Mathematical Programming 67 (1994) 247–264

A trust region algorithm for nonsmooth optimization

T. Bannert

Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen, Lotzestraße 16-18, D-37083 Göttingen, Germany

Received 7 May 1993; revised manuscript received 11 April 1994

Abstract

A trust region algorithm is proposed for minimizing the nonsmooth composite function $F(x) = h(f(x))$, where f is smooth and h is convex. The algorithm employs a smoothing function, which is closely related to Fletcher's exact differentiable penalty functions. Global and local convergence results are given, considering convergence to a strongly unique minimizer and to a minimizer satisfying second order sufficiency conditions.

Keywords: Trust region; Nonsmooth optimization; Exact differentiable penalty function; Maratos effect

1. Introduction

This paper presents a trust region algorithm for solving the unrestricted, nonsmooth optimization problem

$$\text{Minimize} \quad F(x) = h(f(x)) , \qquad x \in \mathbb{R}^n , \tag{1}$$

where $f: \mathbb{R}^n \to \mathbb{R}^m$ is (at least) once continuously differentiable and $h: \mathbb{R}^m \to \mathbb{R}$ is convex. Trust region algorithms for problems of this kind have been considered by Madsen [9], when h is the maximum norm, and by Fletcher [6], Yuan [13], and various other authors for the general case. From the current iteration point x_k, the iteration step is defined by minimizing a model function in a certain range around x_k, called the trust region. The size Δ_k of the trust region is adjusted after every step. The iteration step d_k in [6] as well as in this paper is defined as a solution of the subproblem

$$\begin{aligned} &\text{Minimize} \quad \Phi_k(d) := h(f(x_k) + f'(x_k)d) + \tfrac{1}{2} d^T B_k d \\ &\text{s.t.} \qquad \|d\| \leqslant \Delta_k \end{aligned} \tag{2}$$

SSDI 0025-5610(94)00031-N

with some symmetric matrix $B_k \in \mathbb{R}^{n\times n}$ and some norm $\|\cdot\|$ on $\mathbb{R}^n$. However, we use a different condition to decide whether d_k should be accepted, so $x_{k+1} := x_k + d_k$, or one prefers to define $x_{k+1} := x_k$ and reduce the trust region radius. Usually, d_k is accepted if it leads to a reduction of the objective function, while in this paper a reduction of a smoothing function is demanded. This smoothing function is the direct analogon to Fletcher's exact differentiable penalty functions for equality or inequality constrained smooth optimization problems [3–5]. A slightly different version of these penalty functions was used by Powell and Yuan [10] for a trust region algorithm for equality constrained smooth problems, which greatly influenced this paper.

The reason for introducing a smoothing function is the so-called Maratos effect, which can prevent superlinear convergence of the sequence $\{x_k\}$ generated by the algorithm to a local solution x^* of (1) satisfying second order sufficiency conditions. An example has been presented by Yuan [12]. The point is that, in spite of second order sufficiency conditions holding, there may exist a sequence $\{z_k\} \subset \mathbb{R}^n$ such that $z_k \to x^*$, $\|z_{k+1} - x^*\|_2 = O(\|z_k - x^*\|_2^2)$, but $F(z_{k+1}) > F(z_k)$ for infinitely many k, resulting from nondifferentiability of F. The steps d_k produced by subproblem (2) yield $\|x_k + d_k - x^*\|_2 = o(\|x_k - x^*\|_2)$ or better, if B_k is properly chosen and Δ_k has the right magnitude. Nevertheless, they may fail to reduce F and thus will not be accepted.

The usual way to overcome this difficulty (see Fletcher [7], Yuan [14]) is to put in a second order correction step $\tilde{d}_k$ whenever $F(x_k + d_k)$ is larger than it should be. $\tilde{d}_k$ is a solution of the subproblem

$$\begin{aligned} &\text{Minimize} \quad \tilde{\Phi}_k(d) := h(f(x_k + d_k) + f'(x_k)d) + \tfrac{1}{2}(d_k + d)^T B_k (d_k + d) \\ &\text{subject to} \quad \|d_k + d\| \leqslant \Delta_k, \end{aligned} \tag{3}$$

and x_{k+1} is defined as $x_k + d_k + \tilde{d}_k$, if this choice reduces F. In general, $x_k + d_k + \tilde{d}_k$ is not substantially closer to x^* than $x_k + d_k$, so the correction step does not accelerate convergence by itself. Hence we obtain the same order of convergence without a correction step, allowing that d_k may be accepted even if $F(x_k + d_k) > F(x_k)$. In order to maintain global convergence properties, a reduction of the smoothing function is demanded instead. The given analysis as well as our numerical tests suggest that the use of the smoothing function may be a real alternative to the second order correction step in order to avoid the Maratos effect.

In the next two sections, the smoothing function and the algorithm are presented. The following three sections consist of global and local convergence proofs, considering local convergence to a strongly unique solution and to a solution that satisfies second order sufficiency conditions. In the last section, the costs to evaluate the smoothing function are compared with the costs of the correction step.

Throughout this paper, we use the following notations: $\|x\|_2$ denotes the Euclidean norm of a vector x, and $\|A\|_2 := \sup\{\|Ax\|_2 : \|x\|_2 = 1\}$ for any matrix A. x^T and A^T denote the transpose. Let $\operatorname{im} A := \{Ax : x \in \mathbb{R}^l\}$ for $A \in \mathbb{R}^{k\times l}$. The gradient of a function $g : \mathbb{R}^n \to \mathbb{R}$ is denoted by $\nabla g(x)$ and regarded as a column vector, the Hessian is denoted by $\nabla^2 g(x)$. For $f : \mathbb{R}^n \to \mathbb{R}^m$, f_i denotes the ith component, and $f'(x)$ denotes the $m \times n$-matrix with the rows

$\nabla f_i(x)^T$. We further use directional derivatives and subdifferentials of nonsmooth functions (cf. e.g. [2]). For the convex function h,

$$\partial h(p)=\{\pi\in\mathbb{R}^m:\ h(p+q)\geqslant h(p)+\pi^Tq,\ \forall q\in\mathbb{R}^m\}.$$

For the composite function F, the chain rules

$$F'(x;d)=h'(f(x);f'(x)d),\qquad \partial F(x)=f'(x)^T\partial h(f(x))$$

hold. Finally $F'(x;d)\geqslant u^Td$ holds for all $u\in\partial F(x)$ and the same for h. A point $x^*\in\mathbb{R}^n$ is called a *stationary point* for (1), if $0\in\partial F(x^*)$.

