A Smoothing Newton Method for Solving Absolute Value Equations

Xiaoqin Jiang
Department of public basic, Wuhan Yangtze Business University, Wuhan 430065, P.R. China
392875220@qq.com

Abstract: In this paper, we reformulate the system of absolute value equations as a family of parameterized smooth equations and propose a smoothing Newton method to solve this class of problems. we prove that the method is globally and locally quadratically convergent under suitable assumptions. The preliminary numerical results demonstrate that the method is effective.

Key words: Absolute value equations; smoothing Newton method; global convergence; local quadratic convergence.

1 Introduction

Given $A \in \Re^{n\times n}$, $B \in \Re^{n\times n}$ and $b \in \Re^n$ with positive integer n, we consider the absolute value equation (AVE for short):

$$Ax + B|x| = b, \tag{1}$$

where $|x|$ denotes the component-wise absolute value of $x \in \Re^n$. When $B = -I$ where I is the identity matrix, Eq. 1 induces to $Ax - |x| = b$. It is well known that many problems can be reformulated as AVEs. Eq. 1 was first introduced by Rohn [14]. Recently, AVEs have been extensively studied in the literature. It was shown that determining the existence of a solution to AVE is NP-hard [10]. If AVE is solvable, it has either a unique solution or multiple solutions; and various sufficiency conditions on solvability (or non-solvability) of AVE with unique and multiple solutions were given in [10, 11]. It was also shown that $Ax - |x| = b$ is equivalent to LCP [3, 10], and that the equivalent linear mixed 0-1 reformulations of AVEs were developed in [11]. Various numerical methods for solving AVEs were proposed in the literature (see [8, 9, 15]). Recently, a smoothing Newton method was proposed for solving $Ax - |x| = b$ [1]. It was proved that the method is globally convergent and that the convergence rate is quadratic when the singular values of A exceed 1.

In this paper, we are interested in smoothing Newton method for solving Eq. 1. By using a smooth approximation of the absolute function, we reformulate Eq. 1 as a family of parameterized smooth equations. Then, we propose a smoothing Newton method to solve Eq. 1, which is different from the one in [1]. We show that the method is well-defined under an assumption that the minimal singular value of the matrix A is strictly greater than the maximal singular value of the matrix B. In particular, we show that the method is globally

Published by Atlantis Press, Paris, France.

and locally quadratically convergent without any additional assumption. The preliminary numerical results are reported, which show that the method is effective.

Throughout this paper, we use the following notation. We use e to denote the n-vector of ones and $sgn(\cdot)$ to denote the symbol function. For a matrix $M \in \Re^{n\times n}$, we use $\lambda_i(M)$, $\lambda_{\min}(M)$, and $\lambda_{max}(M)$ to denote the i-th eigenvalue, the smallest eigenvalue, and the largest eigenvalue of M, respectively. For a vector $u \in \Re^n$, denote by $diag\{u_1, u_2, \cdots, u_n\}$ a diagonal matrix with its i-th diagonal element being u_i.

2 Algorithm and Convergence

We define the function $\phi : \Re^2 \to \Re$ by

$$\phi(a, b) := \|(a, b)\|_2 = \sqrt{|a|^2 + |b|^2}, \quad \forall (a, b) \in \Re^2. \tag{2}$$

Then, the function ϕ has the following properties.

Lemma 1 *Let ϕ be defined by Eq. 2, then the following results hold. (i) $\phi(0, b) = |b|$. (ii) ϕ is continuously differentiable on $\Re^2\backslash\{(0, 0)\}$, and when $(a, b) \neq (0, 0)$, $\frac{\partial\phi(a,b)}{\partial a} = \frac{sgn(a)|a|}{\phi(a,b)}$ and $\frac{\partial\phi(a,b)}{\partial b} = \frac{sgn(b)|b|}{\phi(a,b)}$. (iii) ϕ is strongly semismooth on $\Re^2$.*

Proof. The result (i) is obvious; and the proofs of results (ii)-(iii) from [4, Proposition 2.3]. □

We define the function $\Phi : \Re^{n+1} \to \Re^n$ by

$$\Phi(\mu, x) := (\phi(\mu, x_1), \ldots, \phi(\mu, x_n))^T, \quad \forall \mu \in \Re,\ \forall x \in \Re^n \tag{3}$$

and the function $F : \Re^{n+1} \to \Re^{n+1}$ by

$$F(\mu, x) := \begin{pmatrix} \mu \\ Ax + B\Phi(\mu, x) - b \end{pmatrix}, \quad \forall \mu \in \Re,\ \forall x \in \Re^n. \tag{4}$$

Then, we have the following results.

