

Research Article

A New Smoothing Nonlinear Conjugate Gradient Method for Nonsmooth Equations with Finitely Many Maximum Functions

Yuan-yuan Chen^{1,2} and Shou-qiang Du¹

¹ College of Mathematics, Qingdao University, Qingdao 266071, China

² School of Management, University of Shanghai for Science and Technology, Shanghai 200093, China

Correspondence should be addressed to Yuan-yuan Chen; usstchenyuan@163.com

Received 8 March 2013; Accepted 25 March 2013

Academic Editor: Yisheng Song

Copyright © 2013 Y.-y. Chen and S.-q. Du. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The nonlinear conjugate gradient method is of particular importance for solving unconstrained optimization. Finitely many maximum functions is a kind of very useful nonsmooth equations, which is very useful in the study of complementarity problems, constrained nonlinear programming problems, and many problems in engineering and mechanics. Smoothing methods for solving nonsmooth equations, complementarity problems, and stochastic complementarity problems have been studied for decades. In this paper, we present a new smoothing nonlinear conjugate gradient method for nonsmooth equations with finitely many maximum functions. The new method also guarantees that any accumulation point of the iterative points sequence, which is generated by the new method, is a Clarke stationary point of the merit function for nonsmooth equations with finitely many maximum functions.

1. Introduction

In this paper, we consider the following nonsmooth equations with finitely many maximum functions:

$$\begin{aligned} \max_{j \in J_1} \{h_{1j}(x)\} &= 0, \\ &\vdots \\ \max_{j \in J_n} \{h_{nj}(x)\} &= 0, \end{aligned} \quad (1)$$

where $h_{ij} : R^n \rightarrow R$ for $i = 1, \dots, n$ and $j \in J_i$, $i = 1, \dots, n$ are continuously differentiable functions, and J_i are finite index sets. Denote that

$$h_i(x) = \max_{j \in J_i} \{h_{ij}(x)\}, \quad i = 1, \dots, n, \quad (2)$$

$$H(x) = (h_1(x), \dots, h_n(x))^T,$$

then (1) can be reformulated as

$$H(x) = 0, \quad (3)$$

where $H : R^n \rightarrow R^n$ is nonsmooth equations and $J_i(x) = \{j_i \in N \mid h_{ij}(x) = h_i(x)\}$, $i = 1, \dots, n$. Nonsmooth equations have been studied for decades, which is proposed in the study of the optimal control, the variational inequality and complementarity problems, equilibrium problems, and engineering mechanics [1–3], because many practical problems, such as stochastic complementarity problems, variational inequality problems, KKT systems of constrained nonlinear programming problems, and many problems in equilibrium problems, can be reformulated into (1). In the past few years, there has been a growing interest in the study of (1) (such as [4, 5]). Due to its simplicity and global convergence, the iterative methods, such as nonsmooth Levenberg-Marquardt method, general Newton method, and smoothing methods for solving nonsmooth equations, have been widely studied [4–11].

In this paper, we give a new smoothing nonlinear conjugate gradient method for (1). In the following section, we will recall some definitions and some background of nonlinear conjugate gradient methods. And we also give the new smoothing nonlinear conjugate gradient method for (1), which guarantees that any accumulation point of the iterative points sequence is a Clarke stationary point of the merit

function for (1). In the last section, some discussions are also given.

Notation. In the following, a quantity with a subscript k denotes that quantity is evaluated at x_k , the norm is the 2 norm, and $R_{++} = \{t \mid t > 0\}$.

2. Preliminaries and New Method

In this section, firstly, we give some definitions and some backgrounds of nonlinear conjugate gradient method. Secondly, we propose the new methods for (1) and give the convergence analysis.

In the following, we give two definitions, which will be used in this paper.

Definition 1. Let $F : R^n \rightarrow R$ be a locally Lipschitz function. Then, F is almost everywhere differentiable. Denote D_F be the set of points where F is differentiable, then the general gradient of F at x in the sense of Clark is

$$\partial F(x) = \text{conv} \left\{ \lim_{x_k \rightarrow x, x_k \in D_F} \nabla F(x_k) \right\}, \quad (4)$$

where conv denotes the convex set.

Definition 2. Let $r : R^n \rightarrow R$ be a locally Lipschitz continuous function. We call $\tilde{r} : R^n \times R_{++} \rightarrow R$ a smoothing function of r , if $\tilde{r}(\cdot, \mu)$ is continuously differentiable for any fixed $\mu \in R_{++}$ and

$$\lim_{x \rightarrow \tilde{x}, \mu \downarrow 0} \tilde{r}(x, \mu) = r(\tilde{x}). \quad (5)$$

In the following, we will give the new method for (1). In order to describe the method clearly, we divide this section into two parts. In Case 1, we give the new nonlinear conjugate gradient method for smooth objective function. In Case 2, we give the new smoothing nonlinear conjugate gradient method for nonsmooth objective function.