2. The smoothing function

Let $\sigma>0$ be fixed. For any $x\in\mathbb{R}^n$, let $\delta(x)$ denote the unique solution of the problem

$$\text{Minimize}\quad \phi_x(d):=h(f(x)+f'(x)d)+\tfrac{1}{2}\sigma d^Td,\quad d\in\mathbb{R}^n \tag{4}$$

and define the *smoothing function*

$$\psi(x):=h(f(x)+f'(x)\delta(x))+\sigma\delta(x)^T\delta(x). \tag{5}$$

A necessary and sufficient condition for $\delta(x)$ to be the solution of (4) is that

$$0\in\partial\phi_x(\delta(x))=f'(x)^T\partial h(f(x)+f'(x)\delta(x))+\sigma\delta(x)$$

or, equivalently,

$$\exists\pi\in\partial h(f(x)+f'(x)\delta(x)):\quad f'(x)^T\pi+\sigma\delta(x)=0. \tag{6}$$

Lemma 2.1. *For any* $x\in\mathbb{R}^n$,

(a) $\psi(x)\leqslant F(x)$,

(b) $\delta(x)=0$ *if and only if x is a stationary point for* (1),

(c) $\|\delta(x)\|_2\leqslant(1/\sigma)\|f'(x)\|_2\cdot\min\{\|\pi\|_2:\ \pi\in\partial h(f(x))\}$,

(d) $\phi_x(d)\geqslant\phi_x(\delta(x))+\tfrac{1}{2}\sigma\|d-\delta(x)\|_2^2,\ \forall d\in\mathbb{R}^n$.

Proof. (a) can easily be derived from (5) and (6):

$$\psi(x)=h(f(x)+f'(x)\delta(x))-\pi^Tf'(x)\delta(x)\leqslant h(f(x))=F(x).$$

The chain rule for subdifferentials provides

$$\partial\phi_x(0)=f'(x)^T\partial h(f(x))=\partial F(x).$$

This proves (b), because $\delta(x)=0$ holds if and only if $0\in\partial\phi_x(0)$, while $0\in\partial F(x)$ is the condition for x to be a stationary point for (1).

For any $\pi\in\partial h(f(x))$, we have

$$F(x)\geqslant\psi(x)\geqslant h(f(x))+\pi^Tf'(x)\delta(x)+\sigma\delta(x)^T\delta(x).$$

Hence $\|\delta(x)\|_2 \leqslant (1/\sigma)\|f'(x)\|_2\|\pi\|_2$, and the minimum in (c) is attained, because $\partial h(f(x))$ is a closed convex set.

Finally, (d) follows from the fact that ϕ_x is uniformly convex: (6) implies

$$\begin{aligned}\phi_x(d)-\phi_x(\delta(x)) &= h(f(x)+f'(x)d)-h(f(x)+f'(x)\delta(x))\\ &\quad +\tfrac{1}{2}\sigma\|d\|_2^2-\tfrac{1}{2}\sigma\|\delta(x)\|_2^2\\ &\geqslant \pi^T f'(x)(d-\delta(x))+\sigma\delta(x)^T(d-\delta(x))+\tfrac{1}{2}\sigma\|d-\delta(x)\|_2^2\\ &= \tfrac{1}{2}\sigma\|d-\delta(x)\|_2^2\end{aligned}$$

for any $d\in\mathbb{R}^n$, which completes the proof. □

Lemma 2.2. *The functions δ and ψ defined by* (4) *and* (5) *are continuous.*

Proof. It is sufficient to show that δ is continuous. For any fixed $x\in\mathbb{R}^n$, let U be a small neighbourhood of x and $y\in U$. Part (d) of the previous lemma implies

$$\phi_x(\delta(y))\geqslant\phi_x(\delta(x))+\tfrac{1}{2}\sigma\|\delta(x)-\delta(y)\|_2^2,$$
$$\phi_y(\delta(x))\geqslant\phi_y(\delta(y))+\tfrac{1}{2}\sigma\|\delta(x)-\delta(y)\|_2^2.$$

Taking the sum, we obtain

$$\begin{aligned}\sigma\|\delta(x)-\delta(y)\|_2^2 &\leqslant |\phi_y(\delta(x))-\phi_x(\delta(x))|+|\phi_y(\delta(y))-\phi_x(\delta(y))|\\ &= |h(f(y)+f'(y)\delta(x))-h(f(x)+f'(x)\delta(x))|\\ &\quad +|h(f(y)+f'(y)\delta(y))-h(f(x)+f'(x)\delta(y))|\,.\end{aligned}$$

Since h as a convex function is locally Lipschitz, the subdifferentials $\{\partial h(f(y)): y\in U\}$ are uniformly bounded, and so are $\{\delta(y): y\in U\}$ from part (c) of Lemma 2.1. Therefore all arguments of h appearing in the above inequality are elements of a bounded set, on which we can choose a Lipschitz constant L for h. Now

$$\begin{aligned}&\sigma\|\delta(x)-\delta(y)\|_2^2\\ &\leqslant 2L\cdot\|f(x)-f(y)\|_2+L\cdot\|f'(x)-f'(y)\|_2\cdot(\|\delta(x)\|_2+\|\delta(y)\|_2)\xrightarrow{y\to x}0\,,\end{aligned}$$

which means that δ is continuous. □

If h is a polyhedral function, $f\in C^k$, and for some $x\in\mathbb{R}^n$ nondegeneracy and strict complementarity conditions hold in the solution of (4), then ψ is C^{k-1} in a neighbourhood of x. This corresponds to a property of Fletcher's [5] penalty function for smooth problems with inequality constraints. Since we will not explicitly use smoothness of ψ, we omit the proof; the main arguments are given in Section 6, applying to the case when x is a local solution satisfying second order sufficiency conditions.

There is a second version of the smoothing function: Define $\delta(x)$ as above and let

$$\hat{\psi}(x) := \phi_x(\delta(x)) = \psi(x) - \tfrac{1}{2}\sigma\delta(x)^T\delta(x) .$$

$\hat{\psi}$ can be used for a trust region algorithm in almost the same manner as ψ and has even better smoothness properties [1]. However, ψ seems to be the more natural choice, because applied to a smooth function $F(x) = f(x)$ (with $h(p) = p$), we have $\psi = F$, while $\hat{\psi} \neq F$.

3. The algorithm

Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^n$ and let $\rho \in (0, 0.5)$, $\tau_1 \in (0, 1)$, and $\tau_2 > 1$ be constant parameters. Consider the following algorithm:

Algorithm 3.1.

step 0: *Choose arbitrary starting values* $x_0 \in \mathbb{R}^n$, $\Delta_0 > 0$, $\sigma > 0$, *calculate* $\delta_0 := \delta(x_0)$, *let* $k := 0$.

step 1: *If* $\delta_k = 0$: *Stop*, x_k *is a stationary point for* (1).

step 2: *Choose* $B_k \in \mathbb{R}^{n \times n}$ *symmetric.*
If $\sigma < \|B_k\|_2$: *Let* $\sigma := 2\|B_k\|_2$, *recalculate* δ_k.

step 3a: *If* $\|\delta_k\| > \Delta_k$, *define* $d_k := \delta_k$;

3b: *otherwise let* d_k *be a solution of the subproblem*

$$\begin{array}{ll} \text{Minimize} & \Phi_k(d) := h(f(x_k) + f'(x_k)d) + \tfrac{1}{2}d^T B_k d \\ \text{s.t.} & \|d\| \leqslant \Delta_k . \end{array}$$

step 4: *Let* $s_k := 4 \cdot |F(x_k + d_k) - \Phi_k(d_k)| / \|d_k\|_2^2$.
If $\sigma < s_k$: *Let* $\sigma := 2s_k$, *recalculate* δ_k.

step 5: *Calculate* $\delta(x_k + d_k)$ *and* $q_k := (\psi(x_k) - \psi(x_k + d_k)) / (\psi(x_k) - \Phi_k(d_k))$.

step 6a: *If* $q_k \geqslant \rho$, *let* $x_{k+1} := x_k + d_k$, $\delta_{k+1} := \delta(x_k + d_k)$,
and choose $\Delta_{k+1} \in [\|d_k\|, \max(\tau_2\|d_k\|, \Delta_k)]$;

6b: *otherwise let* $x_{k+1} := x_k$, $\delta_{k+1} := \delta_k$, $\Delta_{k+1} := \tau_1\|d_k\|$.

step 7: *Increment k by one, go to step* 1.