Lemma 2 *Let Φ be defined by Eq. 3, then the following results hold. (i) $F(\mu, x) = 0$ if and only if x solves Eq. 1. (ii) F is continuously differentiable on $\Re^{n+1}\backslash\{0\}$, and when $(\mu, x) \neq 0$, the Jacobian matrix of F at (μ, x) is given by*

$$F'(\mu, x) := \begin{pmatrix} 1 & 0 \\ B\frac{\partial\Phi(\mu,x)}{\partial\mu} & A + B\frac{\partial\Phi(\mu,x)}{\partial x} \end{pmatrix}, \tag{5}$$

where $\frac{\partial\Phi(\mu,x)}{\partial\mu} := \left(\frac{\partial\phi(\mu,x_1)}{\partial\mu}, \ldots, \frac{\partial\phi(\mu,x_n)}{\partial\mu}\right)^T$ and $\frac{\partial\Phi(\mu,x)}{\partial x} := diag\left\{\frac{\partial\phi(\mu,x_1)}{\partial x_1}, \ldots, \frac{\partial\phi(\mu,x_n)}{\partial x_n}\right\}$. (iii) F is strongly semismooth on $\Re^{n+1}$.

Proof. The result (i) holds from Eq. 1 and Lemma 1(i); the result (ii) holds from Lemma 1(ii); and the result (iii) holds from Lemma 1(iii) and the fact that the composition of strongly semismooth functions is strongly semismooth. □

By Lemma 2(i)(ii), instead of solving Eq. 1, one may solve $F(\mu, x) = 0$ by using some Newton-type method, and make $\mu \downarrow 0$ so that a solution of Eq. 1 can be found.

Published by Atlantis Press, Paris, France.

Algorithm 1 *(A Smoothing Newton Method)*

Step 0 *Choose* $\delta, \sigma \in (0,1), \mu_0 > 0, x^0 \in \Re^n$. *Set* $z^0 := (\mu_0, x^0)$. *Denote* $e^0 := (1,0) \in \Re \times \Re^n$. *Choose* $\beta > 1$ *such that* $(\min\{1, \|F(z^0)\|\})^2 \leq \beta\mu_0$. *Set* $k := 0$.

Step 1 *If* $\|F(z^k)\| = 0$, *stop. Otherwise, set* $\tau_k = min\{1, \|F(z^k)\|\}$.

Step 2 *Compute* $\Delta z^k := (\Delta\mu_k, \Delta x^k) \in \Re \times \Re^n$ *by*

$$F(z^k) + F'(z^k)\Delta z^k = (1/\beta)\tau_k^2 e^0, \tag{6}$$

where $F'(\cdot)$ *is defined by Eq. 5.*

Step 3 *Let* α_k *be the maximum of the values* $1, \delta, \delta^2, \cdots$ *such that*

$$\|F(z^k + \alpha_k \Delta z^k)\| \leq [1 - \sigma(1 - 1/\beta)\alpha_k]\|F(z^k)\|. \tag{7}$$

Step 4 *Set* $z^{k+1} := z^k + \alpha_k \Delta z^k$ *and* $k := k+1$. *Go to Step 1.*

The following proposition can be easily proved. We omit its proof here.

Lemma 3 *Let* $\{z^k\}$ *be generated by Algorithm 1. Then, the following results hold. (i) Both sequences* $\{\|F(z^k)\|\}$ *and* $\{\tau_k\}$ *are monotonically decreasing. (ii)* $\tau_k^2 \leq \beta\mu_k$ *holds for all* k. *(iii) The sequence* $\{\mu_k\}$ *is monotonically decreasing, and* $\mu_k > 0$ *for all* k.

In order to show the solvability of Newton equations Eq. 6, we need the following lemma. The proof of this lemma is easy, so we omit it here.

Lemma 4 *Suppose that* $M \in \Re^{n\times n}$ *and* $N \in \Re^{n\times n}$ *are symmetric real matrices. If the minimal singular value of the matrix* M *is strictly greater than the maximal singular value of the matrix* N, *then the matrix* $M - N$ *is positive definite.*

Assumption 1 *The minimal singular value of the matrix* A *is strictly greater than the maximal singular value of the matrix* B.

Now, we show the solvability of Newton equations Eq. 6.

Theorem 1 *Let* F *and* F' *be defined by Eq. 4 and Eq. 5, respectively. Suppose that Assumption 1 holds. Then,* $F'(\mu, x)$ *is invertible at any* $(\mu, x) \in \Re \times \Re^n$ *with* $\mu > 0$.

Proof. From Eq. 5, it is easy to see that $F'(\mu, x)$ is invertible if and only if $A + B\frac{\partial\Phi(\mu,x)}{\partial x}$ is invertible. Suppose that there exists $y \neq 0$ such that $[A + B\frac{\partial\Phi(\mu,x)}{\partial x}]y = 0$. Then,

$$y^T A^T A y = (Ay)^T A y = \left[B\frac{\partial\Phi(\mu,x)}{\partial x}y\right]^T B\frac{\partial\Phi(\mu,x)}{\partial x}y = y^T C^T B^T B C y, \tag{8}$$

where $C := diag\left\{\frac{sgn(x_1)|x_1|}{\|(\mu,x_1)\|_2}, \ldots, \frac{sgn(x_n)|x_n|}{\|(\mu,x_n)\|_2}\right\}$. By Corollary 4.5.11 in [7], there exists a constant ξ such that $0 \leq \lambda_{\min}(C^T C) \leq \xi \leq \lambda_{\max}(C^T C) < 1$ and $\lambda_{\max}(C^T B^T B C) = \xi\lambda_{\max}(B^T B)$. This, together with the assumption that $\lambda_{min}(A^T A) > \lambda_{max}(B^T B) > 0$, implies that $\lambda_{min}(A^T A) > \lambda_{max}(C^T B^T B C)$. Thus, it follows by Lemma 4 that $y^T A^T A y \geq y^T C^T B^T B C y$. This contradicts Eq. 8. The proof is complete. □

