数理解析研究所講究録
1198 巻 2001 年 137-144

A Superlinearly and Globally Convergent Method for Reaction and Diffusion Problems with a Non-Lipschitzian Operator

Xiaojun Chen (陳　小君) *

To appear in Computing

Abstract

This paper proposes a superlinearly and globally convergent method for reaction and diffusion problems with a non-Lipschitz operator. We reformulate the problem as a system of equations with locally Lipschitzian functions. Then the system is solved by using a smoothing Newton method which converges superlinearly and globally. Numerical examples illustrate the reformulation and the smoothing Newton method.

Key word: Non-Lipschitzian operator, smoothing Newton method.
AMS Subject Classifications: 65H10.

1 Introduction

We consider the following system of nonlinear equations

$$F(x) := Ax + Cf(x) - b = 0 \tag{1.1}$$

where A is an $n \times n$ symmetric positive definite matrix, C is an $n \times n$ diagonal matrix with positive diagonal entries $c_i, i = 1, 2, \ldots, n$,

$$f_i(x) = f_i(x_i) = \begin{cases} x_i^p, & x_i \geq 0 \\ 0, & x_i < 0, \end{cases}$$

and b is a vector in R^n. Here $p \in (0,1)$ is a constant. System (1.1) arises from finite element approximations or finite difference approximations for reaction-diffusion

*Department of Mathematics and Computer Science, Shimane University, Matsue 690-8504, Japan. 島根大学総合理工学部数理情報システム学科。This work was supported in part by the Scientific Research Grant C11640119 from the Japan Society of the Promotion of Science.

problems. A typical problem is as follows [1, 2]. Let Ω be a bounded domain in R^2 with a Lipschitz boundary $\partial\Omega$. Given a positive number λ, find u such that

$$\begin{aligned} -\Delta u + \lambda\xi(u) &= 0 \quad \text{in } \Omega \\ u &= 1 \quad \text{on } \partial\Omega, \end{aligned}$$

where

$$\xi(u) = \begin{cases} u^p, & u \geq 0 \\ 0, & u < 0. \end{cases}$$

The difficulty to solve (1.1) is that F is not local Lipschitz. In the last decade, many superlinearly and globally convergent algorithms for nonsmooth equations defined by a locally Lipschitzian operator have been developed [6, 7, 9]. The Rademacher theorem, the Clarke generalized Jacobian and the semismoothness play key roles in convergence analysis of Newton type methods for nonsmooth equations with locally Lipschitzian operators. The Rademacher theorem states that a locally Lipschitzian operator is almost everywhere differentiable. According to the Rademacher theorem, if F is local Lipschitz, the Clarke generalized Jacobian can be defined by [6]

$$\partial F(x) = co\{\lim_{\substack{x^k \to x \\ x^k \in D_F}} F'(x^k)\},$$

where D_F denotes the set of points at which F is differentiable and co denotes the covex hull. A locally Lipschitzian function is called semismooth at x^* if the limit

$$\lim_{\substack{V \in \partial F(x^* + th') \\ h' \to h, t \downarrow 0}} \{Vh'\}$$

exists for any $h \in R^n$ [9]. Unfortunately, F defined in (1.1) is not local Lipschitz. These powerful tools are not applicable for solving (1.1).

Globally convergent methods for solving (1.1) have been studied in [1, 2]. However, the convergence rate of these methods is not expected faster than linear.

The aim of this paper is to present a superlinearly and globally convergent method for solving (1.1). In section 2, we reformulate the system of equations (1.1) as a system of equations defined by a locally Lipschitzian function, and study the Clarke generalized Jacobian and the semismoothness of the new function. In section 3, we give a smooth approximation function of the locally Lipschitzian function. In section 4, we study a smoothing Newton method for solving (1.1) and show that the method is superlinearly and globally convergent. Moreover, we illustrate the reformulation and the method by a numerical example.

The set of all positive real numbers is denoted by $R_{++} = \{t \mid t > 0, t \in R\}$. The index set is denoted by $N = \{1, 2, \ldots, n\}$. We use $\|\cdot\|$ to denote the Euclidean norm.

