Comput Optim Appl (2011) 48: 45–58
DOI 10.1007/s10589-009-9242-9

A globally and quadratically convergent method for absolute value equations

Louis Caccetta · Biao Qu · Guanglu Zhou

Received: 13 August 2008 / Revised: 9 March 2009 / Published online: 21 March 2009

Abstract We investigate the NP-hard absolute value equation (AVE) $Ax - |x| = b$, where A is an arbitrary $n \times n$ real matrix. In this paper, we propose a smoothing Newton method for the AVE. When the singular values of A exceed 1, we show that this proposed method is globally convergent and the convergence rate is quadratic. Preliminary numerical results show that this method is promising.

Keywords Absolute value equations · Smoothing Newton method · Global convergence · Convergence rate

1 Introduction

We consider the absolute value equations (AVE) of the type:

$$Ax - |x| = b, \tag{1.1}$$

where $A \in \Re^{n\times n}, b \in \Re^n$ and $|x|$ denotes the vector with absolute values of each component of x. A slightly more general form of the AVE, $Ax + B|x| = b$ was introduced in [10] and investigated in a more general context in [4]. As was shown in [7],

L. Caccetta · G. Zhou (✉)
Western Australian Centre of Excellence in Industrial Optimisation (WACEIO), Department of Mathematics and Statistics, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia
e-mail: G.Zhou@curtin.edu.au

L. Caccetta
e-mail: caccetta@maths.curtin.edu.au

B. Qu
Institute of Operations Research, Qufu Normal University, Rizhao 276826, Shandong, P.R. China
e-mail: qubiao001@163.com

the general NP-hard linear complementarity problem (LCP) [1–3] which subsumes many mathematical programming problems can be formulated as an AVE (1.1). This implies that the AVE (1.1) is NP-hard in its general form. The AVE (1.1) was investigated theoretically in [7] and a bilinear program was prescribed there for the special case when the singular values of A are not less than one. In [7], some results about the existence of solutions for the AVE (1.1) are given, which we list as follows:

(i) The AVE (1.1) is uniquely solvable for any $b \in \Re^n$ if the singular values of A exceed 1.
(ii) If 1 is not an eigenvalue of A, the singular values of A are merely greater than or equal to 1 and $\{x|(A+I)x-b \geq 0, (A-I)x-b \geq 0\} \neq \emptyset$, then the AVE (1.1) is solvable.
(iii) The AVE (1.1) is uniquely solvable for any b if $\|A^{-1}\| < 1$.
(iv) If $b < 0$ and $\|A\| < \gamma/2$, where $\gamma = \frac{\min_i |b_i|}{\max_i |b_i|}$, then AVE (1.1) has exactly 2^n distinct solutions, each of which has no zero components and a different sign pattern.

Recently, some computational methods have been presented for the AVE (1.1). See [4–6]. In particular, a parametric successive linearization algorithm is proposed in [4, 5] based on a new reformulation of the AVE (1.1) as a parameter-free piecewise linear concave minimization problem. At each iteration, a linear programming problem needs to be solved. It is proved that the algorithm terminates at a point satisfying a necessary optimality condition of the linear concave minimization problem. To our best knowledge, it is unknown if this method has a global convergence property. In [6], a generalized Newton algorithm is proposed for the AVE (1.1). It is proved that the generalized Newton method [6] converges linearly from any starting point to the unique solution of the AVE (1.1) under the condition that $\|A^{-1}\| < \frac{1}{4}$.

In this paper, we present in Sect. 3 a smoothing Newton algorithm to solve the AVE (1.1). This algorithm is proved to be globally convergent and the convergence rate is quadratic under the condition that the singular values of A exceed 1. This condition is weaker than the one used in [6]. Preliminary numerical results given in Sect. 4 show that this smoothing Newton algorithm is very promising.

A word about our notation. All vectors will be column vectors unless transposed to a row vector by a prime T. The inner product of two vectors x and y in the n-dimensional real space $\Re^n$ will be denoted by $x^T y$. For $x \in \Re^n$ the ∞-norm will be denoted by $\|x\|_\infty$ and the 2-norm by $\|x\|$. The notation $A \in \Re^{m\times n}$ will signify a real $m \times n$ matrix. For such a matrix A^T will denote the transpose of A. The identity matrix of arbitrary dimension will be denoted by I. For simplicity, the dimensionality of some vectors and matrices will not be explicitly given. For a square matrix $A \in \Re^{n\times n}$, $\|A\| = \max_{\|x\|=1} \|Ax\|$. $\lambda_{\min}(A)$ denotes the least eigenvalue of A. For a vector $x \in \Re^n$, $\operatorname{diag}(x)$ denotes the $n \times n$ diagonal matrix generated by x. For a vector $d \in \Re^n$, $o(\|d\|)$ stands a vector function of d, satisfying

$$\lim_{d\to 0} \frac{o(\|d\|)}{\|d\|} = 0,$$

while $O(\|d\|^2)$ stands a vector function of d, satisfying

$$\|O(\|d\|^2)\| \le M\|d\|^2$$

for all d satisfying $\|d\| \le \delta$, and some $M > 0$ and $\delta > 0$.