In steps 0, 2, 4, and 5, calculate $\delta(\cdot)$ and $\psi(\cdot)$ according to (4) and (5), using the current value of the parameter σ.

In step 2, it is not necessary to calculate $\|B_k\|_2$ exactly. The following theorems remain valid, if $\|B_k\|_2$ is replaced by any sequence $\{b_k\}$ of upper bounds for $\|B_k\|_2$ satisfying $b_k \leqslant C \cdot \|B_k\|_2$ for some constant C.

The iteration step d_k is normally defined in step 3b, which consists of the usual trust region subproblem (2). Note that B_k is not demanded to be positive semidefinite. The alternative step selection in step 3a is seldom used, but important for convergence analysis.

It is motivated by Lemma 4.3: Since $d_k = \delta_k$ always yields a sufficient descent, a step within the trust region radius Δ_k would be less efficient in case $\Delta_k < \|\delta_k\|$.

The main difference to Fletcher's algorithm [6], however, is the definition of the ratio q_k in step 5. The success of the step d_k is valued by the change in the smoothing function ψ instead of F. q_k is the ratio of the actual reduction $\psi(x_k) - \psi(x_k + d_k)$ and the predicted reduction $\psi(x_k) - \Phi_k(d_k)$. $\Phi_k(d_k)$ turns out to be an even better estimate for $\psi(x_k + d_k)$ than for $F(x_k + d_k)$, at least in the situation considered in Section 6, which may give rise to the Maratos effect.

In steps 2 and 4 the parameter σ is adjusted, if necessary. Roughly spoken, the quadratic term $\frac{1}{2}\sigma d^T d$ in (4) must dominate all other second order terms occurring.

4. Global convergence

If Algorithm 3.1 terminates in step 1 of the kth iteration, then x_k is a stationary point for (1) according to Lemma 2.1. From now on we assume that an infinite sequence $\{x_k\}$ is generated. Our aim is to show that every accumulation point of $\{x_k\}$ is a stationary point for (1). We assume that

- the sequences $\{x_k\}$ and $\{B_k\}$ are bounded,
- f' is Lipschitz on a convex superset of $\{x_k\}$ with Lipschitz constant L_f.

As a consequence of the first assumption and Lemma 2.1, the trust region radii $\{\Delta_k\}$ and the sequence $\{\delta_k\}$ remain bounded, too. Hence all arguments of h appearing in the following will be contained in a bounded set, on which h as a convex function is also Lipschitz with some Lipschitz constant L_h.

Lemma 4.1. *The parameter σ is changed only a finite number of times.*

Proof. Since σ is at least doubled with every change, it is sufficient to show that it remains bounded. Let $\bar{b}$ be an upper bound for $\{\|B_k\|_2\}$. Then $\sigma \leqslant 2\bar{b}$ after a change in step 2, and

$$\sigma = 8\,\frac{|F(x_k + d_k) - \Phi_k(d_k)|}{\|d_k\|_2^2} \leqslant 4(L_h L_f + \bar{b})$$

after a change in step 4. Thus σ remains bounded and the lemma is proved. □

Lemma 4.2. *For all k,*

$$\psi(x_k) - \Phi_k(d_k) \geqslant \tfrac{1}{2}\sigma\|\delta_k\|_2^2$$

holds, so the denominator in step 5 is nonzero.

Proof. From step 3 of the algorithm, $\Phi_k(d_k) \leqslant \Phi_k(\delta_k)$ holds for all k. Therefore,

$$\psi(x_k) - \Phi_k(d_k) \geqslant \psi(x_k) - \Phi_k(\delta_k) = \sigma\delta_k^T\delta_k - \tfrac{1}{2}\delta_k^T B_k \delta_k \geqslant \tfrac{1}{2}\sigma\|\delta_k\|_2^2$$

regarding that $\sigma \geqslant \|B_k\|_2$ from step 2. □

Lemma 4.3. *If $d_k = \delta_k$ and σ is not changed in step 4, then d_k will be accepted. This means the algorithm does not infinitely loop with a constant x_k.*

Proof. If $d_k = \delta_k$, from Lemma 2.1(a) and Lemma 4.2 we have

$$1 - q_k = \frac{\psi(x_k + \delta_k) - \Phi_k(\delta_k)}{\psi(x_k) - \Phi_k(\delta_k)} \leqslant \frac{F(x_k + \delta_k) - \Phi_k(\delta_k)}{\frac{1}{2}\sigma\|\delta_k\|_2^2} \leqslant \tfrac{1}{2},$$

because we assumed σ to be unchanged in step 4. Therefore $q_k \geqslant \frac{1}{2} > \rho$, and δ_k will be accepted. □

Theorem 4.4. *Let the sequences $\{x_k\}$ and $\{B_k\}$ from Algorithm 3.1 be bounded, and let f' be Lipschitz on a convex superset of $\{x_k\}$. Then every accumulation point of $\{x_k\}$ is a stationary point for (1).*

Proof. Without loss of generality, σ can be regarded as a constant. From the previous lemma we know that there is an infinite number of k for which $q_k \geqslant \rho$. For each of these,

$$\psi(x_k) - \psi(x_{k+1}) \geqslant \rho \cdot (\psi(x_k) - \Phi_k(d_k)) \geqslant \tfrac{1}{2}\rho\sigma\|\delta_k\|_2^2$$

from Lemma 4.2. Since $\{\psi(x_k)\}$ is monotone and bounded, we obtain the limit $\delta_k \to 0$ (note that $\delta_k = \delta_{k+1}$ if $q_k < \rho$). Now continuity of δ from Lemma 2.2 supplies that $\delta(x^*) = 0$ for every accumulation point x^* of the sequence $\{x_k\}$, which completes the proof. □

5. Local convergence to a strongly unique solution

Let $x^* \in \mathbb{R}^n$ be a strongly unique solution of (1), that is, there exist constants $\alpha > 0$, $\epsilon > 0$ such that

$$F(x) - F(x^*) \geqslant \alpha \cdot \|x - x^*\|_2 \qquad \forall x \in \mathbb{R}^n\colon \quad \|x - x^*\|_2 \leqslant \epsilon. \tag{7}$$

We assume that

- $x_k \to x^*$, the sequence $\{B_k\}$ is bounded,
- f' is Lipschitz in a neighbourhood of x^* with Lipschitz constant L_f.