2 Lipschitz Reformulation

Let $\omega : R \to R$ be defined by

$$\omega(t) := \begin{cases} t^{1/p}, & t \geq 0 \\ t, & t < 0. \end{cases}$$

The function ω is strictly monotonically increasing. Hence the inverse of ω exists and has the form

$$\omega^{-1}(s) = \begin{cases} s^p, & s \geq 0 \\ s, & s < 0. \end{cases}$$

Moreover $\omega(t) \geq 0$ if and only if $t \geq 0$.

Let

$$g(y) = (\omega(y_1), \omega(y_2), \ldots, \omega(y_n))^T.$$

Then by definitions of f and g, we have

$$\begin{aligned} f_i(g_i(y_i)) &= \begin{cases} g_i(y_i)^p, & y_i \geq 0 \\ 0, & y_i < 0 \end{cases} \\ &= \begin{cases} y_i, & y_i \geq 0 \\ 0, & y_i < 0 \end{cases} \\ &= \max(0, y_i), \qquad i \in N. \end{aligned}$$

Now we define a Lipschitz function $H : R^{2n} \to R^{2n}$ as

$$H(x,y) = \begin{pmatrix} Ax + C\max(0,y) - b \\ x - g(y) \end{pmatrix},$$

where "max" denotes the componentwise maximum.

It is easy to see that if (x, y) is a solution of

$$H(x,y) = 0 \tag{2.1}$$

then x is a solution of (1.1). Conversely, if x is a solution of (1.1) then $(x, g^{-1}(x))$ is a solution of (2.1).

The function H is not differentiable only at points (x, y) where $y_i = 0$ for some $i \in N$. In other words, the set of points at which H is differentiable is

$$D_H = \{(x,y) \mid y_i \neq 0, \text{ for all } i \in N\}.$$

Since H is local Lipschitz, we can define the Clarke generalized Jacobian of H. Let $r : R \to R$ be defined by

$$r(t) = \begin{cases} 1 + t^{(1-p)/p}/p & t > 0 \\ 1 & t \leq 0. \end{cases}$$

Let I be the $n \times n$ identity matrix and

$$R(y) = \mathrm{diag}(r(y_1), r(y_2), \ldots, r(y_n)).$$

Theorem 2.1 *The Clarke generalized Jacobian of H at (x, y) is equal to the set of matrices*

$$\begin{pmatrix} A & CQ_y \\ I & -R(y) + \max(0, Q_y) \end{pmatrix},$$

where

$$Q_y = \mathrm{diag}(q_1, q_2, \ldots, q_n)$$

and

$$q_i \in \partial \max(0, y_i) = \begin{cases} \{1\} & t > 0 \\ [0,1] & t = 0 \\ \{0\} & t < 0 \end{cases} \qquad i \in N.$$

Theorem 2.2 *At every point $(x, y) \in R^{2n}$, all elements of $\partial H(x, y)$ are nonsingular.*

Remark 2.1 The function H is a piecewise continuously differentiable function. According to Theorem 4.1 in [9], H is semismooth in R^{2n}.

3 Smoothing Function of H

In this section, we study smoothing functions of H. The nonsmoothness of H appreas in two terms: $\max(0, y)$ and $g(y)$. To define a smoothing function of H, we set

$$\theta(t) = \begin{cases} t^{1/p} & t \geq 0 \\ 0 & t < 0. \end{cases}$$

It is easy to see that θ is continuously differentiable in R. Moreover,

$$\omega(t) = \min(\theta(t), t) = t - \max(t - \theta(t), 0), \quad \text{for} \quad t \leq 1.$$

Now for $\max(0, t)$, we use the following smoothing function

$$\phi(t, \epsilon) = \begin{cases} \max(0, t) & |t| \geq \epsilon \\ \frac{1}{4\epsilon}(t + \epsilon)^2 & |t| < \epsilon. \end{cases}$$

Let $\alpha \in (0, p^{p/(1-p)}]$. We define the following smoothing function for ω.

$$\psi(t, \epsilon) = \begin{cases} t - \frac{1}{4\epsilon}(t - \theta(t) + \epsilon)^2, & t \leq \alpha \,\&\, |t - \theta(t)| < \epsilon \\ \omega(t), & \text{otherwise.} \end{cases}$$

Proposition 3.1 *Functions ψ, $\psi : R \times R_{++} \to R$ satisfy the following properties.*

1. *For every fixed ϵ, $\phi(t, \epsilon)$ and $\psi(t, \epsilon)$ are continuously differentiable with respect to t in R.*