2 A smoothing function and its properties

Define $H : \Re^n \to \Re^n$ by

$$H(x) := Ax - |x| - b. \tag{2.1}$$

It is clear that x is a solution of the AVE (1.1) if and only if $H(x) = 0$. H is a nonsmooth function due to the non-differentiability of the absolute value function. In this section we give a smoothing function of H and study its properties. We first give some properties of H which will be used in the next section.

Lemma 2.1 [7] *For a matrix $A \in \Re^{n\times n}$, the following conditions are equivalent.*

(i) *The singular values of A exceed* 1.
(ii) *The minimum eigenvalue of $A^T A$ exceeds* 1.
(iii) $\|A^{-1}\| < 1$.

Lemma 2.2 [11] *Suppose that A is nonsingular and $\|A^{-1}E\| < 1$. Then, $A + E$ is nonsingular and $(A + E)^{-1}$ can be written in the form*

$$(A + E)^{-1} = (I + F)A^{-1},$$

where

$$\|F\| \le \frac{\|A^{-1}E\|}{1 - \|A^{-1}E\|}.$$

Moreover,

$$\frac{\|A^{-1} - (A + E)^{-1}\|}{\|A^{-1}\|} \le \frac{\|A^{-1}E\|}{1 - \|A^{-1}E\|}.$$

Lemma 2.3 *Let $D = \mathrm{diag}(x)$ with $x_i \in [-1, 1], i = 1, 2, \ldots, n$. Suppose that $\|A^{-1}\| < 1$. Then, $A + D$ is nonsingular.*

Proof Since $\|A^{-1}D\| \le \|A^{-1}\|\|D\| < \|D\| \le 1$, by Lemma 2.2, we have $A + D$ is nonsingular. □

Theorem 2.1 *Suppose that the singular values of A exceed 1. Then, the set $L_1 := \{x \in \Re^n : \|H(x)\| \le \alpha\}$ is bounded for any $\alpha > 0$.*

Proof If the singular values of A exceed 1, then from Lemma 2.1, we know that $\lambda_{\min}(A^T A) > 1$. Using the fact that $\|Ax\| = \sqrt{x^T A^T Ax}$ and $A^T A$ is a symmetric matrix, we have

$$\begin{aligned}\|Ax - |x| - b\| &\geq \|Ax - |x|\| - \|b\|\\ &\geq \|Ax\| - \||x|\| - \|b\|\\ &= \|Ax\| - \|x\| - \|b\|\\ &\geq (\sqrt{\lambda_{\min}(A^T A)} - 1)\|x\| - \|b\|.\end{aligned}$$

Thus, for any $x \in L_1$,

$$(\sqrt{\lambda_{\min}(A^T A)} - 1)\|x\| - \|b\| \leq \alpha,$$

that is,

$$\|x\| \leq \frac{\alpha + \|b\|}{\sqrt{\lambda_{\min}(A^T A)} - 1},$$

which means the set L_1 is bounded. □

Remark 2.1 We can not guarantee that the set $L_1 := \{x \in \Re^n : \|H(x)\| \leq \alpha\}$ is bounded for any $\alpha > 0$ if $\|A^{-1}\| = 1$. For example, if we set $A = I, \alpha = \|b\|$, then for all $x \geq 0$, we can obtain that $x \in L_1$. Obviously, the set L_1 is unbounded.