Thus we can keep on using the Lipschitz constant L_h for h from the previous section. Because of Lemma 4.1, σ can be regarded as a constant. We start with an easy consequence of the trust region adjustment strategy in step 6.

Lemma 5.1. *If the sequence $\{x_k\}$ generated by Algorithm 3.1 converges to a stationary point, then $d_k \to 0$.*

Proof. Let $\mathscr{A} := \{k \in \mathbb{N}\colon q_k \geqslant \rho\}$ denote the set of all iteration indices belonging to accepted steps. The statement is trivial for the subsequence $\{d_k\}_{k\in\mathscr{A}}$. If, however, $\mathbb{N}\backslash\mathscr{A}$ is infinite, even $\Delta_k \to 0$ holds: For every $k\in\mathbb{N}\backslash\mathscr{A}$, Δ_k is reduced by the constant factor $\tau_1 < 1$, while for $k\in\mathscr{A}$ either $\Delta_{k+1} \leqslant \Delta_k$ or $\Delta_{k+1} \leqslant \tau_2\|d_k\| \to 0$ holds. $\Delta_k \to 0$ implies $d_k \to 0$, because $\delta_k \to 0$ comes from Lemma 2.2. □

Theorem 5.2. *Let the sequence $\{x_k\}$ generated by Algorithm* 3.1 *converge to a strongly unique minimizer x^* of F, let f' be Lipschitz in a neighbourhood of x^*, and let the sequence of matrices $\{B_k\}$ be bounded. Then the order of convergence $x_k \to x^*$ is at least quadratic.*

Proof. The above made assumptions imply that

$$\phi_{x_k}(d) = F(x_k + d) + O(\|d\|_2^2),$$
$$\Phi_k(d) = F(x_k + d) + O(\|d\|_2^2)$$

uniformly for all k. Since $\delta_k \to 0$, we have that

$$\begin{aligned}
\|x_k + \delta_k - x^*\|_2 &\leqslant \frac{1}{\alpha}\,(F(x_k+\delta_k) - F(x^*)) \\
&= \frac{1}{\alpha}\,(\phi_{x_k}(\delta_k) - F(x^*)) + O(\|\delta_k\|_2^2) \\
&\leqslant \frac{1}{\alpha}\,(\phi_{x_k}(x^* - x_k) - F(x^*)) + O(\|\delta_k\|_2^2) \\
&= O(\|x_k - x^*\|_2^2) + O(\|\delta_k\|_2^2)\,.
\end{aligned}$$

Therefore

$$\|x_k + \delta_k - x^*\|_2 = O(\|x_k - x^*\|_2^2)\,. \tag{8}$$

Similarly, $d_k \to 0$ implies

$$\begin{aligned}
\|x_k + d_k - x^*\|_2 &\leqslant \frac{1}{\alpha}\,(F(x_k+d_k) - F(x^*)) \\
&= \frac{1}{\alpha}\,(\Phi_k(d_k) - F(x^*)) + O(\|d_k\|_2^2) \\
&\leqslant \frac{1}{\alpha}\,(\Phi_k(\delta_k) - F(x^*)) + O(\|d_k\|_2^2) \\
&= \frac{1}{\alpha}\,(F(x_k+\delta_k) - F(x^*)) + O(\|d_k\|_2^2) + O(\|\delta_k\|_2^2) \\
&= O(\|x_k - x^*\|_2^2) + O(\|d_k\|_2^2)\,,
\end{aligned}$$

which shows that

$$\|x_k + d_k - x^*\|_2 = O(\|x_k - x^*\|_2^2)\,. \tag{9}$$

Finally, we have to show that d_k is accepted, i.e. $q_k \geqslant \rho$ for large k. From step 4 of the algorithm (note that σ was assumed to be constant) and Lemma 4.2, we have

$$1-q_k=\frac{\psi(x_k+d_k)-\Phi_k(d_k)}{\psi(x_k)-\Phi_k(d_k)}\leqslant\frac{F(x_k+d_k)-\Phi_k(d_k)}{\psi(x_k)-\Phi_k(d_k)}\leqslant\frac{\frac{1}{4}\sigma\|d_k\|_2^2}{\frac{1}{2}\sigma\|\delta_k\|_2^2}.$$

Since $\|d_k\|_2/\|\delta_k\|_2\to 1$ from (8) and (9),

$$\liminf_{k\to\infty} q_k \geqslant \tfrac{1}{2} > \rho,$$

which ensures that d_k is accepted for sufficiently large k. Thus the proof is complete. □

6. Local convergence under second order sufficiency conditions

In this section, let h be a polyhedral convex function:

$$h(p)=\max_{i=1,\ldots,\kappa} g_i+h_i^T p, \qquad g_i\in\mathbb{R}, \quad h_i\in\mathbb{R}^m \quad (i=1,\ldots,\kappa), \quad \kappa>0.$$

For any $p\in\mathbb{R}^m$, let $I(p):=\{i: g_i+h_i^Tp=h(p)\}$ denote the set of *active components*. The derivatives of h are given by

$$h'(p;q)=\max\{h_i^Tq: i\in I(p)\}, \qquad \partial h(p)=\mathrm{co}\{h_i: i\in I(p)\},$$

where co A stands for the convex hull of a set A. We gradually state second order sufficiency conditions on a point $x^*\in\mathbb{R}^n$ to be a local solution for (1). Let $p^*:=f(x^*)$. Since we will be able to restrict ourselves to an arbitrarily small neighbourhood of p^*, assume for simplicity that $I(p^*)=\{1,\ldots,\kappa\}$. We call

$$\begin{aligned}\mathscr{E}(h)&:=\{p\in\mathbb{R}^m: I(p)=I(p^*)\}\\&=\{p\in\mathbb{R}^m: (h_i-h_1)^T(p-p^*)=0\ (i=2,\ldots,\kappa)\}\end{aligned}$$

the *edge* of h through p^*. By the corresponding edges of F resp. ϕ_x, we mean the sets

$$\begin{aligned}\mathscr{E}(F)&:=\{x\in\mathbb{R}^n: f(x)\in\mathscr{E}(h)\},\\\mathscr{E}(\phi_x)&:=\{d\in\mathbb{R}^n: f(x)+f'(x)d\in\mathscr{E}(h)\},\end{aligned}$$

respectively. Introducing a maximum subset $\{i_1,\ldots,i_l\}\subseteq\{2,\ldots,\kappa\}$ such that the matrix

$$R:=(h_{i_1}-h_1|\ldots|h_{i_l}-h_1)\in\mathbb{R}^{m\times l} \tag{10}$$

has full rank l, we can write

$$\begin{aligned}\mathscr{E}(h)&=\{p\in\mathbb{R}^m: R^T(p-p^*)=0\},\\\mathscr{E}(F)&=\{x\in\mathbb{R}^n: R^T(f(x)-p^*)=0\}.\end{aligned}$$

The first condition we impose on x^* is that x^* be a stationary point for (1), i.e.

