第20卷 第3期
1994年7月

曲阜师范大学学报
Journal of Qufu Normal University

Vol. 20 No. 3
July 1994

A MODIFICATION OF A SUCCESSIVE APPROXIMATION METHOD FOR NONSMOOTH EQUATIONS*

Xu Dachuan Sun Defeng

(Institute of Applied Mathematics, Academia Sinica 100080, Beijing)

Abstract A successive approximation method for nonsmooth equations was provided. In this paper, by introducing a positive number sequence. The methud for computing the upper bound of a nonsmooth equations, which is very difficult to implements is avoided, and the global convergence is also proved.

Key words global convergence successive approximation nonsmooth equation integration convolution.

1 Introduction

Let $F: R^n \rightarrow R^n$ be a continuous function. We consider the system of nonlinear equations

$$F(x)=0, \quad x \in R^n. \tag{1}$$

To solve such nonsmooth equations caused many authors' attention, for example, see([1—14]). Qi and Chen proposed a globally convergent successive approximation method for nonsmooth equations in [1]. At the kth step, they approximate F by a smooth function f_k such that $F=f_k+g_k$, where

$$\| g_k \| \equiv \sup\{ \| g_k(x) \| : x \in R^n\} \leqslant \alpha \| F(x_k) \|,$$

and $\alpha \in (0,1)$ is a fixed constant. Such a decomposition is called a normal decomposition of F. Their method can be described as follows.

Let

$$\theta(x)=\frac{1}{2}F(x)^T F(x)$$

and

* 国家自然科学基金资助项目
收稿日期：1994—01—22

$$\theta_k(x)=\frac{1}{2}f_k(x)^T f_k(x).$$

The successive approximation method (SAM)

Given $\rho,\alpha\in(0,1)$, an initial vetor $x_0\in R^n$ and a normal decomposition $F=f_0+g_0$ with $\|g_0\|\leqslant\frac{\alpha}{2}\|F(x_0)\|$, let $0<\sigma<1-\alpha$. For $k\geqslant 0$:

1 Solve $F(x_k)+f_k'(x_k)d=0$ to get d_k.

2 Set $x_{k+1}=x_k+\rho^{m_k}d_k$,

where m_k is the smallest nonnegative integer m such that

$$\theta_k(x_k+\rho^m d_k)-\theta_k(x_k)\leqslant-2\sigma\rho^m\theta(x_k).$$

3 If $F(x_{k+1})=0$, stop. If $\|g_k\|<\alpha\|F(x_{k+1})\|$, we let $f_{k+1}=f_k$ and $g_{k+1}=g_k$. Otherwise, we construct a new normal decomposition

$$F=f_{k+1}+g_{k+1},$$

with $\|g_{k+1}\|<\min\{\frac{\alpha}{2}\|F(x_{k+1})\|,\frac{1}{2}\|g_k\|\}$.

The most outstanding advantage of the above algorithm over existing method is that it keeps feature of linearization at each step such that the subproblem is a system of linear equations. This feature is not possessed by known globally convergent methods for solving nonsmooth equations. In the above algorithm they need to compute the value of $\|g_k\|$ in k th step, which is not a easy, especially for the nonsmooth functions. However, we can easily compute an upper bound of $\|g_k\|$ to implement. In this paper, our main attention is concentrated on avoiding computing $\|g_k\|$.

We use $f_k'(x_k)$ in the algorithm, wherever a derivative of F at x_k is needed. In the whole paper, we denote $\|\cdot\|_2$ by $\|\cdot\|$.

2 Method and Global Convergence

For convenience, we also call the following decomposition of F a normal decomposition.

Definition 1 Let $\alpha\in(0,1)$, β_k be a constant. At the k th step of the iteration methods described in this section and the next section, we call

$$F=f_k+g_k$$

a normal decomposition of F, if f_k is smooth and

$$\|g_k(x_k)\|\leqslant\alpha\|F(x_k)\|,$$

$$\|g_k\|\leqslant\beta_k,$$

whenever $F(x_k)\neq 0$.

Our method can be described as follows:

The modified successive approximation method (MSAM)

Given $\rho, \alpha, \delta \in (0,1)$, an initial vector $x_0 \in R^n$ and a normal decomposition $F = f_0 + g_0$ witn

$$\| g_0 \| < \beta_0 \equiv \frac{\alpha}{2} \| F(x_0) \|,$$

let $0 < \sigma < 1 - \alpha$. For $k \geq 0$:

1 Solve $F(x_k) + f_k'(x_k) d = 0$ to get d_k.

2 Set $x_{k+1} = x_k + \rho^{m_k} d_k$,

where m_k is the smallest nonnegative integer m suvh that

$$\theta_k(x_k + \rho^m d_k) - \theta_k(x_k) \leqslant -2\sigma\rho^m \theta(x_k).$$

