American Journal of Applied Mathematics
2018; 6(2): 71-77
http://www.sciencepublishinggroup.com/j/ajam
doi: 10.11648/j.ajam.20180602.17
ISSN: 2330-0043 (Print); ISSN: 2330-006X (Online)

A Nonmonotone Trust Region Method for Nonsmooth Composite Programming Problems

Min Xi[*], Ailing Xiao

School of Finance, Guangdong University of Foreign Studies, Guangzhou, China

Email address:
sheamee@hotmail.com (Min Xi), alxiao181@126.com (Ailing Xiao)
[*]Corresponding author

To cite this article:
Min Xi, Ailing Xiao. A Nonmonotone Trust Region Method for Nonsmooth Composite Programming Problems. *American Journal of Applied Mathematics*. Vol. 6, No. 2, 2018, pp. 71-77. doi: 10.11648/j.ajam.20180602.17

Received: March 20, 2018; **Accepted**: April 24, 2018; **Published**: June 26, 2018

Abstract: A class of nonsmooth composite minimization problems are considered in this paper. In practice, many problems in computational science, engineering and other enormous areas fall into this class of problems. Two obvious applications of this class of problems are least square problems with nonsmooth data and the problem of solving a system of nonsmooth equations. Because of its practical importance, this class of problem has received wide attention from the mathematical optimization community. In particular, Sampaio, Yuan and Sun has proposed a trust region method for this class of problems and also provided the convergence analysis to support their algorithm. Motivated partly by their work, in this paper a nonmonotone trust region algorithm for this class of nonsmooth composite minimization problems is presented. Different from most existing monotone line search and trust region methods, this method combines the nonmonotone technique to improve the efficiency of the trust region method. After a brief introduction of the class of problems in the first section, some fundamental concepts and properties which will be used in this paper are presented. Then, the new nonmonotone trust region algorithm for the class of problem is described followed by the global convergence analysis of the new algorithm. A simple application of this algorithm is discussed in the last part of this paper.

Keywords: Nonmonotone, Trust Region Method, Composite, Nonsmooth Optimization

1. Introduction

In this paper, a class of nonsmooth composite minimization problems are considered. More formally, the problems are

$$\text{(P)}\ \ \min\ \ h(f(x)), \tag{1}$$

where $f: R^n \to R^n$ is a locally Lipschitzian function, and $h: R^n \to R$ is a continuously differentiable convex function bounded below. This class of problems arises frequently in computational science, engineering and other enormous areas. For instance, the problem of solving a system of nonsmooth equations and the least squares problem with nonsmooth data are all special cases of (P).

The research on nonsmooth optimization was originated in the work of Clarke [2], and then there has been a number of research in the area [5, 6, 10, 12, 16]. In [5], Fletcher considered trust region methods for a class of composite nondifferentiable optimization problems

$$\min\ \ F(x) = g(x) + h(f(x)),$$

where $g: R^n \to R$ and $f: R^n \to R^n$ are continuously differentiable functions and $h: R^n \to R$ is a convex but nonsmooth function bounded below. Qi and Sun [10] extended the classical trust region algorithm to the nonsmooth case where the objective function is only locally Lipschitzian. The convergence result of this work extends the result of Powell for minimization of smooth functions, the result of Yuan for minimization of composite convex functions. However, many optimization problems in practice come down to the class of problems (P) and it has received extensive attention from the mathematical optimization community. In particular, a trust region method for (P) was proposed by Sampaio, Yuan and Sun in [12]. This work also provided the convergence analysis of the method and extended the theory of trust region method for nonsmooth optimization given by Fletcher, Powell and Yuan. Motivated partly by this work [12], a nonmonotone trust region method for this class of problems (P) is proposed in this paper.

