau g(x_k)): \|\tau g(x_k)\| \leq \|p_k\|\} \leq -\tfrac{1}{2}\|g(x_k)\| \min \left\{ \|p_k\|, \frac{\|g(x_k)\|}{\|B_k\|} \right\}.$$

But from Theorem 2.1 we have

$$\varphi_k(p_k) \leq \min \{\varphi_k(\tau g(x_k)): \|\tau g(x_k)\| \leq \|p_k\|\}$$

with equality only if $i = 0$ in (2.1) or (2.2). Hence the direction p_k from Step 2 satisfies (3.7) with $c_1 = \tfrac{1}{2}$. Note that when Δ_k is reduced, $x_k + p_k$ is rejected and B_k is unchanged, then we have the above inequalities using (2.15). We only include the part of the proof that differs from Powell's original proof. The algorithm includes the generalization of Thomas [15].

THEOREM 3.1. *If $\xi_k \leq \xi < 1$, then*

$$\liminf_{k \to \infty} \|g(x_k)\| = 0.$$

Proof. The result is proved by obtaining a contradiction. Assume that $\{\|g(x_k)\|\}$ is bounded away from zero. First it is shown that

(3.8)
$$\sum_{k \geq 0} \|p_k\| < +\infty,$$

hence $\{\|B_k\|\}$ is uniformly bounded. The next major step is to show that

(3.9)
$$\lim_{k \to \infty} \rho_k = 1.$$

Hence from (3.3) and (3.6) we can conclude that for k sufficiently large,

$$\Delta_{k+1} \geq \|p_k\|.$$

We will now show that for k sufficiently large, under the above assumption we do *not* use an inexact quasi-Newton step[1], i.e.,

$$(3.10) \qquad \frac{\|r_k\|}{\|g(x_k)\|} > \xi_k \quad \text{for } k \geqq k_0.$$

From (3.8) we have that $\|p_k\| \to 0$. Since we have assumed that $\{\|g(x_k)\|\}$ is bounded away from zero, it follows that the fraction

$$\frac{\|B_k p_k + g(x_k)\|}{\|g(x_k)\|}$$

is arbitrarily close to 1, hence for sufficiently large k we have that

$$\frac{\|r_k\|}{\|g(x_k)\|} > \xi \quad \text{for } k \geqq k_0.$$

Thus we have the desired result (3.10).

For k_0 sufficiently large we are using termination (2.1) or (2.2), and we have

$$\Delta_{k+1} \geqq \Delta_k = \|p_k\| \quad \text{for } k \geqq k_0.$$

Hence $\|p_{k+1}\| \geqq \|p_k\| > 0$ for $k \geqq k_0$. This contradicts (3.8) and shows that $\{\|g(x_k)\|\}$ is not bounded away from zero. Q.E.D.

The next theorem follows from Thomas [15].

THEOREM 3.2. *Let* $\alpha_1 > \alpha_2 > 0$ *and assume that* $\{\|B_k\|\}$ *is bounded. If* $\xi_k \leqq \xi < 1$, *then*

$$\lim_{k \to \infty} g(x_k) = 0.$$

Proof. The proof in Thomas [15, Thm. 4.5] only uses that p_k satisfies (3.7).

Q.E.D.

Powell has pointed out [12, p. 6] that if one of the iterates falls into a bounded region in which f is convex, and the subsequent points remain there, then the sequence $\{x_k\}$ converges to the local minimizer.

To explore this remark further, we will discuss under what assumptions there will be local convergence to a strong minimizer.

THEOREM 3.3. *Let* f *be twice continuously differentiable in an open convex set* Ω. *Assume that* x^* *is a local minimizer in* Ω *and the Hessian matrix of* f *at* x^* *is positive definite. Then there exist positive constants* ε *and* Δ *so that if* $\|x_0 - x^*\| \leqq \varepsilon$ *and* $\Delta_0 \leqq \Delta$ *then* $\{x_k\}$ *converges to* x^*.

Proof. Put $B^*(\delta) = \{x : \|x - x^*\| \leqq \delta\}$, and let $\delta > 0$ so that f is strictly convex in $B^*(\delta)$ and $B^*(\delta)$ is contained in Ω. Put

$$\mu = \frac{\delta}{3\gamma_3}.$$

Choose $\varphi > f(x^*)$ and $\varepsilon > 0$ so that

$$(3.11) \qquad B^*(\varepsilon) \subseteq L_\varphi \subseteq B^*(\mu),$$

where

$$L_\varphi = \{x \in B^*(\delta) : f(x) \leqq \varphi\}.$$

---

[1] This is the only modification of the original proof [12, pp. 11–12].