3 If $F(x_{k+1}) = 0$, stop. If $\| g_k(x_{k+1}) \| < \alpha \| F(x_{k+1}) \|$, we let $f_{k+1} = f_k$ and $g_{k+1} = g_k$. Otherwise, let $\beta_{k+1} = \delta\beta_k$, we construct a new normal decompositipon

$$F = f_{k+1} + g_{k+1},$$

with

$$\| g_{k+1}(x_{k+1}) \| \leqslant \frac{\alpha}{2} \| F(x_{k+1}) \|,$$

$$\| g_{k+1} \| \leqslant \beta_{k+1}.$$

Assumption 1 The level set

$$D_0 = \{x \in R^n : \theta(x) \leqslant (1+\alpha)^2 \theta(x_0)\}$$

is bounded.

Assumption 2 $f_k'(x_k)$ are nonsingular for all k.

Lemma 1 Suppose that $F(x_k) \neq 0$ and $F = f_k + g_k$ is a normal decomposition of F. Then, there exists a scalar $t_k \in (0,1]$ such that for all $t \in (0, t_k]$

$$\theta_k(x_k + t d_k) - \theta_k(x_k) \leqslant -2\sigma t \theta(x_k). \qquad (2)$$

Proof Notice $\theta_k'(x_k) = f_k'(x_k)^T f_k(x_k)$ and $f_k'(x_k) d_k = -F(x_k)$. We have

$$\begin{aligned}\theta_k(x_k + t d_k) - \theta_k(x_k) &= \frac{1}{2}(f_k(x_k + t d_k)^T f_k(x_k + t d_k) - f_k(x_k)^T f_k(x_k)) \\ &\quad - t d_k^T f_k'(x_k)^T f_k(x_k) + o(t) \\ &= t F(x_k)^T F(x_k) + t F(x_k)^T g_k(x_k) + o(t).\end{aligned}$$

Since $\sigma < 1 - \alpha$, there exists $t_k \in (0,1]$ such that for all $t \in (0, t_k]$, (3) holds.

Lemma 1 indicates that the SAM is well-defined under Assumption 2.

Theorem 1 Suppose that Assumption 1 and 2 hold. Then the SAM is well-defined and for all k,

$$x_k \in D_0. \qquad (3)$$

Let $\{x_k\}$ be a sequence produced by the SAM. If furthermore for an accumulation point x^* of $\{x_k\}$, $f_k'(x^*)$ is nonsingular for large K, then

$$\lim_{k\to\infty} F(x_k)=0 \tag{4}$$

and

$$F(\tilde{x})=0$$

for all accumulation points $\tilde{x}$ of $\{x_k\}$.

Proof Without loss of generality, we may assume that F is not smooth. Hence $\|g_k\|>0$ for any k.

By Lemma 1, the SAM is well-defined. We now prove (3). Without loss of generality, we assume that $F(x_k)\neq 0$ for all k. Let $K=\{0\}\cup\{k: \|g_{k-1}(x_k)\|\geqslant\alpha\|F(x_k)\|\}$. Assume that K consists of $k_0=0<k_1<k_2<\cdots$ Let k be an arbitrary nonegative integer. Let k_j be the largest number in K such that $k_j\leqslant k$. Then

$$f_k=f_{k_j},\qquad g_k=g_{k_j}$$

and

$$\begin{aligned}\|F(x_k)\| &= \|f_k(x_k)+g_k(x_k)\| - \|f_{k_j}(x_k)+g_{k_j}(x_k)\|\\ &\leqslant \|f_{k_j}(x_k)\|+\|g_{k_j}(x_k)\|\leqslant\|f_{k_j}(x_{k_j})\|+\beta_{k_j}\\ &=\|F(x_{k_j})-g_{k_j}(x_{k_j})\|+\beta_{k_j}\leqslant\|F(x_{k_j})\|+\|g_{k_j}(x_{k_j})\|+\beta_{k_j}\\ &\leqslant\|F(x_{k_j})\|+2\beta_{k_j}.\end{aligned}$$

If $j=0$, then $\|F(x_k)\|\leqslant\|F(x_0)\|+\alpha\|F(x_0)\|$, since $\|\beta_0\|\equiv\frac{\alpha}{2}\|F(x_0)\|$.

