International Journal of Control and Automation
Vol.9, No.2 (2016), pp.119-132
http://dx.doi.org/10.14257/ijca.2016.9.2.12

A Smoothing Newton Method for Absolute Value Equation

Longquan Yong

School of Mathematics and Computer Science, Shaanxi University of Technology, Hanzhong 723001, China
yonglongquan@126.com

Abstract

This paper investigate the NP-hard absolute value equation (AVE) $Ax - |x| = b$, where A is an arbitrary square matrix whose singular values exceed one. The significance of the absolute value equation arises from the fact that linear programs, quadratic programs, bimatrix games and other problems can all be reduced to the linear complementarity problem that in turn is equivalent to the absolute value equation. This paper present a new smoothing function to the AVE. Based on this function, a smoothing Newton method is proposed for solving the AVE under the less stringent condition that the singular values of A exceed 1. The global convergence of the method is established under appropriate conditions. Preliminary numerical results indicate that this method is promising.

***Keywords**: Absolute value equation, smoothing function, singular value, smoothing Newton method*

1. Introduction

We consider the absolute value equation (AVE):

$$Ax - |x| = b \tag{1}$$

where $A \in R^{n\times n}$, $x, b \in R^n$, and $|x|$ denotes the vector with absolute values of each component of x. A slightly more general form of the AVE (1) was introduced in John (2004) and investigated in a more general context in Mangasarian (2007a).

As were shown in Cottle .(1968,1992), the general NP-hard linear complementarity problem (LCP) that subsumes many mathematical programming problems can be formulated as an absolute value equation such as (1). This implies that AVE (1) is NP-hard in general form. Theoretical analysis focuses on the theorem of alternatives, various equivalent reformulations, and the existence and nonexistence of solutions. John (2004) provided a theorem of the alternatives for a more general form of AVE (1), $Ax + B|x| = b$, and enlightens the relation between the AVE (1) and the interval matrix. The AVE (1) is shown to be equivalent to the bilinear program, the generalized LCP, and the standard LCP if 1 is not an eigenvalue of A by Mangasarian (2006). Based on the LCP reformulation, sufficient conditions for the existence and nonexistence of solutions are given.

Prokopyev (2009) proved that the AVE (1) can be equivalently reformulated as a standard LCP without any assumption, and discussed unique solvability of AVE (1). Hu and Huang (2009) reformulated a system of absolute value equation as a standard linear complementarity problem without any assumption and give some existence and convexity results for the solution set of the AVE (1).

It is worth mentioning that any LCP can be reduced to the AVE (1), which owns a very special and simple structure. Hence how to solve the AVE directly attracts much attention. Based on a new reformulation of the AVE (1) as the minimization of a

ISSN: 2005-4297 IJCA

parameter-free piecewise linear concave minimization problem on a polyhedral set, Mangasarian (2007b) proposed a finite computational algorithm that is solved by a finite succession of linear programs. In the recent interesting paper of Mangasarian (2009a), a semismooth Newton method is proposed for solving the AVE (1), which largely shortens the computation time than the succession of linear programs (SLP) method. It shows that the semismooth Newton iterates are well defined and bounded when the singular values of A exceed 1. However, the global linear convergence of the method is only guaranteed under more stringent condition than the singular values of A exceed 1. Mangasarian (2009b) formulated the NP-hard *n*-dimensional knapsack feasibility problem as an equivalent AVE (1) in an *n*-dimensional noninteger real variable space and proposed a finite succession of linear programs for solving the AVE (1).

A generalized Newton method, which has global and finite convergence, was proposed for the AVE by Zhang 2009). The method utilizes both the semismooth and the smoothing Newton steps, in which the semismooth Newton step guarantees the finite convergence and the smoothing Newton step contributes to the global convergence. A smoothing Newton algorithm to solve the AVE (1) was presented by Louis Caccetta (2011). The algorithm was proved to be globally convergent and the convergence rate was quadratic under the condition that the singular values of A exceed 1. This condition was weaker than the one used in Mangasarian (2009a).

Recently, AVE (1) has been investigated in Jiri Rohn (2009a, 2009b), Yong (2009,2010), and Noor . (2011a, 2011b). Yong (2010, 2011a) adopted particle swarm optimization (PSO) and harmony search (HS) algorithm to AVE (1), and smoothing Newton method for AVE (1) based on aggregate function in Yong (2011b). Noor (2011a, 2011b) proposed two iterative methods for solving AVE (1).

In this paper, we present a new method for solving AVE (1). We replace the absolute value function by a smooth one given by Li (2011). With this smoothing technique, the non-smooth AVE (1) is formulated as a smooth nonlinear equations, furthermore, we studied properties of the smooth problem. Then we adopt smoothing Newton method to AVE (1). This algorithm is proved to be globally convergent and the convergence rate is linearly at least under the condition that the singular values of A exceed 1. The numerical experiments show that the proposed algorithm is effective in dealing with the AVE (1).