To authors' knowledge, most existing algorithms for nonsmooth optimization problems are monotone algorithms that enforce the sequence of objective function values decrease monotonically. However, some researches indicated that, enforcing monotonically decrease at each iteration can considerably slow down the convergence rate in the minimization process, especially in the presence of narrow curved valley. If the iteration is trapped near the narrow curved valley, forcing the value of objective function to decrease at each iteration would result in very short steps or zigzagging. In other words, monotonicity can cause serious loss of efficiency.

Differing from the monotone methods, the nonmonotone methods relax the monotone limitation and allow the sequence of objective function values to increase occasionally. The idea of constructing nonmonotone method was originally developed by Chamberlain *et al.* [1] in the context of constrained optimization to overcome the so-called Maratos effect, which could lead to the rejection of superlinear convergent steps since these steps could cause an increase in both the objective function value and the constraint violation. In 1980s, Grippo *et al.* [6] first applied the nonmonotone line search technique for Newton's method. Numerical results in this work indicate that the nonmonotone line search technique may allow a considerable saving of computation both in the number of line search and function evaluation. Since then, various nonmonotone algorithms have been presented and proved to be efficient and competitive in practice and theory. In 1993, Deng et al. [4] extended the nonmonotone technique from line search methods to trust region methods. Then Toint [13, 14] introduced some variants of nonmonotone techniques combined with line search and trust region algorithms respectively.

In this paper, the nonmonotone technique is combined with the trust region methods to solve the class of nonsmooth composite minimization problems (P). And it is proved that the algorithm is globally convergent. The remainder of this paper is organized as follows. In section 2, some fundamental concepts and properties which will be used in this paper are presented and the framework of trust region method is also introduced. In section 3, the new nonmonotone trust region algorithm for problem (P) is proposed. The global convergence of the new algorithm is established in section 4. Before the conclusion part, a simple application of the new algorithm is discussed in section 5.

2. Preliminaries

Throughout this paper, the notation $\|\cdot\|$ is used to denote the 2-norm for a vector and the induced 2-norm for a matrix. All functions are assumed to be locally Lipschitzian in this paper. The following are some basic definitions and properties that will be used in this paper and they can be found in [2].

Definition 1. Let $f: R^n \to R$, and $x \in R^n$ be such that $f(x)$ is finite, the lower and upper Dini directional derivatives of f at x in the direction $v \in R^n$ are given respectively as

$$f_d^- = \liminf_{t\downarrow 0}[f(x+tv)-f(x)]/t,$$

$$f_d^+ = \limsup_{t\downarrow 0}[f(x+tv)-f(x)]/t.$$

Definition 2. The Clarke directional derivative $f^0(x,v)$ of f at x in the direction $v \in R^n$ is defined by

$$f^0(x,v) = \lim_{x'\to x}\sup_{t\downarrow 0}[f(x'+tv)-f(x')]/t.$$

The Clarke sub-differential $\partial f(x)$ of f at x is defined by

$$\partial f(x) = \{\xi \in R^n \mid \xi^T v \le f^0(x,v), \forall v \in R^n\}.$$

Definition 3. The function $f: R^n \to R$ is said to be lower semicontinuous at x_0, if

$$f(x_0) \le \liminf_{x\to x_0} f(x).$$

The function $f: R^n \to R$ is said to be upper semicontinuous at x_0, if

$$f(x_0) \ge \limsup_{x\to x_0} f(x).$$

Lemma 1. Let $f: R^n \to R$ be locally Lipschitzian function and $x \in R^n$. Then

(1) $f^0(x,v) = \max\{\xi^T d \mid \xi \in \partial f(x), \forall d \in R^n\}$.

(2) $\partial f(x)$ is a compact set and $\partial f(x) \ne \varnothing$.

(3) $f^0(x,\cdot)$ is positively homogeneous and sublinear.

Definition 4. Let $F = (f_1, f_2, \ \ldots, \ f_m)^T : R^n \to R^m$ be continuously differential at $x \in R^n$, then the Jacobi matrix $J(x) \in R^{m\times n}$ is defined as

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

$\nabla F(x) = J(x)^T$ is said to be the gradient of F at x.