This can be done in view of the strict convexity of f in $B^*(\delta)$.

We will now show that for

$$\|x_0 - x^*\| \leqq \varepsilon, \qquad \Delta_0 \leqq \tfrac{2}{3}\delta,$$

then

(3.12) $$\|x_k - x^*\| \leqq \mu$$

and

(3.13) $$\Delta_k \leqq \tfrac{2}{3}\delta.$$

Hence we know that $\{x_k\}$ has limit points. Further, from Theorem 3.1 we know that there exists a subsequence converging to x^*. But $f(x_k)$ is decreasing, hence the sequence $\{x_k\}$ is converging to x^*.

The proof is by induction. Clearly (3.12) and (3.13) are satisfied for $k = 0$. Assume (3.12) and (3.13) hold for k and consider

$$\|x_k + p_k - x^*\| \leqq \|x_k - x^*\| + \|p_k\| \leqq \frac{\delta}{3\gamma_3} + \frac{2\delta}{3} \leqq \delta$$

and $x_k + p_k \in B^*(\delta)$. We will now consider the two different cases; the new function value is less than or larger than the current function value. If $f(x_k + p_k) \leqq f(x_k)$, then $x_k + p_k \in L_\varphi$, hence from (3.11)

$$\|x_k + p_k - x^*\| \leqq \frac{\delta}{3\gamma_3} = \mu,$$

and we have the desired result (3.12). From (3.4) and (3.6) we have that the maximum change in Δ_{k+1} is

$$\Delta_{k+1} \leqq \gamma_3 \|p_k\| \leqq \gamma_3 [\|x_k + p_k - x^*\| + \|x_k - x^*\|] \leqq \gamma_3 \left[\frac{\delta}{3\gamma_3} + \frac{\delta}{3\gamma_3} \right] \leqq \tfrac{2}{3}\delta,$$

and we have the desired (3.13).

If $f(x_k + p_k) > f(x_k)$, then the step is $x_{k+1} = x_k$ and the stepbound is decreased

$$\Delta_{k+1} \leqq \|p_k\| \leqq \Delta_k \leqq \tfrac{2}{3}\delta,$$

and we have shown (3.12) and (3.13).   Q.E.D.

This theorem shows that we get local convergence independent of B_0. To find a suitable Δ_0, we can choose

$$\Delta_0 = \gamma \|g(x_0)\|$$

for any $\gamma > 0$. Hence for x_0 sufficiently close to x^* we have $\Delta_0 \leqq \Delta$.

4. Superlinear rate of convergence. In Steihaug [13] and Eisenstat and Steihaug [4] it is shown that if f is sufficiently smooth and $\{x_k\}$ converges to a stationary point, where the Hessian is nonsingular and

(4.1) $$\lim_{k \to \infty} \frac{\|B_k p_k + g(x_k) - g(x_k + p_k)\|}{\|p_k\|} = 0,$$

and the residual $r_k = B_k p_k + g(x_k)$ satisfies

$$\frac{\|r_k\|}{\|g(x_k)\|} \leqq \xi_k,$$

where $x_{k+1} = x_k + p_k$ and $\xi_k \to 0$, then the sequence $\{x_k\}$ converges superlinearly. In this section, we will show that if we have convergence to a local minimizer where the Hessian is positive definite, then the trust region algorithm also yields a superlinear rate of convergence.

In this section we assume that

A.1. $x_k \to x^*$ as $k \to \infty$.

A.2. f is twice continuously differentiable in an open neighborhood Ω of x^*.

A.3. The Hessian matrix $H(x^*)$ of f is positive definite.

A.4. The sequences $\{p_k\}$ and $\{B_k\}$ satisfy (4.1).

The conditions (3.5) and (4.1) are satisfied if $B_k = H(x_k)$ and if B_k is a least change secant update in Frobenius norm. The proofs that these updates satisfy (3.5) depend on H being Lipschitz continuous. The proofs that the updates satisfy (4.1) depend on $\sum_{k \geq 0} \|p_k\| < \infty$. This follows from Powell [12, Thm. 2] (see [12], [14], [16]).