In section 2, we give some propositions or lemmas that ensure the solution to AVE (1) exists. In section 3, we give a smoothing function and study its properties which will be used in the next section. In section 4 we describe smoothing Newton method to AVE (1). Convergence analyses is demonstrated in section 5. Effectiveness of the method is demonstrated in section 6 by solving some given AVE (1) with singular values of A exceeding 1. Section 7 concludes the paper.

We now describe our notation. All vectors will be column vectors unless transposed to a row vector. The scalar (inner) product of two vectors x and y in the n-dimensional real space R^n will be denoted by $x^T y$. For $x \in R^n$ the 2-norm will be denoted by $\|\mathbf{x}\|$, while $|\mathbf{x}|$ will denote the vector with absolute values of each component of x. The notation $A \in R^{m \times n}$ will signify a real $m \times n$ matrix. For such a matrix A^T will denote the transpose of A. We write I for the identity matrix, e for the vector of all ones. A vector of zeros in a real space of arbitrary dimension will be denoted by **0**. $X = diag\{x_i\}$ for the diagonal matrix whose elements are the coordinates x_i of $x \in R^n$.

2. Preliminaries

The following results by Mangasarian . (2006) and Jiri Rohn (2009a) characterize solvability of AVE (1).

Proposition 2.1 (Mangasarian,2007a). (Existence solution)

(i) If 1 is not an eigenvalue of A and the singular values of A are merely greater or equal to 1, then the AVE (1) is solvable if the set $S \neq \varnothing$, where

$$S = \{x | (A+I)x - b \geq 0, (A-I)x - b \geq 0\}.$$

(ii) If $b < 0$ and $\|A\|_\infty < \gamma/2$, where $\gamma = min_i |b_i|/max_i |b_i|$, then AVE (1) has exactly 2^n distinct solutions, each of which has no zero components and a different sign pattern.

Proposition 2.2 (Mangasarian, 2006). (Unique solvability of AVE (1)).

(i) The AVE (1) is uniquely solvable for any $b \in R^n$ if the singular values of A exceed 1.

(ii) The AVE (1) is uniquely solvable for any $b \in R^n$ if $\|A^{-1}\| < 1$.

Proposition 2.3 (Mangasarian, 2006). (Existence of nonnegative solution).

Let $A \geq 0$, $\|A\| < 1$ and $b \leq 0$, then a nonnegative solution to the AVE (1) exists.

Proposition 2.4 (Jiri Rohn,2009b). If the interval matrix $[A - I, A + I]$ is regular, then for each right-hand side b the equation $Ax - |x| = b$ has a unique solution.

Lemma 2.1 [22] For a matrix $A \in R^{n\times n}$, the following conditions are equivalent.

(i) The singular values of A exceed 1.

(ii) The minimum eigenvalue of $A^T A$ exceeds 1.

$$\|A^{-1}\| < 1.$$

Lemma 2.2 Suppose that A is nonsingular and $\|A^{-1}B\| < 1$. Then $A \pm B$ is nonsingular.

Proof We first show that $I \pm A^{-1}B$ is nonsingular.

For, if not, then for some non-zero vector $x \in R^n$ we have that $(I \pm A^{-1}B)x = 0$, which shows $\|x\| \leq \|A^{-1}Bx\| \leq \|A^{-1}B\|\|x\|$, so $\|A^{-1}B\| \geq 1$, a contradiction.

Since $A(I \pm A^{-1}B)$ is nonsingular, we have $A \pm B$ is nonsingular.

Lemma 2.3 Let $D = diag(d_1, d_2, \cdots, d_n)$ with $d_i \in [-1,1], i = 1, 2, \cdots, n$. Suppose that $\|A^{-1}\| < 1$. Then $A \pm D$ is nonsingular.

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

Let $H := R^n \to R^n$ be a locally Lipschitzian vector function. Sun and Han (1997) introduced a generalized Jacobian $\partial_c H$. Let Clarke's generalized Jacobian be ∂H. Then $\partial H \subset \partial_c H$.

For the function $H(x) = Ax - |x| - b$, at any $x \in R^n$, by simple computation, we have

$$\partial_c H(x) = \{A - diag(d) : d_i \in [-1,1], i = 1, 2, \cdots, n\}.$$

Hence, by Lemma 2.3, we have

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

Lemma 2.5 If $F \in R^{n\times n}$ and $\|F\| < 1$, then $I - F$ is nonsingular and $(I - F)^{-1} = \sum_{k=0}^{\infty} F^k$ With $\|(I - F)^{-1}\| \leq \frac{1}{1 - \|F\|}$.

Proof Suppose $I-F$ is singular. It follows that $(I-F)x=0$ for some nonzero x. But then $\|x\|=\|Fx\|$ implies $\|F\|\geq 1$, a contradiction. Thus $I-F$ is nonsingular. To obtain an expression for its inverse consider the identity

$$\left(\sum_{k=0}^{N} F^k\right)(I-F)=I-F^{N+1}.$$

Since $\|F\|<1$ it follows that $\lim_{k\to\infty} F^k=0$ because $\|F^k\|\leq\|F\|^k$.