Definition 5. Let $F = (f_1, f_2, \ \ldots, \ f_m)^T : R^n \to R^m$, and each $f_i, i = 1,2,...,m$ be locally Lipschitzian continuous function, then the generalized Jacobi matrix $J(x) \in R^{m\times n}$ is defined as

$$\partial F(x)^T = \{Z \in R^{m\times n} \mid J(x) = \lim_{x_k \to x} J(x_k)\}.$$

Before proposing the new nonmonotone trust region method in the next section, the framework of the trust region method will be briefly described first. A traditional trust region algorithm works as follows. At each iterate x_k, a model function $m_k(x)$ which is much easier to handle than the objective function, say $f(x)$, is defined to approximate $f(x)$ within a suitable neighborhood of x_k. The neighborhood is referred to as the trust region and is often represented by a ball centered at the current iterate x_k with radius Δ_k. Then solve a trust region sub-problem, i.e., minimizing the model $m_k(x)$ within a trust region ball to obtain a trial step d_k. If the trial step yields a good decrease in the objective function relative to the decrease in the model and the trust region radius Δ_k is sufficiently small relative to the size of the model gradient, then the step is taken and Δ_k is possibly increased. Otherwise the step is rejected and the trust region radius Δ_k is decreased. Then solve the sub-problem within a smaller trust region to get a new trial step. The ratio of the objective function reduction to its model function reduction, i.e.,

$$r_k = \frac{f(x_k) - f(x_k + d_k)}{m_k(x_k) - m_k(x_k + d_k)} \tag{2}$$

is computed to decide the acceptance of the trial iterate $x_k + d_k$. Due to the strong global convergent properties and robustness, the trust region methods now is an important ingredient of many standard optimization textbooks [8, 17].

3. The Nonmonotone Trust Region Algorithm

In this section the new nonmonotone trust region algorithm for the class of nonsmooth composite minimizing problems (P) is described in detail.

First, some parameters and constants need to be explained. $\Delta_0, c_0, c_1, c_2, \eta_1$ and η_2 are given positive constants satisfying $c_0 \le 1, c_2 < c_1 < 1, 0 < \eta_1 < 1 < \eta_2$.

At the k th iteration, given x_k, H_k and Δ_k, approximate the objective function $h(f(x))$ by the following model function:

$$Q_k(d) =: h(f(x_k) + Z_k{}^T d) + \frac{1}{2} d^T H_k d, \tag{3}$$

where $Z_k \in \partial f(x_k)$ is the generalized Jacobian of f at x_k and $H_k \in R^{n\times n}$ is a given symmetric matrix that carry certain second-order information of $h(f(x_k))$.

Then solve the following sub-problem:

$$\text{(SP)} \quad \min_{\|d\| \le \Delta_k} Q_k(d) =: h(f(x_k) + Z_k{}^T d) + \frac{1}{2} d^T B_k d, \tag{4}$$

to gain a trial step. An exact solution of (4) is too expensive and unnecessary, the new algorithm will be satisfied with an inexact solution. Assume d_k^* is the exact solution of (4), then an inexact solution d_k satisfies

$$h(f(x_k)) - Q_k(d_k) \ge c_0[h(f(x_k)) - Q_k(d_k^*)] \tag{5}$$

and

$$\|d_k\| \le \Delta_k$$

will be accepted by the new algorithm. If $d_k = 0$, then stop. Otherwise, r_k is computed to decide the acceptance of the trial iterate $x_k + d_k$.

As mentioned above, to improve the efficiency of the trust region methods, a nonmonotone technique is applied into the framework of the trust region methods. The ratio r_k used in nonmonotone trust region methods is different from the classical one. The objective function value at current iterate in r_k is replaced by the reference function value whose definition follows [6].