2. *For every $t \in R$,*

$$|\phi(t, \epsilon) - \max(0, t)| \leq \epsilon.$$

$$|\psi(t, \epsilon) - \omega(t)| \leq \epsilon.$$

3. For every fixed $t \in R$,

$$\lim_{\epsilon \downarrow 0} \frac{\partial \phi(t,\epsilon)}{\partial t} = \phi^o(t) \in \partial \max(0,t)$$

and

$$\lim_{\epsilon \downarrow 0} \frac{\partial \psi(t,\epsilon)}{\partial t} = \psi^o(t) \in \partial \omega(t).$$

Using the functions ϕ and ψ, we can define a smoothing function of H. Let

$$\Phi(y,\epsilon) = (\phi(y_1,\epsilon),\ldots,\phi(y_n,\epsilon))$$

$$\Psi(y,\epsilon) = (\psi(y_1,\epsilon),\ldots,\psi(y_n,\epsilon))$$

and

$$\mathcal{H}(x,y,\epsilon) = \begin{pmatrix} Ax + C\Phi(y,\epsilon) - b \\ x - \Psi(y,\epsilon) \end{pmatrix}.$$

For brevity we let $z = (x,y)$.

According to Proposition 3.1, we have the following theorem.

Theorem 3.1 *Function $\mathcal{H} : R^{2n} \times R_{++} \to R^{2n}$ satisfies the following properties:*

1. For every fixed $\epsilon > 0$, $\mathcal{H}$ is continuously differentiable with respect to z in R^{2n}.

2. For every $z \in R^{2n}$,

$$\|\mathcal{H}(z,\epsilon) - H(z)\| \le \epsilon\sqrt{n(\|C\|^2+1)}.$$

3. For every fixed $z \in R^{2n}$,

$$\lim_{\epsilon \downarrow 0} \mathcal{H}_z(z,\epsilon) =: \mathcal{H}^o(z) \in \partial H(z).$$

Theorem 3.2 *$\mathcal{H}_z(z,\epsilon)$ is nonsingular at every point $(z,\epsilon) \in R^{2n} \times R_{++}$.*

4 An algorithm and an example

In this section we study an algorithm which is an application of Algorithm 3.1 in [5] to the system of equations (2.1).

Algorithm 1 *Given $\rho, \tau, \eta \in (0,1)$, and a starting point $z^0 \in R^{2n}$. Choose a scalar $\sigma \in (0, 1-\tau)$. Let $\nu = \tau/(2\sqrt{2n}\max\{1, \|C\|\})$. Let $\beta_0 = \|H(z^0)\|$ and $\epsilon_0 = \nu\beta_0$.*

For $k \ge 0$:

1. Find a solution $\hat{d}^k$ of the system of linear equations

$$H(z^k) + \mathcal{H}^o(z^k)d = 0.$$

If $\|H(z^k + \hat{d}^k)\| \le \eta\beta_k$, let $z^{k+1} = z^k + \hat{d}^k$ and perform Step 3. Otherwise perform Step 2.

2. Find a solution d^k of the system of linear equations

$$H(z^k) + \mathcal{H}_z(z^k, \epsilon_k)d = 0.$$

Let m_k be the smallest nonnegative integer m such that

$$\|\mathcal{H}(z^k + \rho^m d^k, \epsilon_k)\|^2 - \|\mathcal{H}(z^k, \epsilon)\|^2 \leq -\sigma\rho^m \|H(z^k)\|^2.$$

Set $t_k = \rho^{m_k}$ and $z^{k+1} = z^k + t_k d^k$.

3. 3.1 If $\|H(z^{k+1})\| = 0$, terminate.

3.2 If

$$0 < \|H(z^{k+1})\| \leq \max\{\eta\beta_k, \tau^{-1}\|H(z^{k+1}) - \mathcal{H}(z^{k+1}, \epsilon_k)\|\},$$

let

$$\beta_{k+1} = \|H(z^{k+1})\| \quad \text{and} \quad \epsilon_{k+1} = \min\{\nu\beta_{k+1}, \frac{\epsilon_k}{2}\}.$$

3.3 Otherwise, let $\beta_{k+1} = \beta_k$ and $\epsilon_{k+1} = \epsilon_k$.

Theorem 4.1 *The system of equations (2.1) has a unique solution.*

Theorem 4.2 *For any $\gamma > 0$, the set*

$$S_\gamma = \{z \mid \|H(z)\|^2 \leq \gamma\}$$

is nonempty and bounded.