Thus,

$$\left(\lim_{N\to\infty}\sum_{k=0}^{N} F^k\right)(I-F)=I.$$

It follows that $(I-F)^{-1}=\sum_{k=0}^{\infty} F^k$. From this it is easy to show that

$$\left\|(I-F)^{-1}\right\|\leq\sum_{k=0}^{\infty}\|F\|^k=\frac{1}{1-\|F\|}.$$

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

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

$$\|H(x)\|=\|Ax-|x|-b\|\geq\|Ax-|x|\|-\|b\|\geq\|Ax\|-\|x\|-\|b\|\geq\left(\sqrt{\lambda_{\min}(A^TA)}-1\right)\|x\|-\|b\|$$

.

Thus, for any $x\in L_1$,

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

That is,

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

This is means that L_1 is bounded.

Remark 2.1 We can not guarantee that the set $L_1=\{x\in R^n:\|H(x)\|\leq\alpha\}$ 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.

3. Properties of Smoothing Function

Define $H:=R^n\to R^n$ by

$$H(x):=Ax-|x|-b. \qquad (2)$$

It is clear that x is a solution of the AVE (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.

Firstly, we give some basic concepts.

Definition 3.1 A function $H_\varepsilon(x):=R^n \to R^n$ is called a smoothing function of a nonsmooth function $H(x):=R^n \to R^n$ if, for any $\varepsilon>0$, $H_\varepsilon(x)$ is continuously differentiable and, for any $x \in R^n$, $\lim_{\varepsilon\downarrow 0, y\to x} H_\varepsilon(y)=H(x)$.

Definition 3.2 [24] Let $H(x):=R^n \to R^n$ be a locally Lipschitz continuous function.

(i) $H_\varepsilon(x):=R^n \to R^n$ is called a regular smoothing function of $H(x)$ if, for any $\varepsilon>0$, $H_\varepsilon(x)$ is continuously differentiable and, for any compact set $D \subset R^n$ and $\bar{\varepsilon}>0$, there exists a constant $L>0$ such that, for any $x \in D$ and $\varepsilon \in (0,\bar{\varepsilon}]$, $\|H_\varepsilon(x)-H(x)\| \le L\varepsilon$.

(ii) $H_\varepsilon(x)$ is said to approximate $H(x)$ at x superlinearly if, for any $y \to x$ and $\varepsilon \downarrow 0$, we have $H_\varepsilon(y)-H(x)-H_\varepsilon^{'}(y)(y-x)=o(\|y-x\|)+O(\varepsilon)$.

(iii) $H_\varepsilon(x)$ is said to approximate $H(x)$ at x linearly if, for any $y \to x$ and $\varepsilon \downarrow 0$, we have $H_\varepsilon(y)-H(x)-H_\varepsilon^{'}(y)(y-x)=O(\|y-x\|)+O(\varepsilon)$.

(iv) $H_\varepsilon(x)$ is said to approximate $H(x)$ at x quadratically if, for any $y \to x$ and $\varepsilon \downarrow 0$, we have $H_\varepsilon(y)-H(x)-H_\varepsilon^{'}(y)(y-x)=O(\|y-x\|^2)+O(\varepsilon)$.

It is clear that a regular smoothing function of $H(x)$ is a smoothing function of $H(x)$.

In the following, we recall some good properties of a smoothing approximation function to the absolute value function proposed by Li (2011). Let

$$\phi_\varepsilon(t)=\frac{2}{\pi}\arctan\left(\frac{t}{\varepsilon}\right),$$

$$\psi_\varepsilon(t)=\int \phi_\varepsilon(t)dt = t\phi_\varepsilon(t)-\frac{1}{\pi}\varepsilon\ln\left(1+\frac{t^2}{\varepsilon^2}\right).$$

More precisely, for any $\varepsilon>0$, we have the following proposition from Li (2011).

Proposition 3.1.

(i) $\dfrac{d[\psi_\varepsilon(t)]}{dt}=\phi_\varepsilon(t)$;

(ii) $\left|\dfrac{d[\psi_\varepsilon(t)]}{dt}\right|=|\phi_\varepsilon(t)|\le 1$;

(iii) $\psi_\varepsilon(t)\le |t|$;

(iv) $0<|t|-\psi_\varepsilon(t)<M_{\alpha\beta}\varepsilon, t\in[\alpha,\beta]\cup[-\beta,-\alpha], 0<a\le\beta$, $M_{\alpha\beta}>0$ is a constant.

Let $\varphi(x)=|x|$, where $\varphi(x_i)=|x_i|, i=1,2,\cdots,n$, obviously absolute value function $|x_i|$ is non-differentiable. We can apply the function $\psi_\varepsilon(t)$ to smooth the absolute value function.