For each k, let m_k satisfy

$$m_0 = 0 \text{ and } 0 \le m_k \le \min\{m_{k-1} + 1, M\},$$

where $k \ge 1$ and $M \ge 0$ is a given integer parameter. Then define the reference function value as follows:

$$h(f(x_{l(k)})) = \max_{0 \le j \le m_k} \left[h(f(x_{k-j})) \right],$$

where $k - m_k \le l(k) \le k.$

The ratio r_k is defined as:

$$r_k = \frac{h(f(x_{l(k)})) - h(f(x_k + d_k))}{Q_{l(k)}(0) - Q_k(d_k)}. \tag{6}$$

Next, determine whether accept $x_k + d_k$ as the new iterate and update the trust region radius Δ_k:

$$x_{k+1} = \begin{cases} x_k + d_k, & r_k > c_2 \\ x_k, & \textit{otherwise}, \end{cases}$$

$$\Delta_{k+1} = \begin{cases} \eta_1 \Delta_k, & r_k \le c_2, \\ \Delta_k, & c_2 < r_k \le c_1, \\ \min\{\eta_2 \Delta_k, \Delta_0\}, & \textit{otherwise}. \end{cases}$$

Denote

$$\phi(x,d) =: h(f(x)+Z_k^T d) - h(f(x)), \tag{7}$$

$$\psi(x,\Delta) =: \sup\{-\phi(x,d) : \|d\| \le \Delta\}. \tag{8}$$

If for some $\Delta > 0, \psi(x,\Delta) = 0$, then x is said to be a critical point of (P), and the algorithm stop.

Now, the nonmonotone trust region algorithm for the nonsmooth composite minimization problem is presented below.

Algorithm NTR (The Nonmonotone Trust Region Algorithm)

Step 0. (Initialization):

Given $x_0 \in R^n$, $H_0 \in R^{n\times n}$, $\varepsilon > 0, \Delta_0 > 0$, $0 < c_0 \le 1$, $0 < c_2 < c_1 < 1$, $0 < \eta_1 < 1 < \eta_2$ and a nonnegative integer M. Set $k = 0, m_0 = 0$ and compute $h(f(x_0))$.

Step 1. (Model definition):

Compute $\psi(x_k,\Delta_k)$ as (8), if $|\psi(x_k,\Delta_k)| \le \varepsilon$, set $x^* = x_k$ and stop. Otherwise compute $H_k \in R^{n\times n}$ and define the model function $Q_k(d)$ by (3).

Step 2. (Step calculation):

Determine a trial step d_k by solving the trust region sub-problem (4) with the trust region radius Δ_k.

Step 3. (Acceptance of the trial step):

Compute r_k as (6).

If $r_k \le c_2$, then $x_{k+1} = x_k$ and $\Delta_{k+1} = \eta_1 \Delta_k$, update to H_{k+1} and go to step 2.

Otherwise, set $x_{k+1} = x_k + d_k$, update to H_{k+1} and compute $h(f(x_{k+1}))$. If $c_2 < r_k \le c_1$, $\Delta_{k+1} = \Delta_k$. Otherwise, $\Delta_{k+1} = \min\{\eta_2 \Delta_k, \Delta_0\}$.

Step 4. Set $k := k+1$ and go to step 1.

Remark 1. It is easy to see that when $M = 0$, Algorithm NTR is reduced to the monotone trust region method. When $M > 0$, Algorithm NTR has the nonmonotone properties which can relax the monotone limitation and improve the efficiency of the trust region methods.

4. The Global Convergence of Algorithm

In this section, the global convergence results of Algorithm NTR given in the previous section is established. This analysis follows the convergence analysis framework of the trust region methods ([3, 8, 17]). But it is for nonsmooth composite programming problem and is based on the nonmonotone technique.

Let $F = h \circ f$, some standard assumptions which can be found in [10] are given as follows.