Denote that

$$f(x) = \frac{1}{2} \|H(x)\|^2, \quad (6)$$

where H is defined in (3). Then (1) is equivalent to the following unconstrained minimization problem with zero optimal value

$$\min_{x \in R^n} f(x). \quad (7)$$

Case 1. In this section, we assume that f is a continuously differentiable function. Then (7) is a standard unconstrained optimization problem. There are many methods for solving the unconstrained optimization problem, such as Newton method, nonlinear conjugate gradient method, and quasi-Newton method [12–16]. Here, based on [13, 14], we will give a new nonlinear conjugate gradient method to solve (7). The iterates for solving (7) is given by

$$x_{k+1} = x_k + \alpha_k d_k, \quad (8)$$

where d_k is the direction and $\alpha_k > 0$ is a step size; in this paper, we use the Wolfe type line search [14]. Compute $\alpha_k > 0$, such that

$$\begin{aligned} f(x_k + \alpha_k d_k) - f(x_k) &\leq -\rho \alpha_k^2 \|d_k\|^2, \\ g(x_k + \alpha_k d_k)^T d_k &\geq -2\sigma \alpha_k \|d_k\|^2, \end{aligned} \quad (9)$$

where $\rho, \sigma \in (0, 1)$, $\rho < \sigma$, and g is the gradient of f . The direction is defined by

$$d_k = \begin{cases} -g_k, & \text{if } k = 1, \\ -g_k + \beta_k d_{k-1}, & \text{if } k \geq 2, \end{cases} \quad (10)$$

where

$$\beta_k = \frac{g_k^T y_{k-1}}{d_{k-1}^T z_{k-1}} - \delta \frac{\|y_{k-1}\|^2}{(d_{k-1}^T z_{k-1})^2} g_k^T d_{k-1}, \quad (11)$$

$$z_{k-1} = y_{k-1} + t_{k-1} s_{k-1}, \quad t_{k-1} = \epsilon_0 + \max\{0, -s_{k-1}^T y_{k-1} / s_{k-1}^T s_{k-1}\}, \\ \epsilon_0 > 0, \quad \delta > 1/4, \quad y_{k-1} = g_k - g_{k-1}, \quad s_{k-1} = x_k - x_{k-1}.$$

Now we give the new method for (7) as follows.

Method 1. Consider the following steps.

Step 0. Given $x_0 \in R^n$, $d_0 = -g_0$, and $k := 0$, if $\|g_0\| = 0$, then stop.

Step 1. Find $\alpha_k > 0$ satisfying (9), and x_{k+1} is given by (8).

Step 2. Compute d_k by (10). Set $k := k + 1$, and go to Step 1.

We also need the following assumptions.

Assumption 3. Consider the following:

- (i) $\bar{L} = \{x \in R^n \mid f(x) \leq f(x_0)\}$ is level bounded;
- (ii) in the neighborhood of \bar{L} , there exists $L > 0$, such that

$$\|g(x) - g(y)\| \leq L \|x - y\|, \quad (12)$$

where $x, y \in U$, U is a neighborhood of \bar{L} .

In the following, we will give the global convergence analysis about Method 1. Firstly, we give some lemmas.

Lemma 4. Let $\{x_k\}$ be generated by Method 1, then

$$g_k^T d_k \leq -\left(1 - \frac{1}{4\mu}\right) \|g_k\|^2, \quad (13)$$

where $\mu > 1/4$.

By [13, Theorem 2.1], we have Lemma 4.

Lemma 5 (see [14]). Suppose that Assumption 3 holds, α_k is computed by (9), and we have

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \quad (14)$$

By Lemmas 4 and 5, we have Lemma 6.

Lemma 6. *Suppose that Assumption 3 holds, α_k is determined by (9), and we get*

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \quad (15)$$

Now, we give the global convergence theorem for Method 1.