Let $F : \Re^n \to \Re^m$ be a locally Lipschitzian vector function. By Rademacher's theorem, F is differentiable almost everywhere. Let Ω_F denote the set of points where F is differentiable. Then the B-subdifferential of F at $x \in \Re^n$ is defined to be

$$\partial_B F(x) = \left\{ \lim_{\substack{x^k \to x\\ x^k \in \Omega_F}} F'(x_k)\right\}, \tag{2.2}$$

where $F'(x_k)$ is the Jacobian of F at $x^k \in \Omega_F$. Clarke's generalized Jacobian of F at x is defined to be

$$\partial F(x) = \mathrm{conv}\partial_B F(x), \tag{2.3}$$

(see [8]). Sun and Han [12] introduced a generalized Jacobian ∂_C, defined by

$$\partial_C F(x) = \partial F_1(x) \times \cdots \times \partial F_m(x), \tag{2.4}$$

where, for $i = 1, 2, \ldots, m$, $\partial F_i(x) = \mathrm{conv}\{\lim_{x^k \to x,\ x^k \in \Omega_{F_i}} [\nabla F_i(x_k)]^T\}$. By (2.2–2.4), for any x,

$$\partial_B F(x) \subseteq \partial F(x) \subseteq \partial_C F(x). \tag{2.5}$$

For the function H, at any $x \in \Re^n$, by simple computation, we have

$$\partial_C H(x) = \{A - \mathrm{diag}(d) : d_i \in [-1, 1], i = 1, \ldots, n\}.$$

Hence, by (2.5) and Lemma 2.3, we have

Lemma 2.4 *Suppose that* $\|A^{-1}\| < 1$. *Then, all* $V \in \partial H(x)$ *are nonsingular.*

Definition 2.1 [9] A function $S_\epsilon : \Re^n \to \Re^m$ is called a smoothing function of a nonsmooth function $M : \Re^n \to \Re^m$ if, for any $\epsilon > 0$, $S_\epsilon(\cdot)$ is continuously differentiable and, for any $x \in \Re^n$,

$$\lim_{\epsilon \downarrow 0, z \to x} S_\epsilon(z) = M(x).$$

Definition 2.2 [9] Let $M : \Re^n \to \Re^m$ be a locally Lipschitz continuous function.

(i) $S_\epsilon : \Re^n \to \Re^m$ is called a regular smoothing function of M if, for any $\epsilon > 0$, S_ϵ is continuously differentiable and, for any compact set $D \subseteq \Re^n$ and $\bar{\epsilon} > 0$, there exists a constant $L > 0$ such that, for any $x \in D$ and $\epsilon \in (0, \bar{\epsilon}]$,

$$\|S_\epsilon(x) - M(x)\| \le L\epsilon.$$

(ii) S_ϵ is said to approximate M at x superlinearly if, for any $y \to x$ and $\epsilon \downarrow 0$, we have

$$S_\epsilon(y) - M(x) - S'_\epsilon(y)(y - x) = o(\|y - x\|) + O(\epsilon).$$

(iii) S_ϵ is said to approximate M at x quadratically if, for any $y \to x$ and $\epsilon \downarrow 0$, we have

$$S_\epsilon(y) - M(x) - S'_\epsilon(y)(y - x) = O(\|y - x\|^2) + O(\epsilon).$$

It is clear that a regular smoothing function of M is a smoothing function of M.

For any $\epsilon > 0$, let $\sqrt{x^2 + \epsilon^2} = (\sqrt{x_1^2 + \epsilon^2}, \ldots, \sqrt{x_n^2 + \epsilon^2})^T$. Define $G_\epsilon : \Re^n \to \Re^n$ by

$$G_\epsilon(x) = Ax - \sqrt{x^2 + \epsilon^2} - b. \tag{2.6}$$

Clearly, G_ϵ is a smoothing function of H. Now we give some properties of G_ϵ, which will be used in the next section.

By simple computation, we have

Lemma 2.5 *For any* $\epsilon > 0$, *the Jacobian of* G_ϵ *at* $x \in \Re^n$ *is*

$$G'_\epsilon(x) = A - \operatorname{diag}\left(\frac{x_i}{\sqrt{x_i^2 + \epsilon^2}}, i = 1, 2, \ldots, n\right).$$

Now We investigate the nonsingularity of the matrix $G'_\epsilon(x)$. Note that for any $\epsilon > 0$, $|\frac{x_i}{\sqrt{x_i^2+\epsilon^2}}| < 1$. Hence, by Lemma 2.3, we obtain the following result.

Theorem 2.2 *Suppose that* $\|A^{-1}\| < 1$. *Then,* $G'_\epsilon(x) = A - \operatorname{diag}(\frac{x_i}{\sqrt{x_i^2+\epsilon^2}}, i = 1, 2, \ldots, n)$ *is nonsingular.*

The following theorem gives the boundedness of the inverse matrix of $G'_\epsilon(x)$.