The first lemma shows that the approximation $\varphi_k(p_k)$ predicts $f(x_k + p_k)$ provided that the gradient of φ_k predicts the gradient of f at $x_k + p_k$. Let ρ_k be the fraction of the actual function decrease and the predicted decrease given by (3.1).

LEMMA 4.1.

$$\lim_{k \to \infty} \rho_k = 1.$$

Proof. The proof is a part of Powell's [12, Thm. 3, pp. 16–18]. Q.E.D.

This result is important since it specifies that for k_0 sufficiently large we use the trust region revision (3.3) and

$$\Delta_k \geq \|p_{k-1}\|, \qquad k > k_0.$$

Therefore, if at any index $k \geq k_0$

$$\|p_k\| < \|p_{k-1}\|,$$

then we know that the method has used the termination rule

$$\|r_k\| / \|g(x_k)\| \leq \xi_k.$$

This is the basic observation used in this section's main theorem.

THEOREM 4.2. *If* $\xi_k \to 0$ *then* $\{x_k\}$ *converges superlinearly, i.e.,*

$$\lim_{k \to \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

Proof. From the observations after Lemma 4.1 we know that since $\{x_k\}$ converges, we are using the termination rule

$$\|B_k p_k + g(x_k)\| / \|g(x_k)\| \leq \xi_k$$

in the minor iterations an infinite number of times. Further, for k sufficiently large, we have from (3.2) that the step p_k is acceptable. We will show that the norm of the new step p_{k+1} is strictly less than the norm of the current step p_k, hence the step p_{k+1} also satisfies the relative residual termination rule, since otherwise $\|p_{k+1}\| = \Delta_{k+1} \geq \|p_k\|$. This shows that when we start taking inexact quasi-Newton steps, then we will continue with these steps.

From Lemma 4.1, (3.2) and (3.3) it follows that there exists k_0 so that for $k \geq k_0$ then $x_{k+1} = x_k + p_k$ and $\Delta_{k+1} \geq \|p_k\|$. Further, from Powell [12, p. 15] it follows that there exists $\mu > 0$ so that

$$(4.2) \qquad \|p_k\| \leq \mu \|g(x_k)\|, \qquad k \geq k_0.$$

We first note that there exist positive constants λ_0 and λ_1 and a neighborhood $D \subseteq \Omega$ so that for all $x \in D$

$$(4.3) \qquad \lambda_1(p, p) \leq (p, H(x)p) \leq \lambda_0(p, p)$$

and $x_k \in D$ for $k \geq k_0$. Let

$$\gamma_k = \|B_k p_k + g(x_k) - g(x_k + p_k)\| / \|p_k\|$$

and choose $k_1 \geq k_0$ so that

$$(4.4) \qquad \mu[\xi_k \lambda_0 + \gamma_k] < 1 - \xi_k.$$

This can be done since $\xi_k \to 0$ and $\gamma_k \to 0$ as $k \to \infty$.

Consider

$$\frac{\|g(x_k + p_k)\|}{\|p_k\|} \leq \frac{\|B_k p_k + g(x_k)\|}{\|g(x_k)\|} \frac{\|g(x_k)\|}{\|p_k\|} + \|B_k p_k + g(x_k) - (x_k + p_k)\| / \|p_k\|$$

and for $k \geq k_0$

$$(4.5) \qquad \|g(x_k)\| \leq \|g(x_k + p_k) - g(x_k)\| + \|g(x_k + p_k)\| \leq \|p_k\| \left[\lambda_0 + \frac{\|g(x_k + p_k)\|}{\|p_k\|} \right]$$

using (4.3). Hence, for index $k \geq k_1$ such that

$$(4.6) \qquad \frac{\|r_k\|}{\|g(x_k)\|} \leq \xi_k$$

we have using (4.5) that

$$\frac{\|g(x_{k+1})\|}{\|p_k\|} \leq \xi_k \left[\lambda_0 + \frac{\|g(x_{k+1})\|}{\|p_k\|} \right] + \gamma_k$$

and

$$(4.7) \qquad \frac{\|g(x_{k+1})\|}{\|p_k\|} \leq \frac{\xi_k \lambda_0 + \gamma_k}{1 - \xi_k}.$$

Hence

$$(4.8) \qquad \frac{\|p_{k+1}\|}{\|p_k\|} = \frac{\|p_{k+1}\|}{\|g(x_{k+1})\|} \frac{\|g(x_{k+1})\|}{\|p_k\|} \leq \mu \frac{\xi_k \lambda_0 + \gamma_k}{1 - \xi_k} < 1,$$

using (4.7) and (4.4).