The smoothing function to the function $\varphi(x_i)$ is derived as

$$\psi_\varepsilon(x_i) = x_i\phi_\varepsilon(x_i) - \frac{1}{\pi}\varepsilon \ln\left(1+\frac{x_i^{\,2}}{\varepsilon^2}\right) \quad , i = 1, 2, \cdots, n.$$

where $\phi_\varepsilon(x_i) = \frac{2}{\pi}\arctan(\frac{x_i}{\varepsilon})$, $i = 1, 2, \cdots, n$.

For any $\varepsilon > 0$, let $\psi_\varepsilon(x) = \left(\psi_\varepsilon(x_1), \psi_\varepsilon(x_2), \cdots, \psi_\varepsilon(x_n)\right)^T$.

Define $H_\varepsilon := R^n \to R^n$ by

$$H_\varepsilon(x) = Ax - \psi_\varepsilon(x) - b. \tag{3}$$

Clearly, for any $\varepsilon > 0$, $H_\varepsilon(x) = Ax - \psi_\varepsilon(x) - b$ is a smoothing function of $H(x) = Ax - |x| - b$.

Now we give some properties of $H_\varepsilon(x)$. By simple computation, for any $\varepsilon > 0$, the Jacobian of $H_\varepsilon(x)$ at $x \in R^n$ is $H_\varepsilon^{'}(x) = A - diag\left(\phi_\varepsilon(x_1), \phi_\varepsilon(x_2), \cdots, \phi_\varepsilon(x_n)\right)$.

Theorem 3.1 Suppose that $\|A^{-1}\| < 1$. Then $H_\varepsilon^{'}(x)$ is nonsingular for any $\varepsilon > 0$.

Proof Note that for any $\varepsilon > 0$, $|\phi_\varepsilon(x_i)| \le 1$, $i = 1, 2, \cdots, n$.

Hence, by Lemma 2.3, we have $H_\varepsilon^{'}(x)$ is nonsingular.

Let $E = -diag\left(\phi_\varepsilon(x_1), \phi_\varepsilon(x_2), \cdots, \phi_\varepsilon(x_n)\right)$. Then $H_\varepsilon^{'}(x) = A + E$ is nonsingular.

The following theorem gives the boundedness of the inverse matrix $H_\varepsilon^{'}(x)$.

Theorem 3.2 Suppose that $\|A^{-1}\| < 1$. Then, for any $\varepsilon > 0$ and any $x \in R^n$, $\|[H_\varepsilon^{'}(x)]^{-1}\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\|}$.

Proof Since $E = -diag\left(\phi_\varepsilon(x_1), \phi_\varepsilon(x_2), \cdots, \phi_\varepsilon(x_n)\right)$, $|\phi_\varepsilon(x_i)| \le 1$, $i = 1, 2, \cdots, n$, then $\|E\| \le 1$. By Theorem 3.1, we have $A + E$ is nonsingular. Let $F = -A^{-1}E$, $A + E = A(I - F)$, so

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

Thus

$$\|[H_\varepsilon^{'}(x)]^{-1}\| = \|(A+E)^{-1}\| \le \|(I-F)^{-1}\|\|A^{-1}\| \le \frac{\|A^{-1}\|}{1-\|F\|} = \frac{\|A^{-1}\|}{1-\|A^{-1}E\|} \text{ (by Lemma 2.5)}$$

Combining with $\|E\| \le 1$, we have $\|[H_\varepsilon^{'}(x)]^{-1}\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\|}$.

Theorem 3.3 Let $H(x)$ and $H_\varepsilon(x)$ be defined as (2) and (3), respectively. Then,

(i) $H_\varepsilon(x)$ is a regular smoothing function of $H(x)$.

(ii) $H_\varepsilon(x)$ approximates $H(x)$ at x at least linearly.

Proof (i) Since $\|H(x)\text{-}H(y)\| \le \|x - y\|$, $H(x)$ is a locally Lipschitz continuous function. For any $\varepsilon > 0$, $H_\varepsilon(x)$ is continuously differentiable and for any $\varepsilon \in (0, \bar{\varepsilon}]$, and

$$x_i \in D = [\alpha, \beta] \cup [-\beta, -\alpha], 0 < a \le \beta,$$

$$\|H_\varepsilon(x)-H(x)\|=\|-\psi_\varepsilon(x)+\varphi(x)\|=\sqrt{\sum_{i=1}^{n}\left[|x_i|-\psi_\varepsilon(x_i)\right]^2}\le L\varepsilon,$$

Where $L=\sqrt{n}M_{\alpha\beta}$.