Theorem 2.3 *Suppose that* $\|A^{-1}\| < 1$. *Then, for any* $\epsilon > 0$ *and any* $x \in \Re^n$, *there exists a constant* $M > 0$ *such that*

$$\|[G'_\epsilon(x)]^{-1}\| = \left\| \left(A - \operatorname{diag}\left(\frac{x_i}{\sqrt{x_i^2 + \epsilon^2}}, i = 1, \ldots, n \right) \right)^{-1} \right\| \le M.$$

Proof Let $E = -\operatorname{diag}(\frac{x_i}{\sqrt{x_i^2+\epsilon^2}}, i = 1, \ldots, n)$. Then, from Lemma 2.2, we have

$$\|[G'_\epsilon(x)]^{-1}\| = \|(A+E)^{-1}\| \le \|(I+F)A^{-1}\| \le (1 + \|F\|)\|A^{-1}\|,$$

where, $\|F\| \le \frac{\|A^{-1}E\|}{1-\|A^{-1}E\|}$. Thus, it follows from the fact that $\|E\| < 1$, we have

$$\begin{aligned}\|(A+E)^{-1}\| &\le \left(1 + \frac{\|A^{-1}\|\|E\|}{1-\|A^{-1}\|\|E\|}\right)\|A^{-1}\| \le \left(1 + \frac{\|A^{-1}\|}{1-\|A^{-1}\|}\right)\|A^{-1}\| \\ &= \frac{\|A^{-1}\|}{1-\|A^{-1}\|}.\end{aligned}$$

Set $M := \frac{\|A^{-1}\|}{1-\|A^{-1}\|}$ and we get the desired result. □

Lemma 2.6 *Let* $H(x)$ *and* $G_\epsilon(x)$ *be defined as* (2.1) *and* (2.6), *respectively. Then,*

(i) G_ϵ *is a regular smoothing function of* H.
(ii) G_ϵ *approximates* H *at* x *quadratically.*

Proof (i) For any $\epsilon > 0$,

$$\|G_\epsilon(x) - H(x)\| = \|\sqrt{x^2+\epsilon^2} - |x|\| \le \sqrt{n}\epsilon.$$

(ii) By simple computation, we have

$$\begin{aligned}&G_\epsilon(y) - H(x) - G'_\epsilon(y)(y-x) \\ &= \begin{pmatrix} -\sqrt{y_1^2+\epsilon^2} + \sqrt{x_1^2} + \frac{y_1}{\sqrt{y_1^2+\epsilon^2}}(y_1 - x_1) \\ \vdots \\ -\sqrt{y_n^2+\epsilon^2} + \sqrt{x_n^2} + \frac{y_n}{\sqrt{y_n^2+\epsilon^2}}(y_n - x_n) \end{pmatrix}.\end{aligned}$$

Hence, in order to prove that G_ϵ approximates H at x quadratically, we only need to show that

$$-\sqrt{y_i^2+\epsilon^2} + \sqrt{x_i^2} + \frac{y_i}{\sqrt{y_i^2+\epsilon^2}}(y_i - x_i) = O(|y_i - x_i|^2) + O(\epsilon), \quad i = 1, \ldots, n.$$

We consider the following two cases.

Case 1: $x_i = 0$. We have

$$
\begin{aligned}
&-\sqrt{y_i^2+\epsilon^2}+\sqrt{x_i^2}+\frac{y_i}{\sqrt{y_i^2+\epsilon^2}}(y_i-x_i)\\
&=-\sqrt{y_i^2+\epsilon^2}+\frac{y_i^2}{\sqrt{y_i^2+\epsilon^2}}\\
&=-\frac{\epsilon^2}{\sqrt{y_i^2+\epsilon^2}}\\
&=O(\epsilon)\\
&=O(|y_i-x_i|^2)+O(\epsilon).
\end{aligned}
$$