Assumption 1. H_k is uniformly bounded, i.e., there exists a constant $\kappa > 0$ such that, for all k,

$$H_k \le \kappa$$

Assumption 2. The level set

$$L(x_0) = \{x \in R^n \mid F(x) \le F(x_0)\}$$

is bounded and compact, where x_0 is some starting point in R^n. Let $D \subset R^n$ be a bounded open convex set containing $L(x_0)$ and Δ_0 be the diameter of D.

Before proving the main convergence theorem, some relevant propositions and lemmas are given.

Proposition 1. *Let F, ϕ be as above, then*

(1) For all $x \in R^n, \phi(x,0) = 0$ and $\phi(x,\cdot)$ is lower semicontinuous.

(2) There exists $\bar{\Delta} > 0$ such that for all $\|d\| \le \bar{\Delta}$, $-\phi(\cdot,d)$ is lower semicontinuous.

(3) $\phi(x,\alpha d) \le \alpha\phi(x,d)$, $\forall x \in L_0, 0 \le \alpha \le 1$.

(4) For any convergent subsequence $\{x_k : k \in K \subseteq J\}$, if $d_k \to 0$, then

$$f(x_k + d_k) - f(x_k) \le \phi(x_k, d_k) + o(\|d\|). \tag{9}$$

Where $J = \{0,1,2,3,...\}$.

Lemma 2. Let $F = h \circ f$. For all $\Delta \ge \Delta_k$,

$$F(x_k) - Q_k(d_k) \ge \frac{c_0}{2\Delta}\psi(x_k,\Delta)\min\left\{\Delta_k, \frac{\psi(x_k,\Delta)}{\|H_k\|\Delta}\right\},$$

where the second term in the min notation is understood as ∞ if $H_k = 0$.

Proposition 1 and Lemma 2 have been proved true in [10].

Lemma 3. For all $\Delta \ge \Delta_k$,

$$F(x_{l(k)}) - Q_k(d_k) \ge \frac{c_0}{2\Delta}\psi(x_k,\Delta)\min\left\{\Delta_k, \frac{\psi(x_k,\Delta)}{\|H_k\|\Delta}\right\},$$

where the second term in the min notation is understood as ∞ if $H_k = 0$.

This lemma can be deduced directly from the Lemma 3.2 in [10].

Theorem 1. Under assumption 1-2, at least one accumulation point $\bar{x}$ of sequence $\{x_k\}$ which generated by Algorithm NTR is a critical point of F.

Proof. Suppose to the contrary that any accumulation point $\bar{x}$ of sequence $\{x_k\}$ is not the critical point of F. That is to say for any $\Delta > 0$, there exists positive number $\varepsilon_0, \varepsilon_1 > 0$ and $N > 0$, such that for all $k > N$, $\|x_k - \bar{x}\| \le \varepsilon_1$,

$$\psi(x_k,\Delta) \ge \varepsilon_0. \tag{10}$$

In fact, if for positive number $\varepsilon_n = 1/n > 0$, some $\Delta_1 > 0$ and any $N > 0$, $\exists k > N$ satisfying $x'_k \in \{x_k\}$ and $\psi(x'_k,\Delta_1) < \varepsilon_n$. Then it is deduced that

$$\liminf_{k\to\infty} \psi(x'_k,\Delta_1) \le 0.$$

On the other hand, since $L(x_0)$ is a compact set, then there is a convergent subsequence $\{x'_{i_t}\}$ of $\{x_k\}$ such that $x'_{i_t} \to x'$, i.e., x' is an accumulation point of $\{x_k\}$. From assumption, x' is not a critical point of $\{x_k\}$, hence $\psi(x', \Delta_1) > 0$, and

$$\liminf_{k\to\infty} \psi(x'_k, \Delta_1) < \psi(x', \Delta_1),$$

which contradicts the lower semi-continuity of $\psi(\cdot, \Delta_1)$ at x'.

According to Algorithm NTR, the sequence $\{x_k\}$ generated by Algorithm NTR has at least one subsequence satisfying one of the following two cases.