(ii) By simple computation, we have

$$H_\varepsilon(y)-H(x)-H_\varepsilon^{'}(y)(y-x)=\begin{pmatrix}-\psi_\varepsilon(y_1)+|x_1|+\phi_\varepsilon(y_1)(y_1-x_1)\\ -\psi_\varepsilon(y_2)+|x_2|+\phi_\varepsilon(y_2)(y_2-x_2)\\ \vdots\\ -\psi_\varepsilon(y_n)+|x_n|+\phi_\varepsilon(y_n)(y_n-x_n)\end{pmatrix}.$$

Since, for $i=1,2,\cdots,n$,

$$|-\psi_\varepsilon(y_i)+|x_i|+\phi_\varepsilon(y_i)(y_i-x_i)|$$

$$=|-\psi_\varepsilon(y_i)+|y_i|+|x_i|-|y_i|+\phi_\varepsilon(y_i)(y_i-x_i)|$$

$$\le|-\psi_\varepsilon(y_i)+|y_i||+||x_i|-|y_i||+|\phi_\varepsilon(y_i)||y_i-x_i|$$

$$\le M_{\alpha\beta}\varepsilon+2|y_i-x_i|=O(|y_i-x_i|)+O(\varepsilon).$$

Hence $H_\varepsilon(x)$ approximates $H(x)$ at x at least linearly.

Theorem 3.4 Suppose that the singular values of A exceed 1. Then

(i) The set $L_2=\{x\in R^n:\|H_\varepsilon(x)\|\le\alpha\}$ is bounded for any $\varepsilon\ge 0$ and $\alpha>0$.

(i) For any constants $\bar{\varepsilon}>0$ and $\beta>0$, $L_3=\{x\in R^n:\|H_\varepsilon(x)\|\le\beta\varepsilon,0<\varepsilon<\bar{\varepsilon}\}$ is bounded.

Proof

(i) Let $\psi_\varepsilon(x)=\left(\psi_\varepsilon(x_1),\psi_\varepsilon(x_2),\cdots,\psi_\varepsilon(x_n)\right)^T$. Then

$$\|\psi_\varepsilon(x)\|=\sqrt{\psi_\varepsilon^{\,2}(x_1)+\psi_\varepsilon^{\,2}(x_2)+\cdots+\psi_\varepsilon^{\,2}(x_n)}\le\sqrt{x_1^{\,2}+x_2^{\,2}+\cdots+x_n^{\,2}}=\|x\|.$$

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

$$\|H_\varepsilon(x)\|=\|Ax-\psi_\varepsilon(x)-b\|\ge\|Ax-\psi_\varepsilon(x)\|-\|b\|$$
$$\ge\|Ax\|-\|\psi_\varepsilon(x)\|-\|b\|\ge\sqrt{\lambda_{\min}(A^TA)}\,\|x\|-\|x\|-\|b\|$$

Thus, for any $x\in L_2$,

$$\sqrt{\lambda_{\min}(A^TA)}\,\|x\|-\|x\|-\|b\|\le\alpha,$$

That is,

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

This is means that L_2 is bounded.

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

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

Define $\theta := R^n \to R$ by

$$\theta(x) = \frac{1}{2}\|H(x)\|^2.$$

For any $\varepsilon > 0$, Define $\theta_\varepsilon := R^n \to R$ by

$$\theta_\varepsilon(x) = \frac{1}{2}\|H_\varepsilon(x)\|^2.$$

We can get the following theorem.

Theorem 3.5 Suppose that $\|A^{-1}\| < 1$. Then, for any $\varepsilon > 0$ and $x \in R^n$, $\nabla\theta_\varepsilon(x) = 0$ implies that $\theta_\varepsilon(x) = 0$.

Proof For any $\varepsilon > 0$ and $x \in R^n$,

$$\nabla\theta_\varepsilon(x) = [H_\varepsilon^{'}(x)]^T H_\varepsilon(x).$$

By Theorem 3.1, $H_\varepsilon^{'}(x)$ is nonsingular. Hence, if $\nabla\theta_\varepsilon(x) = 0$, then $H_\varepsilon(x)$=0 and $\theta_\varepsilon(x) = 0$.

4. Algorithm

Now, we are going to develop the smoothing Newton method for solving $H_\varepsilon(x) = 0$. The method is very similar to the one in [12]. This algorithm was firstly proposed in [24] for solving complementarity and variational inequality problems. The major difference between our method and the method in [12] lies in the use of the smoothing functions. The smoothing function in [12] is constructed based on the $\sqrt{x^2 + \varepsilon^2}$ smoothing function while our smoothing function is obtained on the basis of the smoothing function $\psi_\varepsilon(x)$.

The steps of our smoothing Newton method are stated below.

Algorithm 4.1

Step 0. Choose constants $\delta \in (0,1)$, $\beta \in (0,+\infty)$, $\sigma \in (0,0.5)$, $\rho_1 \in (0,+\infty)$ and $\rho_2 \in (2,+\infty)$. Let $x^0 \in R^n$ be an arbitrary point; let $k := 0$ and $y^0 := x^0$.