Case 2: $x_i \neq 0$. We have

$$
\begin{aligned}
&-\sqrt{y_i^2+\epsilon^2}+\sqrt{x_i^2}+\frac{y_i}{\sqrt{y_i^2+\epsilon^2}}(y_i-x_i)\\
&=-\sqrt{y_i^2+\epsilon^2}+\sqrt{x_i^2+\epsilon^2}+\frac{y_i}{\sqrt{y_i^2+\epsilon^2}}(y_i-x_i)+\sqrt{x_i^2}-\sqrt{x_i^2+\epsilon^2}\\
&=\frac{x_i^2-y_i^2}{\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2}}+\frac{y_i}{\sqrt{y_i^2+\epsilon^2}}(y_i-x_i)+\sqrt{x_i^2}-\sqrt{x_i^2+\epsilon^2}\\
&=(x_i-y_i)\left(\frac{x_i+y_i}{\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2}}-\frac{y_i}{\sqrt{y_i^2+\epsilon^2}}\right)+\sqrt{x_i^2}-\sqrt{x_i^2+\epsilon^2}\\
&=(x_i-y_i)\left[\frac{x_i\sqrt{y_i^2+\epsilon^2}-y_i\sqrt{x_i^2+\epsilon^2}}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}}\right]+\sqrt{x_i^2}-\sqrt{x_i^2+\epsilon^2}\\
&=(x_i-y_i)^2\left[\frac{x_i\sqrt{y_i^2+\epsilon^2}-y_i\sqrt{x_i^2+\epsilon^2}}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)}\right]+\sqrt{x_i^2}\\
&\quad-\sqrt{x_i^2+\epsilon^2}\\
&=O(|y_i-x_i|^2)+O(\epsilon).
\end{aligned}
$$

Here, the last equality follows from the fact that $|\sqrt{x_i^2} - \sqrt{x_i^2 + \epsilon^2}| \le \epsilon$ and

$$\frac{x_i\sqrt{y_i^2+\epsilon^2} - y_i\sqrt{x_i^2+\epsilon^2}}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)}$$

is bounded when $y_i \to x_i$ and $\epsilon \downarrow 0$. In fact, the latter can be obtained as follows.

$$\begin{aligned}
&\frac{x_i\sqrt{y_i^2+\epsilon^2} - y_i\sqrt{x_i^2+\epsilon^2}}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)} \\
&\quad = \frac{(x_i-y_i)\sqrt{y_i^2+\epsilon^2} - y_i(\sqrt{x_i^2+\epsilon^2}-\sqrt{y_i^2+\epsilon^2})}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)} \\
&\quad = \frac{1}{\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2}} - \frac{y_i(\sqrt{x_i^2+\epsilon^2}-\sqrt{y_i^2+\epsilon^2})}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)} \\
&\quad = \frac{1}{\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2}} - \frac{y_i(x_i^2-y_i^2)}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})^2\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)} \\
&\quad = \frac{1}{\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2}} - \frac{y_i(x_i+y_i)}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})^2\cdot\sqrt{y_i^2+\epsilon^2}}.
\end{aligned}$$

It is easy to see that $\{\frac{1}{\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2}}\}$ and $\{\frac{y_i(x_i+y_i)}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})^2\cdot\sqrt{y_i^2+\epsilon^2}}\}$ are bounded when $y_i \to x_i$ and $\epsilon \downarrow 0$. Hence, $\{\frac{x_i\sqrt{y_i^2+\epsilon^2}-y_i\sqrt{x_i^2+\epsilon^2}}{(\sqrt{x_i^2+\epsilon^2}+\sqrt{y_i^2+\epsilon^2})\cdot\sqrt{y_i^2+\epsilon^2}\cdot(x_i-y_i)}\}$ is bounded for $y_i \to x_i$ and $\epsilon \downarrow 0$. □

Define $\theta : \Re^n \to \Re$ by

$$\theta(x) := (1/2)\|H(x)\|^2.$$

For any $\epsilon > 0$, define $\theta_\epsilon : \Re^n \to \Re$ by

$$\theta_\epsilon(x) := (1/2)\|G_\epsilon(x)\|^2.$$

Following the analogous line as in the proof for Theorem 2.1, we can get the following theorem.

Theorem 2.4 *Suppose that the singular values of A exceed* 1. *Then,*

(i) *The set*

$$L_2 := \{x \in \Re^n : \|G_\epsilon(x)\| \le \alpha\}$$

is bounded for any $\epsilon \ge 0$ and $\alpha > 0$.

(ii) *For any constants* $\bar{\epsilon} > 0$ *and* $\beta > 0$,

$$L_3 := \{x \in \Re^n : \|G_\epsilon(x)\| \leq \beta\epsilon, 0 < \epsilon \leq \bar{\epsilon}\}$$

is bounded.