If $j\geqslant 1$, then

$$\begin{aligned}\|F(x_k)\| &\leqslant \|F(x_{k_j})\|+2\beta_{k_j}\leqslant\frac{1}{\alpha}\|g_{k_j-1}(x_{k_j})\|+2\delta\beta_{k_j-1}\\ &\leqslant\left(\frac{1}{\alpha}+2\delta\right)\beta_{k_j-1}\leqslant\left(\frac{1}{\alpha}+2\delta\right)\delta^{j-1}\beta_0\\ &=\left(\frac{1}{\alpha}+2\delta\right)\frac{\alpha}{2}\delta^{j-1}\|F(x_0)\|\leqslant(1+\alpha)\delta^{j-1}\|F(x_0)\|.\end{aligned} \tag{5}$$

In both cases it follows that $\theta(x_k)\leqslant(1+\alpha)^2\theta(x_0)$. This implies that (3) holds.

We now prove the second part of the theorem. If K is infinite, then for any $k\geqslant 0$, there exists $k_j\subset K$ being the largest number in K such that $k_j\leqslant k$ and (5) holds. The limit in the right-hand side of (5) is zero. This proves (4).

Hence, to prove (4), it suffices to prove that K is infinite. Suppose K is finite and assume $\hat{k}>k$ for all $k\in K$. Then $\|g_{k-1}(x_k)\|<\alpha\|F(x_k)\|$ for all $k\geqslant\hat{k}$. Hence for all $k\geqslant\hat{k}$,

$$f_k\equiv f_{\hat{k}},\qquad g_k\equiv g_{\hat{k}} \tag{6}$$

and

$$\theta(x_k)=\frac{1}{2}\|F(x_k)\|^2>\frac{1}{2\alpha^2}\|g_{k-1}\|^2\equiv\hat{\varepsilon}>0. \tag{7}$$

Suppose that K_0 is a subsequence of $\{0,1,\cdots\}$ such that $\{x_k:k\in K_0\}$ converges to x^*. By (6) and the condition of this theorem, $f_k'(x^*)$ is nonsingular. Since $\lim_{\substack{k\to\infty\\k\in K_0}}x_k=x^*$ and $f_k'(\cdot)$ is a continuous function, $\{\|f_k'(x_k)^{-1}\|:k\subset K_0\}$ is uniformly bounded. Therefore, there exists $L>0$ such that $\|d_k\|=\|f_k'(x_k)^{-1}F(x_k)\|\leqslant L$ for all $k\geqslant\hat{k},k\in K_0$. Since $\theta_k'(\cdot)$ is continuous, we have $\delta>0$ such that for all x satisfying $\|x-x^*\|\leqslant\delta$,

$$|\theta_k'(x)-\theta_k'(x^*)|\leqslant\frac{1-\sigma-\alpha}{L}\hat{\varepsilon}. \tag{8}$$

Since $\lim_{\substack{k\to\infty\\k\in K_0}}x_k=x^*$, we have $\bar{k}>\hat{k}$ such that for all $k>\bar{k}$, $k\in K_0$,

$$\|x_k-x^*\|\leqslant\frac{\delta}{2}. \tag{9}$$

Let $t^*\in(0,1)$ be such that

$$t^*L<\frac{\delta}{2}. \tag{10}$$

By (9) and (10), for all $k>\bar{k}$, $k\in K_0$, $t\in(0,t^*]$ and $\eta\in(0,1)$, we have

$$\|x_k+\eta t d_k-x^*\|\leqslant\delta. \tag{11}$$

Now by (8) and (11), for all $k>\bar{k},k\in K_0$ and $t\in(0,t^*]$, we have

$$\begin{aligned}&|\theta_k(x_k+td_k)-\theta_k(x_k)-td_k^T\theta_k'(x^*)|\\&\leqslant t\|d_k\|\int_0^1|\theta_k'(x_k+\eta td_k)-\theta_k'(x^*)|\mathrm{d}\eta\\&\leqslant t(1-\sigma-\alpha)\hat{\varepsilon}.\end{aligned} \tag{12}$$

Therefore, for all $k\geqslant\bar{k}$, $k\in K_0$ and $t\in(0,t^*]$,

$$\begin{aligned}&\theta_k(x_k+td_k)-\theta_k(x_k)\\&\leqslant td_k^T\theta_k'(x^*)+t(1-\sigma-\alpha)\hat{\varepsilon}\\&\leqslant td_k^T\theta_k'(x_k)+t\|d_k\|\,|\theta_k'(x^*)-\theta_k'(x_k)|+t(1-\sigma-\alpha)\hat{\varepsilon}\\&\leqslant td_k^T\theta_k'(x_k)+tL\frac{1-\sigma-\alpha}{L}\hat{\varepsilon}+t(1-\sigma-\alpha)\hat{\varepsilon}\\&=td_k^Tf_k'(x_k)^Tf_k(x_k)+2t(1-\sigma-\alpha)\hat{\varepsilon}\\&=-tF(x_k)^Tf_k(x_k)+2t(1-\sigma-\alpha)\hat{\varepsilon}\\&=-2t\theta(x_k)+tF(x_k)^Tg_k(x_k)+2t(1-\sigma-\alpha)\hat{\varepsilon}\\&\leqslant-2t\theta(x_k)+t\|F(x_k)\|\,\|g_k(x_k)\|+2t(1-\sigma-\alpha)\theta(x_k)\\&\leqslant-2t\theta(x_k)+2t\alpha\theta(x_k)+2t(1-\sigma-\alpha)\theta(x_k)\end{aligned}$$