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

$$H_{\varepsilon^k}(y^k) + H_{\varepsilon^k}^{'}(y^k)d = 0. \qquad (4.1)$$

If (4.1) is unsolvable or if

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

dose not hold, let

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

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

$$\theta_{\varepsilon^k}(y^k + \delta^l d^k) \le \theta_{\varepsilon^k}(y^k) + \sigma\delta^l \nabla\theta_{\varepsilon^k}(y^k)^T d^k .$$

If

$$\left\| H_{\varepsilon^k}(y^k + \delta^{l_k} d^k) \right\| \le \varepsilon^k \beta ,$$

or if

$$\left\| H(y^k + \delta^{l_k} d^k) \right\| \le \frac{1}{2} \left\| H(x^k) \right\|,$$

let

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

and

$$0 < \varepsilon^{k+1} \le \min\left\{ 0.5\varepsilon^k, \theta(x^{k+1}) \right\}.$$

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.

5. Convergence Analysis

In order to discuss the convergence properties of Algorithm 4.1 we make the following assumption.

Assumption 5.1

(i) There exists a constant $\bar{\varepsilon} > 0$ such that $D_1 := \{x \in R^n : \|H_\varepsilon(x)\| \le \beta\varepsilon, 0 < \varepsilon < \bar{\varepsilon}\}$ is bounded.

(ii) For any constants $\varepsilon > 0$ and $\delta > 0$, the following set: $L_{\varepsilon,\delta} := \{x \in R^n : \theta_\varepsilon(x) \le \delta\}$ is bounded.

(iii) There exists a constant $c > 0$ such that $D_2 := \{x \in R^n : \theta_\varepsilon(x) \le c\}$ is bounded.

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

The convergence results of Algorithm 4.1 are summarized in following Theorems.

Theorem 5.1 [24] Suppose that Assumption 5.1 holds. Then, an infinite bounded sequence $\left\{x^k\right\}$ is generated by Algorithm 4.1 and any accumulation point of $\left\{x^k\right\}$ is a solution of the AVE (1). Furthermore, suppose that $\bar{x}$ is an accumulation point of $\left\{x^k\right\}$ generated by Algorithm 4.1, all $V \in \partial H(\bar{x})$ are nonsingular and $\left\{\left\| \left[H'_{\varepsilon^k}(x^k) \right]^{-1} \right\|\right\}$ is uniformly bounded for all x^k sufficiently close to $\bar{x}$. If $H_\varepsilon(x)$ approximates $H(x)$ at $\bar{x}$ linearly at least, then the whole sequence $\left\{x^k\right\}$ converges to $\bar{x}$ linearly at least.

Theorem 5.2 Suppose that $\|A^{-1}\| < 1$. Then, an infinite bounded sequence $\left\{x^k\right\}$ is generated by Algorithm 4.1 and the whole sequence $\left\{x^k\right\}$ converges to the unique solution of the AVE (1) at least linearly.

Proof By Theorems 2.1, 3.4 and 3.5, assumption 5.1 holds. It follows from Lemma 2.4 and Theorem 3.2 that all $V \in \partial H(\bar{x})$ are nonsingular and $\left\{ \left\| \left[H'_{\varepsilon^k}(x^k) \right]^{-1} \right\| \right\}$ is uniformly bounded for all $\{x^k\}$ sufficiently close to $\bar{x}$. From Theorem 3.3, we have that $H_\varepsilon(x)$ approximates $H(x)$ at $\bar{x}$ linearly at least. Hence, we get the result of this theorem by Theorem 5.1.

Remark 5.1 Recently, a generalized Newton method [9] is proposed for the AVE (1). It is proved in Proposition 7 [9] that the generalized Newton method converges linearly from any starting point to the unique solution of the AVE (1) under the conditions that $\|A^{-1}\| < 0.25$ and $\|D(x^k)\| \neq 0$. From Theorem 5.2, the proposed method in this paper converges linearly at least from any starting point to the unique solution of the AVE (1) under the condition that $\|A^{-1}\| < 1$. This condition is weaker than the ones used in [9].

6. Computational Results

In this section we perform some numerical tests in order to illustrate the implementation and efficiency of the proposed method. The proposed algorithm was implemented in MATLAB 7.6. Throughout the computational experiments, the parameters used in the algorithm were set as δ =0.5, β = 1, σ = 0.0005, $\rho_1 = 10^{-8}$, ρ_2 = 2.1, and $\varepsilon^0 = 1$. We used $\|Ax - |x| - b\|_\infty < 10^{-6}$ as the stopping rule. All the experiments were performed on Windows XP 64 System running on an Hp desktop with Intel(R) Xeon(R) 4×2.4GHz and 6GB RAM. In all instances the Algorithm 4.1 performs extremely well, and finally converges to an optimal solution of the AVE.

Problem 1. Let A be a matrix whose diagonal elements are 500 and the nondiagonal elements are chosen randomly from the interval $[1, 2]$ such that A is symmetric. Let $b = (A - I)e$ where I is the identity matrix of order D and e is $n \times 1$ vector whose elements are all equal to unity such that $x = (1, 1, \cdots, 1)^T$ is the exact solution.