Proof (i) Let $y := (\sqrt{x_1^2+\epsilon^2}, \sqrt{x_2^2+\epsilon^2}, \ldots, \sqrt{x_n^2+\epsilon^2})^T$. Then,

$$\begin{aligned}\|y\| &= \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2 + n\epsilon^2} = \sqrt{\|x\|^2 + n\epsilon^2} \leq \sqrt{(\|x\| + \sqrt{n}\epsilon)^2} \\ &= \|x\| + \sqrt{n}\epsilon.\end{aligned}$$

Suppose that the singular values of A exceed 1. Then, from Lemma 2.1, we have $\lambda_{\min}(A^TA) > 1$. Using the fact that $\|Ax\| = \sqrt{x^TA^TAx}$ and A^TA is a symmetric matrix, we have

$$\begin{aligned}\|G_\epsilon(x)\| &= \|Ax - y - b\| \\ &\geq \|Ax - y\| - \|b\| \\ &\geq \|Ax\| - \|y\| - \|b\| \\ &\geq \sqrt{\lambda_{\min}(A^TA)}\|x\| - (\|x\| + \sqrt{n}\epsilon) - \|b\| \\ &= (\sqrt{\lambda_{\min}(A^TA)} - 1)\|x\| - \sqrt{n}\epsilon - \|b\|. \qquad (2.7)\end{aligned}$$

Thus, for any $x \in L_2$,

$$(\sqrt{\lambda_{\min}(A^TA)} - 1)\|x\| - \sqrt{n}\epsilon - \|b\| \leq \alpha,$$

that is,

$$\|x\| \leq \frac{\alpha + \sqrt{n}\epsilon + \|b\|}{\sqrt{\lambda_{\min}(A^TA)} - 1}.$$

This means that L_2 is bounded.

For any $x \in L_3$, by (2.7), we have

$$\|x\| \leq \frac{\beta\epsilon + \sqrt{n}\epsilon + \|b\|}{\sqrt{\lambda_{\min}(A^TA)} - 1} \leq \frac{\beta\bar{\epsilon} + \sqrt{n}\bar{\epsilon} + \|b\|}{\sqrt{\lambda_{\min}(A^TA)} - 1},$$

which implies that L_3 is bounded. □

Theorem 2.5 *Suppose that* $\|A^{-1}\| < 1$. *Then, for any* $\epsilon > 0$ *and* $x \in \Re^n$, $\nabla\theta_\epsilon(x) = 0$ *implies that* $\theta_\epsilon(x) = 0$.

Proof For any $\epsilon > 0$ and $x \in \Re^n$,

$$\nabla\theta_\epsilon(x) = [G'_\epsilon(x)]^T G_\epsilon(x).$$

By Theorem 2.2, $G'_\epsilon(x)$ is nonsingular. Hence, if $\nabla\theta_\epsilon(x) = 0$, then $G_\epsilon(x) = 0$ and $\theta_\epsilon(x) = 0$. □

3 Algorithm and its convergence

In this section, we give a smoothing Newton algorithm for solving $H(x)=0$. First we state this algorithm as follows. This algorithm was proposed in [9] for solving variational inequality problems.

Algorithm 3.1

Step 0. Choose constants $\delta \in (0,1)$, $\beta \in (0,+\infty)$, $\sigma \in (0,1/2)$, $\rho_1 \in (0,+\infty)$ and $\rho_2 \in (2,+\infty)$. Let $x^0 \in \Re^n$ be an arbitrary point; let $k := 0$ and $y^0 := x^0$.

Step 1. Let $d^k \in \Re^n$ satisfy

$$G_{\epsilon^k}(y^k) + G'_{\epsilon^k}(y^k)d = 0. \tag{3.1}$$

If (3.1) is not solvable or if

$$-(d^k)^T \nabla\theta_{\epsilon^k}(y^k) \geq \rho_1 \|d^k\|^{\rho_2} \tag{3.2}$$

does not hold, let

$$d^k = -\nabla\theta_{\epsilon^k}(y^k).$$

Step 2. Let l_k be the smallest nonnegative integer l satisfying

$$\theta_{\epsilon^k}(y^k + \delta^l d^k) \leq \theta_{\epsilon^k}(y^k) + \sigma\delta^l \nabla\theta_{\epsilon^k}(y^k)^T d^k. \tag{3.3}$$

If

$$\|G_{\epsilon^k}(y^k + \delta^{l_k} d^k)\| \leq \epsilon^k \beta, \tag{3.4}$$

or if

$$\|H(y^k + \delta^{l_k} d^k)\| \leq (1/2)\|H(x^k)\|, \tag{3.5}$$

let

$$y^{k+1} := y^k + \delta^{l_k} d^k, x^{k+1} := y^{k+1},$$

and

$$0 < \epsilon^{k+1} \leq \min\{(1/2)\epsilon_k, \theta(x^{k+1})\}.$$

Replace k by $k+1$ and go to Step 1. Otherwise, let

$$y^k := y^k + \delta^{l_k} d^k,$$

and go to Step 1.