This shows that if $k \geq k_1$ so that (4.6) holds, then all consecutive iterates will satisfy (4.6). But we know that such k exist since $\|p_k\| \to 0$, hence there exist $k_2 \geq k_1$ so that for $k \geq k_2$ then (4.8) holds.

Now since γ_k and ξ_k are converging to zero we have from (4.7) that

$$\lim_{k \to \infty} \frac{\|g(x_{k+1})\|}{\|x_{k+1} - x_k\|} = 0$$

and using (4.2)

$$\lim_{k \to \infty} \frac{\|g(x_{k+1})\|}{\|g(x_k)\|} = 0$$

which shows a superlinear rate of convergence.   Q.E.D.

REFERENCES

[1] O. AXELSSON, *Solution of linear systems of equations: iterative methods*, in Sparse Matrix Techniques, V. A. Barker, ed., Lecture Notes in Mathematics 572, Springer-Verlag, New York, 1976, pp. 1–51.

[2] R. CHANDRA, *Conjugate gradient methods for partial differential equations*, Ph.D. dissertation, Res. Rep. 129, Dept. Computer Science, Yale University, New Haven, CT, 1978.

[3] J. E. DENNIS JR. AND H. H. W. MEI, *Two new unconstrained optimization algorithms which use function and gradient values*, J. Optim. Theory Applications, 28 (1979), pp. 453–482.

[4] S. C. EISENSTAT AND T. STEIHAUG, *Local analysis of inexact quasi-Newton methods*, Rep. TR 82-7, Dept. Mathematical Sciences, Rice University, Houston, TX, 1982.

[5] W. H. ENRIGHT AND S. M. SERBIN, *A note on the efficient solution of matrix pencil systems*, BIT, 18 (1978), pp. 276–281.

[6] D. M. GAY, *Computing optimal locally constrained steps*, SIAM J. Sci. Stat. Comp., 2 (1981), pp. 186–192.

[7] M. D. HEBDEN, *An algorithm for minimization using exact second derivatives*. Tech. Rep. T.P. 515, A.E.R.E., Theoretical Physics Division, Harwell, Berkshire, England, 1973.

[8] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*. J. Res. NBS, 49 (1952), pp. 409–436.

[9] J. J. MORÉ, *The Levenberg–Marquardt algorithm: Implementation and theory*, in Numerical Analysis, G. A. Watson, ed., Lecture Notes in Mathematics 630, Proceedings, Biennial Conference Dundee 1977, Springer-Verlag, New York, 1978, pp. 105–132.

[10] M. J. D. POWELL, *A new algorithm for unconstrained optimization*, in Nonlinear Programming, J. B. Rosen, O. L. Mangasarian and K. Ritter, eds., Academic Press, New York, 1970, pp. 31–65.

[11] ———, *A FORTRAN subroutine for unconstrained minimization, requiring first derivatives of the objective function*. Tech. Rep. AERE-R. 6469, Mathematics Branch, A.E.R.E. Harwell, Berkshire, England, 1970.

[12] ———, *Convergence properties of a class of minimization algorithms*, in Nonlinear Programming 2, O. L. Mangasarian, R. R. Meyer and S. M. Robinson, eds., Academic Press, New York, 1975, pp. 1–27.

[13] T. STEIHAUG, *Quasi-Newton methods for large scale nonlinear problems*, Ph.D. dissertation, Res. Rep. 49, School of Organization and Management, Yale University, New Haven, CT, 1980.

[14] ———, *On the sparse and symmetric least-change secant update*, Tech. Rep. TR82-4, Dept. Mathematical Sciences, Rice University, Houston, TX, 1982.

[15] S. W. THOMAS, *Sequential estimation techniques for quasi-Newton algorithms*, Ph.D. dissertation, Res. Rep. TR 75-227, Dept. Computer Science, Cornell University, Ithaca, NY, 1975.

[16] PH. L. TOINT, *On the superlinear convergence of an algorithm for solving a sparse minimization problem*, this Journal, 16 (1979), pp. 1036–1045.