(A1) $\exists \pi^* \in \partial h(p^*)$: $f'(x^*)^T \pi^* = 0$.

Secondly, we demand the set $\mathscr{E}(F)$ to be regular at x^*:

(A2) $\operatorname{rank}(R^T f'(x^*)) = l$ *(nondegeneracy)*.

Note that the affine hull of $\partial h(p^*)$ is $\pi + \operatorname{im} R$ with an arbitrary $\pi \in \partial h(p^*)$. Hence (A2) implies that π^* is uniquely defined by (A1). We demand that π^* is contained in the relative interior of $\partial h(p^*)$, which means

(A3) $\exists \epsilon > 0$: $\{\pi \in \pi^* + \operatorname{im} R: \|\pi - \pi^*\|_2 \leqslant \epsilon\} \subseteq \partial h(p^*)$

(*strict complementarity*). Because $f'(x^*)^T$ is a one-to-one mapping from $\partial h(p^*)$ to $\partial F(x^*)$, (A3) is equivalent to

$$\exists \epsilon' > 0: \quad \{u \in \operatorname{im} f'(x^*)^T R: \|u\|_2 \leqslant \epsilon'\} \subseteq \partial F(x^*) , \tag{11}$$

regarding $f'(x^*)^T \pi^* = 0$. An arbitrary $d \in \mathbb{R}^n$ can be partitioned into a component $d^{\parallel}$ parallel and a component $d^{\perp}$ orthogonal to the tangent space to $\mathscr{E}(F)$ at x^*:

$$d = d^{\parallel} + d^{\perp} , \qquad d^{\perp} \in \operatorname{im}(f'(x^*)^T R) , \qquad R^T f'(x^*) d^{\parallel} = 0 .$$

Then (11) gives

$$F'(x^*, d) \geqslant \left(\epsilon' \frac{d^{\perp}}{\|d^{\perp}\|_2}\right)^T d = \epsilon' \|d^{\perp}\|_2 . \tag{12}$$

That means there is a strict ascent in F when moving from x^* into any direction that is not contained in the tangent space to $\mathscr{E}(F)$. Thus second order conditions must be imposed only for the elements of the tangent space to $\mathscr{E}(F)$:

(A4) Let f be twice differentiable and let $\nabla^2 f$ be Lipschitz in a neighbourhood of x^*.

(A5) $d^T W^* d > 0 \quad \forall d \in \mathbb{R}^n$: $d \neq 0$, $R^T f'(x^*)^T d = 0$,

where $W^* := \sum_{i=1}^{m} \pi_i^* \nabla^2 f_i(x^*)$.

Ultimately, we assume

(A6) $x_k \to x^*$, $B_k \to W^*$

for the sequences from Algorithm 3.1.

The results of this section remain valid even without strict complementarity holding (see [1]). The Lipschitz condition on $\nabla^2 f$ is needed only for second order convergence, while Theorem 6.4 still holds if $\nabla^2 f$ is only supposed to be continuous. The proof can also be modified in a way that only

$$(B_k - W^*)(x_k - x^*) / \|x_k - x^*\|_2 \to 0 , \qquad (B_k - W^*) d_k / \|d_k\|_2 \to 0 ,$$

and relation (21) are required instead of $B_k \to W^*$.

An important point of the proof is that for large k both δ_k and d_k are in the edge $\mathscr{E}(\phi_{x_k}) = \mathscr{E}(\Phi_k)$. This means that both subproblems (4) and (2) correctly predict the set of components that are active in x^*. Since these results are standard except from the influence of the constraint in (2), we give only sketches of the proofs. For details, see [11].

For all x sufficiently close to x^*, $\delta(x)$ is given by the system

$$\begin{pmatrix} \sigma I & f'(x)^T R \\ R^T f'(x) & 0 \end{pmatrix} \begin{pmatrix} \delta(x) \\ \lambda(x) \end{pmatrix} = \begin{pmatrix} -f'(x)^T \pi^* \\ R^T(p^* - f(x)) \end{pmatrix}. \tag{13}$$

Note first that the matrix is regular for x close to x^* because of nondegeneracy condition (A2). The right-hand side is zero for $x = x^*$, and the solution $(\delta(x), \lambda(x))$ is differentiable as a function of x. From the second equation, we have

$$f(x) + f'(x)\delta(x) \in \mathscr{E}(h), \quad \text{so} \quad \delta(x) \in \mathscr{E}(\phi_x). \tag{14}$$

The first equation yields

$$\sigma\delta(x) + f'(x)^T \pi(x) = 0, \tag{15}$$

defining

$$\pi(x) := \pi^* + R\lambda(x). \tag{16}$$

From strict complementarity condition (A3), $\pi(x) \in \partial h(p^*)$ for x close to x^*. Since (14) implies $\partial h(f(x) + f'(x)\delta(x)) = \partial h(p^*)$, (15) means that the solution $\delta(x)$ of (13) actually satisfies condition (6) and hence is the solution of (4). With a little more care, one can even derive strict complementarity for $\delta(x)$ in a uniform sense: For all x sufficiently close to x^*,

$$\{\pi \in \pi(x) + \operatorname{im} R \colon \|\pi - \pi(x)\|_2 \leqslant \tfrac{1}{2}\epsilon\} \subseteq \partial h(p^*)$$

and

$$\phi_x'(\delta(x); d) \geqslant \tfrac{1}{2}\epsilon' \|d^\perp\|_2 \quad \forall d \in \mathbb{R}^n \tag{17}$$

hold, using the constants and arguments from (A3) and (12), and defining $d^\perp$ as the orthogonal projection of d into $\operatorname{im}(f'(x)^T R)$. We further need

Lemma 6.1. *There is a constant $c > 0$ such that $\|\delta(x)\|_2 \geqslant c \cdot \|x - x^*\|_2$ for all x sufficiently close to x^*.*

Proof. We take first derivatives on both sides of system (13). Observing $\delta(x^*) = 0$, $\lambda(x^*) = 0$, we obtain

$$\begin{pmatrix} \sigma I & f'(x^*)^T R \\ R^T f'(x^*) & 0 \end{pmatrix} \begin{pmatrix} \delta'(x^*) \\ \lambda'(x^*) \end{pmatrix} = \begin{pmatrix} -W^* \\ -R^T f'(x^*) \end{pmatrix}.$$

A short calculation gives

$$\delta'(x^*) = -\frac{1}{\sigma} P^{\parallel} W^* - P^{\perp},$$

where $P^{\perp}$ is the orthogonal projection into $\mathrm{im}(f'(x^*)^T R)$, and $P^{\parallel} := I - P^{\perp}$. (A5) implies that $\delta'(x^*)$ is nonsingular, which proves the inequality stated in the lemma. □

Next we transfer these results to subproblem (2). Define d_k^* by the system

$$\begin{pmatrix} B_k & f'(x_k)^T R \\ R^T f'(x_k) & 0 \end{pmatrix} \begin{pmatrix} d_k^* \\ \mu_k \end{pmatrix} = \begin{pmatrix} -f'(x_k)^T \pi^* \\ R^T(p^* - f(x_k)) \end{pmatrix}. \tag{18}$$