$$= -2t\sigma\theta(x_k).$$

This implies that for all $k \geqslant \bar{k}$, $k \in K_0$, we have $\rho^{m_k - 1} \geqslant t^*$, i.e.,

$$\rho^{m_k} \geqslant \rho t^*. \tag{13}$$

By (7), (13) and the construction of our algorithm, for all $k \geqslant \bar{k}, k \in K_0$,

$$\theta_k(x_{k+1}) - \theta_k(x_k) \leqslant -2\sigma\rho^{m_k}\theta(x_k) \leqslant -2\rho t^* \sigma\hat{\varepsilon} < 0.$$

However, by (6) and the construction of our algorithm, $\theta_k(x_k)$ is nonincreasing for $k > \bar{k}$. This implies $\theta_k(x_k) \to -\infty$ as k tends to infinity. This contradicts the facts that $\theta_k(x_k) \geqslant 0$ for all k. Hence, K cannot be finite. This proves (4). The final conclusion of this theorem simply follows (4) and the continuity of F.

3 Some Discussions

The approximate function f_k can be constructed via convolution (see[1]) for nonsmooth equations arising from the variational inequality problem, the maximal monotone operator problem, the nonlinear complementarity problem and nonsmooth partial differential equations. There are already several superlinearly convergent methods [7—8, 12—14] and a superlinear convergence theory [9—10] for solving nonsmooth equations. One may construct a hybrid globally and superlinearly convergent algorithm by the new algorithm and a known superlinearly convergent algorithm with the methodology proposed in [10]. We do not go into details for such a construction.

The authers are grateful to our supervisor professor Jiye Han for his constant helpfulness and guidance. The second author is indebted to professor Liqun Qi for his valuble suggestions on this subject.

References

1 Qi L, Chen X. A globally convergent successive approximation method for nonsmooth equations, School of Mathematics. The University of New South Wales, (Sydney, Australia) 1993.

2 Chen X. On the convergence of Broyden-like methods for nonlinear equations with nondifferentiable terms. Ann Inst Statist Math, 1990, 42: 387～401

3 Chen X, Qi L. Parameterized Newton method and Broyden-like method for solving nonsmooth equations, Applied Mathematics Preprint. AM 92/16, School of Mathematics, The University of New South Wales (Sydney, Australia), 1992.

4 Chen X, Yamamoto T. On the convergence of some quasi-Newton methods for nonlinear equations with nondifferentiable operators. Computing, 1992, 48: 87～94

5 Han S P, Pang J S, Rangaraj N. Globally convergent Newton methods for nonsmooth equations. Math Oper Res, 1992, 17: 586~607

6 Gabriel S A, Pang J S. A trust region method for constrained nonsmooth equations, to appear in: W. W. Hager, D. W. Hearn and P. M. Pardalos, ed. , Large Scale Optimization: State of the Art, Kluwer Academic Publishers B. V.

7 Pang J S. Newton's method for B-differentiable equations. Math Oper Reach, 1990, 15: 311~341

8 Pang J S. A B-differentiable equation based, globally, and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems. Math Prog, 1991, 51: 101~131

9 Pang J S, Qi L, Nonsmooth equations: motivation and algorithms. to appear in: SIAM J Optimization.

10 Qi L. Convergence analysis of some algorithm for solving nonsmooth equations. Math Oper Res, 1993, 18: 227~244.

11 Qi L. Trust region algorithm for solving nonsmooth equations. Applied Mathematics Preprint, AM 92/20. The University of New South Wales (Sydney, Australia), 1992.

12 Qi L, Sun J. A nonsmooth version of Newton's method. to appear in: Math. Prog.

13 Ralph D. Global convergence of damped Newton's method for nonsmooth equations via the path search, to appear in: Math. Oper. Res.

14 Robinson S M. Newton's method for a class of nonsmooth functions, Industrial Engineering Working Paper, University of Wisconsin (Madison, wisconsin) 1988.

一个求解非光滑方程组的限定逐次逼近法

徐大川　孙德锋

(中国科学院应用数学研究所, 100080, 北京市)

摘要　通过引入一个正数列, 提出了求解非光滑方程组的限定逐次逼近法, 证明了算法的全局收敛性, 改进了已有结果.

关键词　全局收敛性　逐次逼近　非光滑方程　积分卷积

分类号　O221