Theorem 7. *Suppose that Assumption 3 holds and $\{x_k\}$ is generated by Method 1, then*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (16)$$

Proof. Suppose by contradiction that there exists $\epsilon > 0$, such that

$$\|g_k\| \geq \epsilon, \quad (17)$$

holds for $k \geq 1$. By

$$d_{k-1}^T z_{k-1} \geq \epsilon_0 d_{k-1}^T s_{k-1}, \quad (18)$$

and (11), we have

$$|\beta_k| \leq \left(\frac{L\epsilon_0 + \mu L^2}{\epsilon_0^2} \right) \frac{\|g_k\|}{\|d_{k-1}\|}, \quad (19)$$

$$\|d_k\| \leq \|g_k\| + \left(\frac{L\epsilon_0 + \mu L^2}{\epsilon_0^2} \right) \|g_k\|.$$

Denoting that $m = (1 + ((L\epsilon_0 + \mu L^2)/\epsilon_0^2))$, we get $\|d_k\|^2 \leq m^2 \|g_k\|^2$. Then we have

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \sum_{k=1}^{\infty} \frac{\epsilon^2}{m^2} = +\infty, \quad (20)$$

which contradicts (15). So we get the theorem. \square

Case 2. In this section, we assume that f is a nonsmooth function. Equation (7) is the nonsmooth unstrained optimization problem. Denote that $\tilde{f}(x, \mu)$ is the smoothing function for $f(x)$, $\tilde{g}_k = \nabla_x \tilde{f}(x_k, \mu_k)$, then $\alpha_k > 0$ is computed by

$$\begin{aligned} \tilde{f}(x_k + \alpha_k d_k, \mu_k) &\leq \tilde{f}(x_k, \mu_k) - \rho \alpha_k^2 \|d_k\|^2, \\ \tilde{g}_{k+1}^T d_k &\geq -2\sigma \alpha_k \|d_k\|^2, \end{aligned} \quad (21)$$

where $0 < \rho < \sigma < 1$. The direction is defined by

$$d_k = \begin{cases} -\tilde{g}_k, & \text{if } k = 1, \\ -\tilde{g}_k + \beta_k d_{k-1}, & \text{if } k \geq 2, \end{cases} \quad (22)$$

where

$$\beta_k = \frac{\tilde{g}_k^T \tilde{y}_{k-1}}{d_{k-1}^T \tilde{z}_{k-1}} - \delta \frac{\|\tilde{y}_{k-1}\|^2}{(d_{k-1}^T \tilde{z}_{k-1})^2} \tilde{g}_k^T d_{k-1}, \quad (23)$$

$\tilde{z}_{k-1} = \tilde{y}_{k-1} + t_{k-1} s_{k-1}$, $t_{k-1} = \epsilon_0 + \max\{0, -s_{k-1}^T \tilde{y}_{k-1} / s_{k-1}^T s_{k-1}\}$, $\epsilon_0 > 0$, $\delta > 1/4$, $\tilde{y}_{k-1} = \tilde{g}_k - \tilde{g}_{k-1}$, $s_{k-1} = x_k - x_{k-1}$.

We give the following smoothing nonlinear conjugate gradient method.

Method 2. Give $x_0 \in R^n$, $\mu_0 > 0$, $\gamma > 0$, and $\gamma_1 \in (0, 1)$.

Step 1. Find $\alpha_k > 0$ satisfying (21), and x_{k+1} is given by (8).

Step 2. If $\|\nabla_x \tilde{f}(x_{k+1}, \mu_k)\| \geq \gamma \mu_k$, then set $\mu_{k+1} = \mu_k$; otherwise, let $\mu_{k+1} \leq \gamma_1 \mu_k$.

Step 3. Compute d_k by (22). Set $k := k + 1$, and go to Step 1.

Theorem 8. *Suppose that $\tilde{f}(\cdot, \mu)$ satisfies Assumption 3, for fixed $\mu > 0$. Then the sequence $\{x_k\}$ generated by Method 2 satisfies*

$$\lim_{k \rightarrow \infty} \mu_k = 0, \quad (24)$$

$$\liminf_{k \rightarrow \infty} \|\nabla_x \tilde{f}(x_k, \mu_{k-1})\| = 0.$$

Proof. If the set $\{k \mid \mu_{k+1} \leq \gamma_1 \mu_k\}$ is finite, then for a fixed K , $\|\nabla \tilde{f}(x_k, \mu_{k-1})\| \geq \gamma \mu_{k-1}$, for all $k > K$. Denoteing that $\bar{\mu} = \mu_k$, $k > K$, because $\tilde{f}(\cdot, \bar{\mu})$ is a smooth function, the previous method reduces to Method 1, where $f(x) = \tilde{f}(x, \bar{\mu})$. By Theorem 7, we get

$$\liminf_{k \rightarrow \infty} \|\nabla_x \tilde{f}(x_k, \bar{\mu})\| = 0. \quad (25)$$