From (A2) and (A5), the matrix is nonsingular for large k. As above, one has

$$f(x_k) + f'(x_k) d_k^* \in \mathscr{E}(h), \quad d_k^* \in \mathscr{E}(\Phi_k), \tag{19}$$

and $0 \in \partial \Phi_k(d_k^*)$ with strict complementarity holding, which leads to

$$\Phi_k'(d_k^*; d) \geq \tfrac{1}{2}\epsilon' \|d^{\perp}\|_2 \quad \forall d \in \mathbb{R}^n \tag{20}$$

for all sufficiently large k. $d^{\perp}$ denotes the orthogonal projection of d into $\mathrm{im}(f'(x_k)^T R)$. From second order condition (A5) and $B_k \to W^*$ one can derive a constant $\eta > 0$ such that

$$d^T B_k d \geq \eta \|d\|_2^2 \qquad \forall d \in \mathbb{R}^n\colon \ R^T f'(x_k) d = 0 \tag{21}$$

holds for large k. Altogether, d_k^* satisfies conditions (A1) to (A5) for Φ_k and thus is a local minimizer for Φ_k. Note however that Φ_k is not convex if B_k has one or more negative eigenvalues, hence d_k^* is not a global minimizer in general.

In order to estimate the difference between d_k^* and the optimum step $x^* - x_k$ leading directly into the solution, multiply the vector $(x^* - x_k, 0)$ by the matrix from system (18) and compare the result with the right-hand side of (18) (cf. [11]). This leads to

$$\|x_k + d_k^* - x^*\|_2 = O(\|x_k - x^*\|_2 \cdot (\|x_k - x^*\|_2 + \|B_k - W^*\|_2)). \tag{22}$$

Hence the algorithm converges superlinearly, if we can prove that $d_k = d_k^*$ and $q_k \geq \rho$ for large k. We start with the first and show that $d_k = d_k^*$ except it is prevented by a too small trust region radius Δ_k. The theorem holds for the standard algorithm, too, ignoring the third alternative.

Theorem 6.2. *Let h be polyhedral, and let* (A1)–(A6) *be satisfied. Then for all sufficiently large k, at least one of the following alternatives is true*:

(I) $d_k = d_k^*$,

(II) $\|d_k\| = \Delta_k$ *and* $\Delta_k < \|d_k^*\|$,

(III) $d_k = \delta_k$ *and* $\Delta_k < \|\delta_k\|$.

Proof. If for some k both (I) and (III) are false, then d_k is a solution of subproblem (2) and $d_k \neq d_k^*$. We show that

$$\Phi_k'(d_k; d_k^* - d_k) < 0,$$

so (II) must hold, because otherwise $d_k^* - d_k$ were a feasible direction from d_k, so d_k were non-optimal for (2). First we have

$$\begin{aligned}
&\Phi_k'(d_k;\, d_k^* - d_k) + \Phi_k'(d_k^*;\, d_k - d_k^*) \\
&\quad = h'(f(x_k) + f'(x_k)d_k; f'(x_k)(d_k^* - d_k)) + d_k^T B_k(d_k^* - d_k) \\
&\qquad + h'(f(x_k) + f'(x_k)d_k^*; f'(x_k)(d_k - d_k^*)) + d_k^{*T} B_k(d_k - d_k^*) \\
&\quad \leqslant -(d_k - d_k^*)^T B_k(d_k - d_k^*)
\end{aligned} \tag{23}$$

because h is convex. We partition

$$d_k - d_k^* =: a_k^\perp + a_k^\parallel, \qquad a_k^\perp \in \operatorname{im}(f'(x_k)^T R), \qquad R^T f'(x_k) a_k^\parallel = 0.$$

Note that $a_k^\parallel \to 0$ and $a_k^\perp \to 0$ because $d_k \to 0$ from Lemma 5.1 and $d_k^* \to 0$ from (22). From (20), we have

$$\Phi_k'(d_k^*;\, d_k - d_k^*) \geqslant \tfrac{1}{2}\epsilon' \|a_k^\perp\|_2. \tag{24}$$

We subtract (24) from (23) and replace from (21):

$$\begin{aligned}
\Phi_k'(d_k;\, d_k^* - d_k) &\leqslant -\tfrac{1}{2}\epsilon'\|a_k^\perp\|_2 - (a_k^\perp + a_k^\parallel)^T B_k (a_k^\perp + a_k^\parallel) \\
&\leqslant \left(-\tfrac{1}{2}\epsilon' + \|B_k\|_2\|a_k^\perp\|_2 + 2\|B_k\|_2\|a_k^\parallel\|_2\right)\|a_k^\perp\|_2 - (a_k^\parallel)^T B_k a_k^\parallel \\
&\leqslant -\tfrac{1}{4}\epsilon'\|a_k^\perp\|_2 - \eta\|a_k^\parallel\|_2^2 < 0
\end{aligned}$$

for large k. Thus (II) holds and the theorem is proved. □

Since for large k, δ_k and d_k^* are in $\mathscr{E}(\Phi_k)$, the same holds for d_k if alternative (I) or (III) becomes true. However, there seems to be some hope that $d_k \in \mathscr{E}(\Phi_k)$ also in case (II), because δ_k is within the trust region whenever step 3b of Algorithm 3.1 is performed. This is not entirely correct, but close to:

Lemma 6.3. *Let $e_k \in \mathbb{R}^n$ be of minimum Euclidean norm satisfying*

$$d_k + e_k \in \mathscr{E}(\Phi_k), \quad \text{i.e.} \quad R^T(f(x_k) + f'(x_k)(d_k + e_k) - p^*) = 0.$$

Then $\|e_k\|_2 = O(\|x_k - x^*\|_2^2)$.

Proof. Note first that δ_k, d_k^* and d_k are of order $O(\|x_k - x^*\|_2)$ from (13), (22), and Theorem 6.2. If $\|\delta_k\| > \Delta_k$, then $d_k = \delta_k$ and $e_k = 0$, so there is nothing more to prove. Otherwise, we have $\Phi_k'(d_k;\, \delta_k - d_k) \geqslant 0$, because d_k is optimal and δ_k is feasible for (2). From (17),

$$\phi_{x_k}'(\delta_k;\, d_k - \delta_k) \geqslant \tfrac{1}{2}\epsilon'\|e_k\|_2. \tag{25}$$

For suitable constants $c_1, c_2 > 0$ we have

$$\phi_{x_k}'(\delta_k;\, d_k - \delta_k) - \Phi_k'(\delta_k;\, d_k - \delta_k) = \delta_k^T(\sigma I - B_k)(d_k - \delta_k) \leqslant c_1 \cdot \|x_k - x^*\|_2^2 \tag{26}$$

and

$$\Phi_k'(d_k;\ \delta_k - d_k) + \Phi_k'(\delta_k;\ d_k - \delta_k) \leqslant -(d_k - \delta_k)^T B_k (d_k - \delta_k) \leqslant c_2 \cdot \|x_k - x^*\|_2^2 \tag{27}$$

like in (23). Adding (26) to (27) and subtracting (25) yields

$$0 \leqslant \Phi_k'(d_k;\ \delta_k - d_k) \leqslant -\tfrac{1}{2}\epsilon' \|e_k\|_2 + (c_1 + c_2)\|x_k - x^*\|_2^2,$$

so $\|e_k\|_2 = O(\|x_k - x^*\|_2^2)$ is proved. □

Now we are able to prove our main result:

Theorem 6.4. *Let h be polyhedral, assume* (A1)–(A6) *are true. Then* $d_k = d_k^*$ *and* $q_k \geqslant \rho$ *hold for all sufficiently large k, and the sequence* $\{x_k\}$ *converges to* x^* *superlinearly.*

Proof. Assume for the moment we had already proved that $q_k \geqslant \rho$ for all sufficiently large k. Then we can show that $d_k = d_k^*$ holds from some k_0 on: For such k alternative (II) or (III) from Theorem 6.2 hold, step 6 of the algorithm implies $\Delta_{k+1} \geqslant \Delta_k$. Since $\delta_k \to 0$ and $d_k^* \to 0$, after a finite number of iterations (I) must hold. If, however, (I) holds for some large enough k, it will hold for $k+1$, too: Both $\|d_{k+1}^*\|$ and $\|\delta_{k+1}\|$ will be of magnitude $O(\|x_{k+1} - x^*\|_2)$, while $\Delta_{k+1} \geqslant \|d_k^*\|$, which is much larger because of (22). Thus, neither (II) nor (III) can hold, and $d_k = d_k^*$ will hold for all following k. Now (22) provides the bound

$$\|x_{k+1} - x^*\|_2 = O(\|x_k - x^*\|_2 \cdot (\|x_k - x^*\|_2 + \|B_k - W^*\|_2)), \tag{28}$$

so the proof will be complete if we show that $q_k \geqslant \rho$ for large k.

From Lemmas 4.2 and 6.1,

$$1 - q_k = \frac{\psi(x_k + d_k) - \Phi_k(d_k)}{\psi(x_k) - \Phi_k(d_k)} \leqslant \frac{\psi(x_k + d_k) - \Phi_k(d_k)}{\tfrac{1}{2}\sigma c^2 \|x_k - x^*\|_2^2} \tag{29}$$

for large k. Let $z_k \in \mathbb{R}^n$ be of minimum norm satisfying $z_k \in \mathscr{E}(\Phi_{x_k + d_k})$, i.e.

$$R^T(f(x_k + d_k) + f'(x_k + d_k) z_k - p^*) = 0.$$

Remembering $\|d_k\|_2 = O(\|x_k - x^*\|_2)$, which came from Theorem 6.2, we derive from Lemma 6.3

$$0 = R^T(f(x_k) + f'(x_k)(d_k + e_k) - p^*) = R^T(f(x_k - d_k) - p^*) + O(\|x_k - x^*\|_2^2),$$

so $\|z_k\|_2 = O(\|x_k - x^*\|_2^2)$. Lemma 2.1 gives the bound

$$\begin{aligned}\psi(x_k+d_k) &= \phi_{x_k+d_k}(\delta(x_k+d_k)) + \tfrac{1}{2}\sigma\|\delta(x_k+d_k)\|_2^2 \\ &\leqslant \phi_{x_k+d_k}(z_k) - \tfrac{1}{2}\sigma\|z_k-\delta(x_k+d_k)\|_2^2 + \tfrac{1}{2}\sigma\|\delta(x_k+d_k)\|_2^2 \\ &= h(f(x_k+d_k)+f'(x_k+d_k)z_k) + \sigma z_k^T\delta(x_k+d_k) \\ &= h(p^*) + \pi^{*T}(f(x_k+d_k)+f'(x_k+d_k)z_k-p^*) + \sigma z_k^T\delta(x_k+d_k)\ ,\end{aligned}$$

because $h(p) = h(p^*) + \pi^{*T}(p-p^*)$ for all $p \in \mathscr{E}(h)$. A lower bound for $\Phi_k(d_k)$ is

$$\begin{aligned}\Phi_k(d_k) &= h(f(x_k)+f'(x_k)d_k) + \tfrac{1}{2}d_k^TB_kd_k \\ &\geqslant h(p^*) + \pi^{*T}(f(x_k)+f'(x_k)d_k-p^*) + \tfrac{1}{2}d_k^TB_kd_k\,.\end{aligned}$$

Thus the numerator in (29) can be estimated by

$$\begin{aligned}&\psi(x_k+d_k) - \Phi_k(d_k) \\ &\quad\leqslant \pi^{*T}(f(x_k+d_k)+f'(x_k+d_k)z_k-f(x_k)-f'(x_k)d_k) \\ &\quad\quad + \sigma z_k^T\delta(x_k+d_k) - \tfrac{1}{2}d_k^TB_kd_k \\ &\quad= \sum_{i=1}^m \pi_i^*(f_i(x_k+d_k)-f_i(x_k) - \nabla f_i(x_k)^Td_k - \tfrac{1}{2}d_k^T\nabla^2 f_i(x_k)d_k) \\ &\quad\quad + \pi^{*T}f'(x_k+d_k)z_k + \sigma z_k^T\delta(x_k+d_k) - \tfrac{1}{2}d_k^T(B_k-W_k)d_k \\ &\quad= o(\|x_k-x^*\|_2^2)\end{aligned}$$

observing $\|d_k\|_2 = O(\|x_k-x^*\|_2)$, $\|z_k\|_2 = O(\|x_k-x^*\|_2^2)$, the Lipschitz condition on $\nabla^2 f$, $\pi^{*T}f'(x^*)=0$, and the fact that the matrices B_k as well as

$$W_k := \sum_{i=1}^m \pi_i^*\,\nabla^2 f_i(x_k)$$

converge to W^*. Substitution in (29) gives $\liminf_{k\to\infty} q_k \geqslant 1$, so $q_k \geqslant \rho$ for large k, and the theorem is proved. □

Corollary 6.5. *If the matrices B_k are chosen according to*

$$B_k := \sum_{i=1}^m \pi_i^{(k)}\,\nabla^2 f_i(x_k)\ ,$$

where $\pi^{(k)}$ denotes the multiplier to δ_k defined in (6), then the order of convergence $x_k \to x^*$ is at least quadratic.

Proof. (16) implies $\|\pi^{(k)}-\pi^*\|_2 = O(\|x_k-x^*\|_2)$. Hence $\|B_k-W^*\|_2 = O(\|x_k-x^*\|_2)$, and second order convergence follows from (28). □

An alternative choice for B_k, which turned out to be slightly superior in practical tests, is

$$B_k := \sum_{i=1}^{m} \gamma_i^{(k)} \nabla^2 f_i(x_k)$$

with the multipliers $\gamma_i^{(k)}$ to d_k as a solution of (2). Like in [14], we cannot show that $\gamma^{(k)} \to \pi^*$; however, the opposite case seems negligible for practice. Under the assumption $\gamma^{(k)} \to \pi^*$, one can derive

$$\|B_{k+1} - W^*\|_2 = O(\|x_k - x^*\|_2 \cdot (\|x_k - x^*\|_2 + \|B_k - W^*\|_2)),$$

which implies second order convergence $(x_k, B_k) \to (x^*, W^*)$.