Here the data (A, b) can be generated by following Matlab scripts:

```
D=input('Dimension of matrix A=')
rand('state',0);
A1=zeros(D,D);
for i=1:D
  for j=1:D
    if i==j
      A1(i,j)=500;
    elseif i>j
      A1(i,j)=1+rand;
    else
      A1(i,j)=0;
    end
  end
end
A=A1+(tril(A1,-1))';
b=(A-eye(D))*ones(D,1)
```

Figure 1. Generating Data (A, B) By the Matlab Scripts

and we set the random-number generator to the state of 0 so that the same data can be regenerated. Let the initial guess is $x^0 = (0, 0, \cdots, 0)^T$. Numerical results of this problem are presented in Table 1.

Table 1. Computational Results from Algorithm 4.1

Dimension	**Problem 1**		**Problem 2**	
	No. of Iterations.	Optimal $\theta(x)$	No. of Iterations.	Optimal $\theta(x)$
4	22	5.4202e-025	24	1.4037e-025
8	38	5.7104e-028	49	1.9062e-027
16	34	4.9323e-027	168	1.7157e-025
32	47	6.8549e-025	86	2.7625e-024

Problem 2. Let the matrix A is given by

$$a_{ii} = 4D, \quad a_{i,i+1} = a_{i+1,i} = D, \qquad a_{ij} = 0.5, \quad i = 1, 2, \cdots, D.$$

Let $b = (A - I)e$ where I is the identity matrix order D and e is $n \times 1$ vector whose elements are all equal to unity such that $x = (1, 1, \cdots, 1)^T$ is the exact solution. Let the initial guess is equal to $x^0 = (0, 0, \cdots, 0)^T$. The numerical results are shown in Table 1.

The iteration processes of objective function values $\theta(x)$ and $\theta_\varepsilon(x)$ of $D = 4$ are respectively shown in Figure 2 and Figure 3.

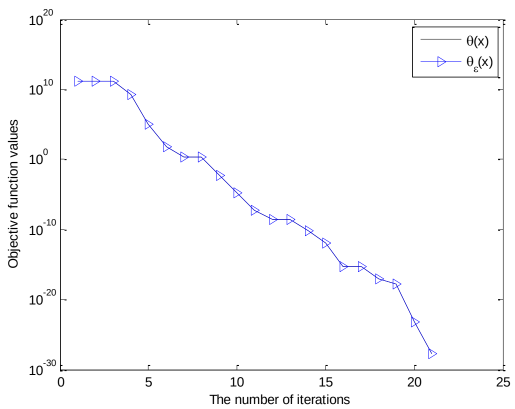

Figure 2. Iteration Processes of Objective Function Values $\theta(x)$ and $\theta_\varepsilon(x)$ for Problem 1

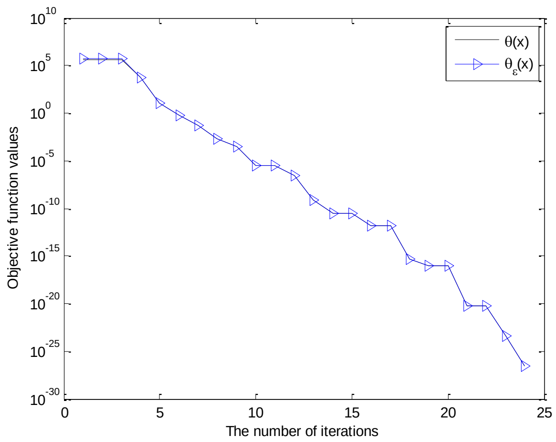

Figure 3. Iteration Processes of Objective Function Values $\theta(x)$ and $\theta_\varepsilon(x)$ for Problem 2

We omitted plots for the others (D = 8, 16, 32) to save space and also in consideration of the fact that they display more or less the same trend.

Problem 3. Following we consider some randomly generated AVE problem with singular values of A exceeding 1 where the data (A, b) are generated by the Matlab scripts:

```
rand('state',0);
b=rand(D,1);
A=rand(D,D)'*rand(D,D)+D*eye(D);
```

and we set the random-number generator to the state of 0 so that the same data can be regenerated. Let the initial guess is equal to $x^0 = (0, 0, \cdots, 0)^T$. The numerical results are shown in Table 2.

Table 2. Computational Results from Algorithm 4.1

Dimension	No. of Iterations.	Optimal $\theta(x)$	Time (seconds)
4	43	1.8698e-026	0.096553
8	67	7.1464e-026	0.169669
16	184	7.0503e-030	0.565197
20	240	3.2137e-030	0.742510
25	402	6.7458e-034	1.532336

7. Conclusion

We have proposed a new smoothing function to the AVE (1). Based on this function, we develop a smoothing Newton method for solving the AVE. Under appropriate conditions, we establish the global convergence of the method. Possible future work may consist of investigating other smoothing function and other optimization algorithm for AVE (1).

Acknowledgments

This work is supported by National Natural Science Foundation of China (Grant No. 11401357), the foundation of Hanzhong Administration of Science & Technology (Grant No.2013hzzx-39), and the foundation of Shaanxi University of Technology (Grant No. SLGKYQD2-14).