Theorem 4.3 *For any starting point $z^0 \in R^{2n}$, Algorithm 4.1 is well defined and the generated sequence $\{z^k\}$ remains in the level set $S_{(1+\tau)\|H(z^0)\|}$ and converges to the unique solution z^* of (2.1). Moreover, the convergence rate is superlinear.*

To illustrate the smoothing Newton method, we consider the following example [2].

Example 4.1

$$-\Delta u + \frac{9}{(1-p)^2}\max(0, u^p) = \frac{9}{r(1-p)^2}\left(\frac{3r-1}{2}\right)^{\frac{2p}{1-p}} h(r - \frac{1}{3}), \quad \text{in}\, \Omega = (0,1)\times(0,1)$$

$$u(r) = \left(\frac{3r-1}{2}\right)^{\frac{2}{1-p}} h(r - \frac{1}{3}), \quad \text{on}\, \partial\Omega,$$

where $r^2 = x^2 + y^2$ and h is the Heaviside function.

This problem has the solution

$$u(r) = \left(\frac{3r-1}{2}\right)^{\frac{2}{1-p}} h(r - \frac{1}{3}).$$

Application of the five-pint difference scheme with mesh size $1/(\sqrt{n}+1)$ to this problem gives a system of equations $F(x) = 0$. We rewrite the system as a system of equations $H(z) = 0$.

Table 1: Numerical result of Example 4.1

p	n	k	$\|H(z^{k-2})\|$	$\|H(z^{k-1})\|$	$\|H(z^k)\|$
0.1	225	21	2.0E-5	1.2E-8	1.3E-14
0.1	625	26	2.4E-6	1.8E-10	2.2E-14
0.3	225	9	5.8E-4	2.5E-6	5.0E-11
0.3	625	12	5.2E-5	8.8E-8	1.7E-12
0.5	225	6	2.1E-3	8.5E-7	1.0E-12
0.5	625	7	7.8E-5	3.7E-8	5.3E-13

We used Algorithm 4.1 to solve $H(z) = 0$. In our numerical experiment, parameters of Algorithm 4.1 were chosen as

$$\rho = 0.8, \quad \tau = 0.6, \quad \eta = 0.96, \quad \sigma = 0.3.$$

We took the starting point $z^0 = 0$ and stoped the algorithm when $\|H(z^k)\| \le 10^{-10}$.

Numerical results are obtained by using Matlab on a IBM PC. Numerical results show that the Lipschitz reformulation is stable and Algorithm 4.1 is superlinearly and globally convergent. In Table 1, we present $\|H(z^k)\|$ in the last three iterations for different n and p.

References

[1] Aziz,A.K., Stephens,A.B., Suri,M.: Numerical methods for reaction-diffusion problems with non-differentiable kinetics. Numer. Math., 53,1-11(1988).

[2] Barrett,J. W., Shanahan,R. M: Finite element approximation of a model reaction-diffusion problem with a non-Lipschitz nonlinearity. Numer. Math., 59,217-242(1991).

[3] Chen, B., Chen, X.: A global linear and local quadratic continuation smoothing method for variational inequalities with box constraints, Comp. Optim. Appl. (2000).

[4] Chen, X.: Smoothing methods for complementarity problems and their applications: a survey. J. Oper. Res. Soc. Japan, 43, 32-47 (2000).

[5] Chen,X., Ye,Y.: On homotopy-smoothing methods for box-constrained variational inequalites, SIAM J. Control. Optim.,37,589-616 (1999).

[6] Clarke, F.H.: Optimization and Nonsmooth Analysis, Jonh Wiley & Sons, Inc., New York, 1983.

[7] Fukushima, M., Qi, L. eds.: Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers 1999.

[8] Gabriel,S.A., Moré,J.J.: *Smoothing of mixed complementarity problems,* in M.C. Ferris and J.S. Pang, eds., "Complementarity and Variational Problems: State of the Art,", SIAM, Philadelphia, Pennsylvania, 1997, pp. 105-116.

[9] Qi,L., Sun,J.: *A nonsmooth version of Newton's method*, Math. Programming, 58, 353-367 (1993).