In order to discuss the convergence properties of Algorithm 3.1 we make the following assumption.

Assumption 3.1

(i) There exists a constant $\bar{\epsilon} > 0$ such that

$$D_1 := \{x \in \Re^n : \|G_\epsilon(x)\| \leq \beta\epsilon, 0 < \epsilon \leq \bar{\epsilon}\}$$

is bounded.

(ii) For any $\epsilon > 0$ and $\delta > 0$, the following set:

$$L_{\epsilon,\delta} := \{x \in \Re^n : \theta_\epsilon(x) \leq \delta\}$$

is bounded.

(iii) There exists a constant $c > 0$ such that

$$D_2 := \{x \in \Re^n : \theta(x) \leq c\}$$

is bounded.

(iv) For any $\epsilon > 0$ and $x \in \Re^n$, $\nabla\theta_\epsilon(x) = 0$ implies that $\theta_\epsilon(x) = 0$.

Now we summarize the convergence results of Algorithm 3.1 as follows.

Theorem 3.6 [9] *Suppose that Assumption 3.1 holds. Then, an infinite bounded sequence $\{x^k\}$ is generated by Algorithm 3.1 and any accumulation point of $\{x^k\}$ is a solution of the AVE (1.1). Furthermore, suppose that $\bar{x}$ is an accumulation point of $\{x^k\}$ generated by Algorithm 3.1, all $V \in \partial H(\bar{x})$ are nonsingular and $\{\|(G'_{\varepsilon^k}(x^k))^{-1}\|\}$ is uniformly bounded for all x^k sufficiently close to $\bar{x}$. If G_ε approximates H at $\bar{x}$ quadratically, then the whole sequence $\{x^k\}$ converges to $\bar{x}$ quadratically.*

Theorem 3.7 *Suppose that $\|A^{-1}\| < 1$. Then, an infinite bounded sequence $\{x^k\}$ is generated by Algorithm 3.1 and the whole sequence $\{x^k\}$ converges to the unique solution of the AVE (1.1) quadratically.*

Proof By Theorems 2.1, 2.4 and 2.5, Assumption 3.1 holds. It follows from Lemma 2.4 and Theorem 2.3 that all $V \in \partial H(\bar{x})$ are nonsingular and $\{\|(G'_{\varepsilon^k}(x^k))^{-1}\|\}$ is uniformly bounded for all x^k sufficiently close to $\bar{x}$. From Lemma 2.6, we have that G_ε approximates H at $\bar{x}$ quadratically. Hence, we get the result of this theorem by Theorem 3.6. □

Remark 3.2 Recently, a generalized Newton method [6] is proposed for the AVE (1.1). It is proved in Proposition 7 [6] that the generalized Newton method [6] converges *linearly* from any starting point to the unique solution of the AVE (1.1) under the conditions that $\|A^{-1}\| < \frac{1}{4}$ and $\|D(x^k)\| \neq 0$. From Theorem 3.7, the proposed method in this paper converges *quadratically* from any starting point to the unique solution of the AVE (1.1) under the condition that $\|A^{-1}\| < 1$. This condition is weaker than the ones used in [6].

4 Preliminary numerical results

In this section we perform some numerical tests in order to show the viability of Algorithm 3.1. The proposed algorithm was implemented in MATLAB 7.1. Throughout the computational experiments, the parameters used in the algorithm were set as $\delta = 0.5$, $\beta = 1$, $\sigma = 0.0005$, $\rho_1 = 10^{-8}$, and $\rho_2 = 2.1$. We used $\|Ax - |x| - b\|_\infty \leq 10^{-6}$ as the stopping rule.

Three cases of the AVE (1.1) were considered. (i) The singular values of A exceed 1. In this case, the AVE (1.1) is uniquely solvable for any $b \in \Re^n$. (ii) $b < 0$ and $\|A\| < \gamma/2$, where $\gamma = \frac{\min_i |b_i|}{\max_i |b_i|}$. In this case, the AVE (1.1) has exactly 2^n distinct solutions, each of which has no zero components and a different sign pattern. (iii) The AVE (1.1) is constructed as the one in [4]. Chose a random A from a uniform distribution on $[-10, 10]$, then chose a random x from a uniform distribution on $[-1, 1]$ and set $b = Ax - |x|$. In each cases, we tested Algorithm 3.1 on 100 consecutively generated solvable random problems (instances) with fully dense matrices $A \in \Re^{1000\times 1000}$. The maximum number of iterations used per instance was itmax = 100 and the maximum number of line search used per step was lsmax = 20.