So we know that $\|\nabla \tilde{f}(x_k, \mu_{k-1})\| \geq \gamma \mu_{k-1}$ for $k > K$ is impossible. Then we can assume that the set $\{k \mid \mu_{k+1} \leq \gamma_1 \mu_k\}$ is infinite, then

$$\lim_{k \rightarrow \infty} \mu_k = 0. \quad (26)$$

By the infinity of $\{k \mid \mu_{k+1} \leq \gamma_1 \mu_k\}$, we can assume that the set is $\{k_0, k_1, \dots \mid k_0 < k_1 < \dots\}$. Then we have

$$\liminf_{i \rightarrow \infty} \|\nabla \tilde{f}(x_{k_i+1}, \mu_{k_i})\| \leq \gamma_1 \lim_{i \rightarrow \infty} \mu_{k_i}. \quad (27)$$

By

$$\lim_{i \rightarrow \infty} \mu_{k_i} = 0, \quad (28)$$

we have

$$\liminf_{k \rightarrow \infty} \|\nabla_x \tilde{f}(x_k, \mu_{k-1})\| = 0. \quad (29)$$

So we complete the proof. \square

Remark 9. From the result of Theorem 8, we know that for some kinds of smoothing functions [9, 10], any accumulation point x^* of $\{x_k\}$ generated by Method 2 is a Clarke stationary point of f ; that is, $0 \in \partial f(x^*)$.

3. Some Discussions

In this paper, we give a new smoothing nonlinear conjugate gradient method for (1). The new method also guarantees that any accumulation point is a Clarke stationary point of the merit function for (1).

Discussion 1. In our Methods 1 and 2, we can use any line search, which is well defined under the condition that the search directions are descent directions.

Discussion 2. There are some kinds of smoothing functions, which satisfied Assumption 3 for fixed $\mu > 0$ (see [9, 10]). When the smoothing function of f has gradient consistent property, then any accumulation point of the sequence $\{x_k\}$ generated by Method 2 is a Clarke stationary point. Under some assumptions, we can also use the methods in [15] to solve nonsmooth equations with finitely many maximum functions (1).

Discussion 3. The new method can also be used for solving nonlinear complementarity problem (NCP). By F-B function

$$\varphi(a, b) = \sqrt{a^2 + b^2} - (a + b), \quad (30)$$

we know that NCP is equivalent to $\psi(x) = 0$, where ψ is a continuously differentiable function, so we can use the smooth version Method 1 to solve it. Method 2 can also be used to solve the vertical complementarity problems

$$\begin{aligned} H_i(x) &\geq 0, \quad i = 1, \dots, m, \\ \prod_{i=1}^m H_i^j(x) &= 0, \quad j = 1, \dots, n, \end{aligned} \quad (31)$$

where $H_i(x) : R^n \rightarrow R^n$ are continuously differentiable functions.

Discussion 4. The new method can also be used for solving Hamilton-Jacobi-Bellman equations (HJB) (see [17]). The Hamilton-Jacobi-Bellman equations (HJB) is used to find $x \in R^n$, such that

$$\begin{aligned} \max_{1 \leq i \leq k} \{L^i x - l^i\} &= 0, \quad \text{in } \Theta, \\ x &= 0, \quad \text{on } \partial\Theta, \end{aligned} \quad (32)$$

where Θ is a bounded domain in R^d , L^i and $i = 1, \dots, k$ are elliptic operators of second order. HJB arise in stochastic control problems and often are used to solve finance and control problems. By finite element method, we can obtain the following discrete HJB equation; find $x \in R^n$, such that

$$\max_{1 \leq j \leq k} \{A^j x - \zeta^j\} = 0, \quad (33)$$

where $A^j \in R^{n \times n}$, $\zeta^j \in R^n$, $j = 1, \dots, k$.

Discussion 5. We also can consider to use the new method to solve the general variational inequality problem (see [18]). The general variational inequality problem is to compute $x \in R^n$, such that

$$q(x) \in X, \quad (y - q(x))^T p(x) \geq 0, \quad \text{for any } y \in X, \quad (34)$$

where $p, q : R^n \rightarrow R^n$ are two continuously differentiable functions and $X \subseteq R^n$ is a closed convex set. Equation (34)

is denoted by $\text{GVI}(X, p, q)$ in [18]. $\text{GVI}(X, p, q)$ is a generalization of complementarity problems, nonlinear variational inequalities problems, and general nonlinear complementarity problems. The variational inequalities problems are to compute