References

[1] J. Rohn, “A Theorem of the Alternatives for the Equation Ax+B|x|=b. Linear and Multilinear Algebra”, vol. 52, **(2004)**, pp. 421-426.
[2] O.L. Mangasarian, “Absolute value programming. Comput. Optim. Application”, **(2007a)**, pp. 36:43-53.
[3] R.W. Cottle and G. Dantzig, “Complementary pivot theory of mathematical programming”, Linear Algebra and its Applications, vol. 1, **(1968)**, pp. 103–125.
[4] R.W. Cottle, J.S Pang, and R E Stone, “The Linear Complementarity Problems”, *Academic Press*, New York, **(1992)**.
[5] O.L. Mangasarian and R.R. Meyer, “Absolute value equations”, Linear Algebra Application, vol. 419, **(2006)**, pp. 359-367.
[6] O. Prokopyev, “On equivalent reformulations for absolute value equations”, Computational Optimization and Applications, vol. 3, **(2009)**, pp. 363-372.
[7] S.L. Hu and Z.H. Huang, “A note on absolute value equations”, Optim. Lett, vol. 3, **(2009)**, pp. 417-424.
[8] O.L. Mangasarian, “Absolute value equation solution via concave minimization”, Optim. Lett, vol. 1, **(2007b)**, pp. 3-8.
[9] O.L Mangasarian, “A generalized Newton method for absolute value equations”, Optim. Lett, vol. 3, **(2009a),** pp. 101-108.
[10] O.L. Mangasarian, “Knapsack feasibility as an absolute value equation solvable by successive linear programming”, Optim. Lett, vol. 2, **(2009b),** pp. 161-170.
[11] C. Zhang and Q.J. Wei, “Global and Finite Convergence of a Generalized Newton Method for Absolute Value Equations”, Journal of Optimization Theory and Applications, vol. 143, **(2009)**, pp. 391-403.
[12] L. Caccetta, B. Qu and G. Zhou, “A globally and quadratically convergent method for absolute value equations”, Computational Optimization and Applications, vol. 48, no. 1, **(2011),** pp. 45-58.
[13] J. Rohn, “An Algorithm for Solving the Absolute Value Equation”, Electronic Journal of Linear Algebra, vol. 18, (2009a), pp. 589-599.
[14] J. Rohn, “On Unique Solvability of the Absolute Value Equation”, Optim. Lett, vol. 4, **(2009b),** pp. 603-606.
[15] L. Yong, “Mixed integer linear programming method for absolute value equations”, Proceedings of the 2009 International Symposium on Information Processing (ISIP’09), Finland: Academy Publisher, **(2009),** pp. 316-318.
[16] L. Yong, “Particle Swarm Optimization for Absolute Value Equations”, Journal of Computational Information Systems, vol. 7, **(2010),** pp. 2359-2366.
[17] L. Yong, S. Liu, S. Zhang and F. Deng, “A New Method for Absolute Value Equations Based on Harmony Search Algorithm”, ICIC Express Letters, Part B: Applications, vol. 2, **(2011)**, pp. 1231-1236.
[18] L. Yong, S. Liu, S. Zhang and F. Deng, “Smoothing Newton method for absolute value equations based on aggregate function”, International. Journal of the Physical Sciences, vol. 23, **(2011)**, pp. 5399 - 5405.
[19] M.A. Noor, J. Iqbal, K.I. Noor and E. Al-Said, “On an iterative method for solving absolute value equations”, Optim. Lett. DOI: 10.1007/s11590-011-0332-0, **(2011)**.
[20] M.A. Noor, J. Iqbal, S. Khattri and E. Al-Said, “A new iterative method for solving absolute value equations”, International Journal of the Physical Sciences, vol. 6, **(2011)**, pp. 1793-1797.
[21] Q. Li, D.H. Li, “A Smoothing Newton Method for Nonlinear Complementarity Problems”, Advanced Modeling and Optimization, vol. 13, **(2011)**, pp. 141-152.
[22] G.W. Stewart, “Introduction to Matrix Computations”, **(1973)**, Academic Press, San Diego.
[23] D. Sun, J. Han, “Newton and quasi-Newton methods for a class of nonsmooth equations and related problems”, SIAM Journal of Optim, vol. 7, **(1997)**, pp. 463-480.
[24] L. Qi and D. Sun, “Smoothing functions and smoothing Newton method for complementarity and variational inequality problems”, Journal of Optim. Theory Application, vol. 113, **(2002)**, pp. 121-147.

Author

Longquan Yong, born in Yang County, Shaanxi Province, China, on April 8, 1980. He received his B.S. in Applied Mathematics from Yantai University, Yantai, China, in 2002, the M.S. degree in Applied Mathematics from Xidian University, Xi’an, in 2005, and the Ph.D. degree in Applied Mathematics from Xidian University, Xi’an, in 2014.

He has been working in Shaanxi University of Technology since April,2005. Now he is an associate professor. His current research interests include optimization theory and methods, and swarm intelligent computation. He has published a series of theses in the fields of optimization and intelligent computing.