Case (a) There exists a subsequence $\{r_k : k \in K_1\}$, such that $r_k \le c_2$ and $\lim\limits_{k\in K_1, k\to\infty} \Delta_k = 0$.

Case (b) There exists a subsequence $\{r_k : k \in K_2\}$ and a positive integer K_0, such that for any $k > K_0$, $k \in K_2$, $r_k > c_2$.

Now, let $\bar{x}$ be an accumulation point of such a subsequence and an infinite subsequence $\{x_k : k \in K_3\}$ converges to $\bar{x}$, but $\bar{x}$ is not a critical point of F.

First, consider Case (a).

Since $\lim\limits_{k\in K_1, k\to\infty} \Delta_k = 0$, then

$$\lim_{k\in K_1, k\to\infty} \|d_k\| = 0.$$

From Lemma 3 and Assumption 1, for any sufficiently large $k \in K_3 \subseteq K_1, k > N$,

$$\frac{F(x_k + d_k) - Q_k(d_k)}{F(x_k) - Q_k(d_k)}$$

$$= \frac{F(x_k + d_k) - F(x_k) - \phi(x_k, d_k) - \frac{1}{2} d_k^T H_k d_k}{F(x_k) - Q_k(d_k)}$$

$$\le \frac{o(\|d_k\|) - \frac{1}{2} d_k^T H_k d_k}{\frac{c_0}{2\Delta_0} \psi(x_k, \Delta_0) \min\left\{\Delta_k, \frac{\psi(x_k, \Delta_0)}{\|H_k\|\Delta_0}\right\}}$$

$$\le \frac{o(\|d_k\|)}{\frac{c_0}{2\Delta_0} \varepsilon_0 \Delta_k} \le \frac{o(\|d_k\|)}{\frac{c_0}{2\Delta_0} \varepsilon_0 \|d_k\|}.$$

Let k be sufficiently large, it can be deduced that

$$\frac{F(x_k + d_k) - Q_k(d_k)}{F(x_k) - Q_k(d_k)} < 1 - c_2.$$

It follows that

$$(1 - c_2) F(x_k) > F(x_k + d_k) - c_2 Q_k(d_k)$$

Notice that $0 < c_2 < 1$ and $F(x_{l(k)}) > F(x_k)$, then

$$(1 - c_2) F(x_{l(k)}) > F(x_k + d_k) - c_2 Q_k(d_k).$$

Thus,

$$\frac{F(x_{l(k)}) - F(x_k + d_k)}{F(x_{l(k)}) - Q_k(d_k)} > c_2.$$

This is to say,

$$r_k > c_2$$

for $k \in K_3 \subseteq K_1$ large enough, which leads to a contradiction according to the assumption of Case (a).

Now, consider Case (b). Since $m_{k+1} \le m_k + 1$, it follows that

$$F(x_{l(k+1)}) = \max_{0\le j\le m_{k+1}} [F(x_{k+1-j})] \le \max_{0\le j\le m_k+1} [F(x_{k+1-j})]$$

$$= \max\{F(x_{l(k)}), F(x_{k+1})\} = F(x_{l(k)}).$$

Since the function F is bounded below, then the sequence $\{F(x_{l(k)})\}$ converges.

From the assumption of Case (b), for any $k > K_0$, $k \in K_3 \subseteq K_2$, it follows that $r_k > c_2$.