7. Discussion

The algorithm was tested with a variety of problems including Yuan's [12] counterexample for only linear convergence of Fletcher's basic algorithm, discrete nonlinear approximation problems, and some hand-made polynomials of size up to 50. The matrices B_k were chosen as mentioned at the end of the previous section, using exact second order derivatives. The tests yielded no significant differences between Algorithm 3.1 and Fletcher's algorithm with correction step, comparing the number of steps until a certain neighbourhood of a solution was reached. Hence the more interesting point are the expenses per step. Table 1 shows the number of evaluations of f and f' and the number of subproblems to be solved per step of Fletcher's basic algorithm [6], Fletcher's algorithm with correction step [7], Algorithm 3.1, and a modification of Algorithm 3.1, which is proposed below. There is a decision made between accepted and rejected steps. Note that about 80–90% of all steps are accepted steps on average.

First regard the number of function evaluations. Fletcher's algorithm needs one more evaluation of f whenever a correction step is performed, while in Algorithm 3.1, f' is evaluated also in rejected steps for the calculation of $\delta(x_k+d_k)$ and $\psi(x_k+d_k)$. If one is willing to avoid this evaluation at the price of solving another additional subproblem, the following modification is possible: In step 5 of Algorithm 3.1, first calculate an estimate $\tilde{q}_k$ to q_k, using $f'(x_k)$ instead of $f'(x_k+d_k)$ for the calculation of $\psi(x_k+d_k)$. If $\tilde{q}_k < \rho$, reject

Table 1
Expenses per step of several algorithms

Algorithm	Accepted step			Rejected step		
	subprob.	eval. f	eval. f'	subprob.	eval. f	eval. f'
Fletcher [6]	1	1	1	1	1	0
Fletcher [7]	1-2	1-2	1	2	2	0
Algorithm 3.1	2	1	1	2	1	1
Modification	3	1	1	2	1	0

d_k and go to step 6b, otherwise continue with step 5 and 6, evaluating $f'(x_k+d_k)$ and q_k exactly and testing $q_k \geqslant \rho$ again. Numerical tests showed that the estimated test and the exact test seldom yield different results. The expenses per step are shown in the last line of Table 1. The analysis of Sections 4 to 6 can be adapted.

Beside these function evaluations, most of the time is spent solving the subproblems (2), (3) and (4) for d_k, $\tilde{d}_k$, and δ_k. In most applications, h is a polyhedral convex function and $\|\cdot\|$ is a polyhedral norm, frequently the maximum norm; so the subproblems can be solved via quadratic programming. The number of subproblems to be solved per step, as noted in Table 1, is not too significant, because the processing time needed to solve one subproblem varies quite much and it is hard to give exact proportions. In the following, we assume that a primal quadratic programming procedure with some active set strategy is used (see e.g. [8]).

Calculating δ_k is less expensive than calculating d_k, because the matrix of the corresponding quadratic program is essentially the identity matrix. A second reason is that $\sigma > \|B_k\|_2$ causes δ_k to be considerably smaller than d_k in most steps, so fewer changes of the active set may be necessary to solve the quadratic program.

Another important point is what starting points are available for the quadratic programming procedure. The second order correction $\tilde{d}_k$, calculated by (3), is often found in one step when starting with the active set from the solution of (2), in which case the costs can almost be neglected. The best starting point for (2) is harder to state. In the end phase of the algorithm, starting with the active set from the last step is surely favourable. In Algorithm 3.1, one can alternatively start in the point δ_k from the solution of (4). By this choice, problems arising from nonconvexity of Φ_k are excluded: The quadratic programming procedure will find a local solution d_k satisfying $\Phi_k(d_k) \leqslant \Phi_k(\delta_k)$, which is sufficient for the analysis given in Sections 4 and 5 (in particular, if one defines $d_k := \delta_k$ in every step, the results of Sections 4 and 5 remain valid).

The parameter σ in Algorithm 3.1 is normally changed at most once or twice until it attains its final value. In Table 1, we neglected the additional subproblem that must be solved when σ is changed. If there is no good estimate for σ available, it is preferable to start with a small value, because a too big value is not reduced by the algorithm and may slow down convergence.

Acknowledgement

I wish to thank Professor J. Werner for many helpful suggestions and his constant guidance and encouragement. I further am grateful to a referee, whose detailed suggestions led to improvement.

References

[1] T. Bannert, "Ein Trust-Region-Verfahren basierend auf der Glättung nichtdifferenzierbarer Funktionen," Ph.D. Thesis, Georg-August-Universität Göttingen (Göttingen, 1993).

[2] F.H. Clarke, *Optimization and Nonsmooth Analysis* (John Wiley & Sons, New York, 1983).
[3] R. Fletcher, "A class of methods for nonlinear programming with termination and convergence properties," in: J. Abadie, ed., *Integer and Nonlinear Programming* (North-Holland, Amsterdam–London, 1970) pp. 157–175.
[4] R. Fletcher, "A class of nonlinear programming III: Rates of convergence," in: F. Lootsma, ed., *Numerical Methods for Nonlinear Optimization* (Academic Press, London–New York, 1972) pp. 371–381.
[5] R. Fletcher, "An exact penalty function for nonlinear programming with inequalities," *Mathematical Programming* 5 (1973) 129–150.
[6] R. Fletcher, "A model algorithm for composite nondifferentiable optimization problems," *Mathematical Programming Study* 17 (1982) 67–76.
[7] R. Fletcher, "Second order corrections for non-differentiable optimization," in: G.A. Watson, ed., *Numerical Analysis Proceedings* (Springer, Berlin, 1982) pp. 85–114.
[8] R. Fletcher, *Practical Methods of Optimization* (John Wiley & Sons, New York, 1987).
[9] K. Madsen, "An algorithm for minimax-solution of overdetermined systems of nonlinear equations," *Journal of the Institute of Mathematics and its Applications* 16 (1975) 321–328.
[10] M.J.D. Powell and Y. Yuan, "A trust region algorithm for equality constrained optimization," *Mathematical Programming* 49 (1990) 189–211.
[11] R.S. Womersley, "Local properties of algorithms for minimizing nonsmooth composite functions," *Mathematical Programming* 32 (1985) 69–89.
[12] Y. Yuan, "An example for only linear convergence of trust region algorithms for non-smooth optimization," *IMA Journal of Numerical Analysis* 4 (1984) 327–335.
[13] Y. Yuan, "Conditions for convergence of trust region algorithms for nonsmooth optimization," *Mathematical Programming* 31 (1985) 220–228.
[14] Y. Yuan, "On the superlinear convergence of a trust region algorithm for non-smooth optimization," *Mathematical Programming* 31 (1985) 269–285.