$$(y - x)^T p(x) \geq 0, \quad \text{for any } y \in X, \quad (35)$$

where $x \in X$. The general nonlinear complementarity problems are to compute $x \in R^n$, such that

$$q(x) \geq 0, \quad p(x) \geq 0, \quad q(x)^T p(x) = 0. \quad (36)$$

We can rewrite (34) as the following nonsmooth equation:

$$q(x) - P_X [q(x) - p(x)] = 0, \quad (37)$$

where P_X is the projection operator onto X under the Euclidean norm. The general nonlinear complementarity problems are widely used in solving engineering problems and economic problems; such that, under some conditions, the n -person noncooperative game problem can be reformulated as (37) (see [19]). Therefore, how to use our new method to solve the general nonlinear complementarity problems and the n -person noncooperative game problem would be an interesting topic and deserves further investigation.

Acknowledgments

This work was supported by National Science Foundation of China (11101231, 70971070), a Project of Shandong Province Higher Educational Science and Technology Program (J10LA05), and International Cooperation Program for Excellent Lecturers of 2011 by Shandong Provincial Education Department.

References

- [1] L. Qi and P. Tseng, "On almost smooth functions and piecewise smooth functions," *Nonlinear Analysis*, vol. 67, no. 3, pp. 773–794, 2007.
- [2] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. 1, Springer, New York, NY, USA, 2003.
- [3] D. Sun and J. Han, "Newton and quasi-Newton methods for a class of nonsmooth equations and related problems," *SIAM Journal on Optimization*, vol. 7, no. 2, pp. 463–480, 1997.
- [4] Y. Gao, "Newton methods for solving two classes of nonsmooth equations," *Applications of Mathematics*, vol. 46, no. 3, pp. 215–229, 2001.
- [5] S.-Q. Du and Y. Gao, "A parametrized Newton method for nonsmooth equations with finitely many maximum functions," *Applications of Mathematics*, vol. 54, no. 5, pp. 381–390, 2009.
- [6] N. Yamashita and M. Fukushima, "Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems," *Mathematical Programming B*, vol. 76, no. 3, pp. 469–491, 1997.
- [7] L. Q. Qi and J. Sun, "A nonsmooth version of Newton's method," *Mathematical Programming A*, vol. 58, no. 3, pp. 353–367, 1993.

- [8] Y. Elfoutayeni and M. Khaladi, "Using vector divisions in solving the linear complementarity problem," *Journal of Computational and Applied Mathematics*, vol. 236, no. 7, pp. 1919–1925, 2012.
- [9] X. Chen and W. Zhou, "Smoothing nonlinear conjugate gradient method for image restoration using nonsmooth nonconvex minimization," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 765–790, 2010.
- [10] X. Chen, "Smoothing methods for nonsmooth, nonconvex minimization," *Mathematical Programming B*, vol. 134, no. 1, pp. 71–99, 2012.
- [11] M. C. Ferris, O. L. Mangasarian, and J. S. Pang, *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic, Dordrecht, The Netherlands, 2001.
- [12] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.
- [13] Z. Dai and F. Wen, "Global convergence of a modified Hestenes-Stiefel nonlinear conjugate gradient method with Armijo line search," *Numerical Algorithms*, vol. 59, no. 1, pp. 79–93, 2012.
- [14] C.-Y. Wang, Y.-Y. Chen, and S.-Q. Du, "Further insight into the Shamanskii modification of Newton method," *Applied Mathematics and Computation*, vol. 180, no. 1, pp. 46–52, 2006.
- [15] Y. Y. Chen and S. Q. Du, "Nonlinear conjugate gradient methods with Wolfe type line search," *Abstract and Applied Analysis*, vol. 2013, Article ID 742815, 5 pages, 2013.
- [16] Y. Xiao, H. Song, and Z. Wang, "A modified conjugate gradient algorithm with cyclic Barzilai-Borwein steplength for unconstrained optimization," *Journal of Computational and Applied Mathematics*, vol. 236, no. 13, pp. 3101–3110, 2012.
- [17] S. Zhou and Z. Zou, "A relaxation scheme for Hamilton-Jacobi-Bellman equations," *Applied Mathematics and Computation*, vol. 186, no. 1, pp. 806–813, 2007.
- [18] X. Chen, L. Qi, and D. Sun, "Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities," *Mathematics of Computation*, vol. 67, no. 222, pp. 519–540, 1998.
- [19] M. C. Ferris and J. S. Pang, "Engineering and economic applications of complementarity problems," *SIAM Review*, vol. 39, no. 4, pp. 669–713, 1997.