Let $\delta = \min\{\Delta_k \mid k \in K_3\}$, $\eta = \frac{c_0 c_2}{2\Delta_0} \varepsilon_0 \min\left\{\delta, \frac{\varepsilon_0}{\kappa\Delta_0}\right\}$. By Lemma 2-3 and Assumption 1, for all $k \in K_3$,

$$F(x_{l(k)}) - F(x_{k+1}) \ge \frac{c_0 c_2}{2\Delta_0} \psi(x_k, \Delta_0) \min\left\{\Delta_k, \frac{\psi(x_k, \Delta_0)}{\kappa\Delta_0}\right\}.$$

It follows from (10) that for all large $k \in K_3$, $k > \max\{N, K_0\}$,

$$F(x_{k+1}) \le F(x_{l(k)}) - \frac{c_0 c_2}{2\Delta_0} \psi(x_k, \Delta_0) \min\left\{\Delta_k, \frac{\psi(x_k, \Delta_0)}{\kappa\Delta_0}\right\}$$

$$\le F(x_{l(k)}) - \frac{c_0 c_2}{2\Delta_0} \varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{\kappa\Delta_0}\right\}$$

$$\le F(x_{l(k)}) - \frac{c_0 c_2}{2\Delta_0} \varepsilon_0 \min\left\{\delta, \frac{\varepsilon_0}{\kappa\Delta_0}\right\}$$

$$\le F(x_{l(k)}) - \eta. \tag{11}$$

Thus,

$$F(x_{l(k)}) \le F(x_{l(l(k)-1)}) - \eta. \tag{12}$$

Since the function F is bounded below and the sequence $\{F(x_{l(k)})\}$ is convergent, it follows from (12) that $\eta \le 0$, which contradicts to the fact that $\eta > 0$.

Combing the Case (a) and Case (b), the theorem holds.

5. Applications

The above algorithm and theory have two obvious applications: the least square problems with nonsmooth data and nonsmooth equations.

In this section, consider a least square problem

$$\min_{x \in R^n} \frac{1}{2}\|F(x)-b\|^2. \quad (13)$$

Take $h(\cdot) = 1/2\|\cdot\|^2$ and $f(\cdot) = F(\cdot) - b$, where $F(x)^T = (f_1(x), f_2(x), ...f_m(x))$ and $b^T = (b_1, b_2, ...b_m)$ are not necessarily smooth. Then (13) is a special case of problem (P) since h is a continuously differentiable convex function bounded below, and $f(\cdot) = F(\cdot) - b$ is locally Lipschitzian. The new nonmonotone trust region method to solve (13) can be applied. The algorithm framework is as follows:

Algorithm

Step 0. (Initialization):

Given $x_0 \in R^n$, $H_0 \in R^{n \times n}$, $\varepsilon > 0, \Delta_0 > 0$, $0 < c_0 \le 1$, $0 < c_2 < c_1 < 1$, $0 < \eta_1 < 1 < \eta_2$ and a nonnegative integer M. Set $k = 0, m_0 = 0$ and compute $\frac{1}{2}\|F(x_0) - b\|^2$.

Step 1. (Model definition):

Compute $\psi(x_k, \Delta_k)$ as (8), where

$$\phi(x_k, d) =: \frac{1}{2}\left\|F(x_k) - b + Z_k{}^T d\right\|^2 - \frac{1}{2}\|F(x_k) - b\|^2$$

If $|\psi(x_k, \Delta_k)| \le \varepsilon$, set $x^* = x_k$ and stop. Otherwise compute $H_k \in R^{n \times n}$ and define the model function

$$Q_k(d) = \frac{1}{2}\left\|F(x_k) - b + Z_k{}^T d\right\|^2 + \frac{1}{2} d^T H_k d$$

Step 2. (Step calculation):

Determine a trial step d_k by solving the trust region sub-problem (4) with the trust region radius Δ_k.

Step 3. (Acceptance of the trial step):

Compute

$$r_k = \frac{\|F(x_{l(k)}) - b\|^2 - \|F(x_k + d_k) - b\|^2}{Q_{l(k)}(0) - Q_k(d_k)}$$

If $r_k \le c_2$, then $x_{k+1} = x_k$ and $\Delta_{k+1} = \eta_1 \Delta_k$, update H_k and go to step 2. Otherwise, set $x_{k+1} = x_k + d_k$, update H_k and compute $h(f(x_{k+1}))$. If $c_2 < r_k \le c_1$, $\Delta_{k+1} = \Delta_k$. If $r_k > c_1$, $\Delta_{k+1} = \min\{\eta_2 \Delta_k, \Delta_0\}$.