In each case, we broke the 100 problems into 10 groups each containing 10 problems (instances). Computational results are summarized in Tables 1–3, respectively. In the following tables, the term "tonvi" denotes the total number of violated instances in each group of 10 problems. "k" denotes the total number of iterations utilized for each group of 10 problems. "tot" denotes the total time for solving each group of 10 problems.

We note the following:

(i) Out of 300 instances, 297 instances were solved to an accuracy of 10^{-6}.
(ii) The average number of iterations per instance is 5.67.
(iii) The overall average time for solving each instance was 2.1139 seconds.

Clearly, Algorithm 3.1 outperforms the successive linear programming algorithm in [4, 5] for these test problems. The performance of Algorithm 3.1 and the generalized Newton method in [6] is similar. Both algorithms can solve the problems in very few iterations.

Table 1 Results from Algorithm 3.1 on case (i), shown in groups of 10

Instances	tonvi	k	tot (seconds)
1–10	0	67	24.59
11–20	0	65	23.7
21–30	0	68	24.88
31–40	0	67	24.54
41–50	0	62	22.63
51–60	0	66	24.053
61–70	0	67	24.51
71–80	0	65	24.14
81–90	0	67	24.45
91–100	0	65	23.77

Table 2 Results from Algorithm 3.1 on case (ii), shown in groups of 10

Instances	tonvi	k	tot (seconds)
1–10	0	49	19.16
11–20	0	50	17.79
21–30	0	49	18.95
31–40	0	49	17.04
41–50	0	48	18.83
51–60	0	51	19.55
61–70	0	47	17.94
71–80	0	49	19.12
81–90	0	50	19.65
91–100	0	49	19.04

Table 3 Results from Algorithm 3.1 on case (iii), shown in groups of 10

Instances	tonvi	k	tot (seconds)
1–10	0	57	21.32
11–20	1	50	18.65
21–30	0	49	18.26
31–40	0	52	19.34
41–50	0	59	22.27
51–60	1	49	18.47
61–70	1	49	18.31
71–80	0	54	20.33
81–90	0	56	21.19
91–100	0	57	21.36

5 Conclusions

In this paper, we have proposed a smoothing Newton method for the NP-hard absolute value equation (AVE) $Ax - |x| = b$. Compared with existing methods in [4–6], the proposed method in this paper has some nice convergence properties. In particular, when the singular values of A exceed 1, we have shown that this proposed method is globally convergent and the convergence rate is quadratic.

Acknowledgements The work of L. Caccetta and G. Zhou is supported by the Australian Research Council Discovery Project DP0665946. The research of B. Qu is supported by the National Natural Science Foundation of China (Grant No. 10701047 and 10571106). Zhou's work is supported by the National Natural Science Foundation of China (Grant No. 10771120).

References

1. Chung, S.J.: NP-completeness of the linear complementarity problem. J. Optim. Theory Appl. **60**, 393–399 (1989)
2. Cottle, R.W., Dantzig, G.: Complementary pivot theory of mathematical programming. Linear Algebra Appl. **1**, 103–125 (1968)

3. Cottle, R.W., Pang, J.S., Stone, R.E.: The Linear Complementarity Problem. Academic Press, New York (1992)
4. Mangasarian, O.L.: Absolute value programming. Comput. Optim. Appl. **36**, 43–53 (2007)
5. Mangasarian, O.L.: Absolute value equation solution via concave minimization. Optim. Lett. **1**, 3–8 (2007)
6. Mangasarian, O.L.: A generalized Newton method for absolute value equations. Optim. Lett. **3**, 101–108 (2009)
7. Mangasarian, O.L., Meyer, R.R.: Absolute value equations. Linear Algebra Appl. **419**, 359–367 (2006)
8. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244 (1993)
9. Qi, L., Sun, D.: Smoothing functions and smoothing Newton method for complementarity and variational inequality problems. J. Optim. Theory Appl. **113**, 121–147 (2002)
10. Rohn, J.: A theorem of the alternatives for the equation $Ax + B|x| = b$. Linear Multilinear Algebra **52**, 421–426 (2004)
11. Stewart, G.W.: Introduction to Matrix Computations. Academic Press, San Diego (1973)
12. Sun, D., Han, J.: Newton and quasi-Newton methods for a class of nonsmooth equations and related problems. SIAM J. Optim. **7**, 463–480 (1997)