Step 4. Set $k := k+1$ and go to step 1.

6. Conclusion

In this paper, a new nonmonotone trust region methods is presented for the class of nonsmooth composite minimization problems (P). The nonmonotone technique is combined with the trust region method to improve the algorithm's efficiency. Under the mild conditions, the global convergence property of the new algorithm is proved. For the future research, this idea can be extended to other optimization problems, such as constrained optimization.

Acknowledgements

This work was supported by a Project Funded by Guangdong University of Foreign Studies (Grant No. 17QN32).

References

[1] R. M. Chamberlain, M. J. D. Powell, C. Lemarechal and H. C. Pedersen, The watchdog technique for forcing convergence in algorithms for constrained optimization, Mathematical Programming Studies, 1982, vol. 16: pp. 1-17.

[2] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley, New York, 1983.

[3] A. R. Conn, N. I. M. Gould and P. L. Toint, Trust Region Methods. Philadelphia: SIAM, 2000.

[4] N. Y. Deng, Y. Xiao and F. J. Zhou, Nonmonotonic trust region algorithm, Journal of Optimization Theory and Applications, 1993, vol. 76, pp. 259-285.

[5] R. Fletcher, Practical Methods of Optimization, Volume 2, John Wiley and Sons, New York, 1981.

[6] L. Grippo, F. Lampariello and S Lucidi, A nonmonotone line search technique for Newton's method 1986, vol. 23: pp. 88.

[7] M. Kimiaei, A new class of nonmonotone adaptive trust-region methods for nonlinear equations with box constraints, Calcolo, 2016, vol. 54: pp. 1-44

[8] J. Nocedal and S. J. Wright, Numerical Optimization, 2nd edition, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006.

[9] J. S. Pang, S. P. Han, and N. Rangaraj, Minimization of locally Lipschitzian function, SIAM Journal on Optimization, 1991, vol. 1, pp. 57-82.

[10] L. Qi and J. Sun, A trust region algorithm for minimization of locally Lipschitzian functions, Mathematical Programming, 1994, vol. 66: pp. 25-43.

[11] S. Rezaee and S. Babaie-Kafaki, A modified nonmonotone trust region line search method, Journal of Applied Mathematics and Computing, 2017, https://doi.org/10.1007/s12190-017-1113-4.

[12] R. J. de Sampaio, J. Yuan and W. Sun, Trust region algorithm for nonsmooth optimization, Applied Mathematics and Computation, 1997, vol. 85: pp. 109-116.

[13] Ph. L. Toint, An assessment of nonmonotone line search techniques for unconstrained optimization, SIAM Journal on Scientific Computing, 1996, vol. 17: pp. 725-739.

[14] Ph. L. Toint, A non-monotone trust-region algorithms for nonlinear optimization subject to convex constraints, Mathematical Programming, 1997, vol. 17: pp. 69-94.

[15] X. Wang and Y. Yuan, An augmented Lagrangian trust region method for equality constrained optimization, Optimization Methods and Software, 2015, vol. 30: pp. 559-582.

[16] Y. Yuan, Conditions for convergence of trust region algorithms for nonsmooth optimization, Mathematical Programming, 1985, vol. 31: pp. 220-228.

[17] Y. Yuan and W. Sun, Optimization Theory and Methods, Science Press, Beijing, 1997.

[18] Q. Zhou, W. Sun and H. Zhang, A new simple model trust-region method with generalized Barzilai-Borwein parameter for large-scale optimization, Science China-mathematics, 2016, vol. 59: pp. 2265-2280.

[19] H. Zhu, Q. Ni, L. Zhang and W. Yang, A fractional trust region method for linear equality constrained optimization, Discrete Dynamics in Nature and Society, 2016, pp. 1-10.