By using Theorem 1 we obtain that the system of Newton equations is solvable. Similar to the one in [6, Remark 2.1 (v)], we have that the line search Eq. 7 is well-defined. Thus, Algorithm 1 is well-defined if Assumption 1 holds. In the following, we show convergence of Algorithm 1.

Published by Atlantis Press, Paris, France.

Lemma 5 *Suppose that Assumption 1 holds. Then the sequence $\{z^k\}$ generated by Algorithm 1 is bounded.*

Proof. By Lemma 3 (iii), it is easy to see that the sequence $\{\mu_k\}$ is bounded. Thus, we need only to show that the sequence $\{x^k\}$ is bounded. From Lemma 3 (i), it follows that the sequence $\{\|F(z^k)\|\}$ is bounded. This, together with Eq. 4, implies that the sequence $\{\|Ax^k + B\Phi(\mu_k, x^k) - b\|\}$ is bounded, and hence, the sequence $\{\|Ax^k + B\|(\mu_k, x^k)\|\|\}$ is bounded. Since $\|Ax^k\| - \|B\Phi(\mu_k, x^k)\| \leq \|Ax^k + B\Phi(\mu_k, x^k)\|$, we assume, with loss of generality, that there exists a constant $\zeta > 0$ such that

$$\|Ax^k\| - \|B\Phi(\mu_k, x^k)\| \leq \zeta. \tag{9}$$

In addition, since, for all k, $\|Ax^k\|^2 = (x^k)^T A^T Ax^k > \lambda_{\min}(A^T A)\|x^k\|^2$, $\|B\Phi(\mu_k, x^k)\|^2 < \lambda_{\max}(B^T B)\left\|\Phi(\mu_k, x^k)\right\|^2$, and $\left\|\Phi_2(\mu_k, x^k)\right\|^2 \leq \|x^k\|^2 + n\mu_k^2$, it follows by Eq. 9 that for all k,

$$\begin{aligned}\zeta &\geq \|Ax^k\| - \|B\Phi(\mu_k, x^k)\| > \sqrt{\lambda_{\min}(A^T A)}\|x^k\| - \sqrt{\lambda_{\max}(B^T B)(\|x^k\|^2 + n\mu_k^2)} \\ &\geq (\sqrt{\lambda_{\min}(A^T A)} - \sqrt{\lambda_{\max}(B^T B)})\|x^k\| - \sqrt{n\mu_k^2 \lambda_{max}(B^T B)}.\end{aligned}$$

This and the assumption imply that $\|x^k\| \leq \frac{\zeta + \sqrt{n\mu_k^2 \lambda_{max}(B^T B)}}{\sqrt{\lambda_{\min}(A^T A)} - \sqrt{\lambda_{\max}(B^T B)}}$ holds for all k. Thus, the sequence $\{x^k\}$ is bounded. The proof is complete. □

Theorem 2 *Suppose that Assumption 1 holds and the sequence $\{z^k\}$ is generated by Algorithm 1. Then any accumulation point of $\{x^k\}$ is a solution of Eq. 1.*

Proof. By Lemma 5, we can assume, without loss of generality, that $\lim_{k\to\infty} z^k = z^* = (\mu_*, x^*)$. By Lemma 3 (i), we have $F^* := F(z^*) = \lim_{k\to\infty} F(z^k)$ and $\tau_* := \min\{1, F^*\} = \lim_{k\to\infty} \min\{1, F(z^k)\}$. Now, we show $F^* = 0$. Assume that $F^* > 0$, we derive a contradiction. In this case, it follows by Lemma 3 (ii) that $\mu_* > 0$. The proof is divided into the following two cases.

Case 1. Suppose that $\alpha_k \geq \alpha^* > 0$ for all k, where α^* is a constant. By Eq. 7, we have $\|F(z^{k+1})\| \leq \|F(z^k)\| - \sigma(1 - 1/\beta)\alpha^*\|F(z^k)\|$. Since the sequence $\{F(z^k)\|\}$ is bounded, we have $\sum_{k=0}^{\infty} \alpha^*\sigma(1 - 1/\beta)\|F(z^k)\| < \infty$, which implies that $\lim_{k\to\infty} \|F(z^k)\| = 0$. This contradicts $F^* > 0$.

Case 2. Suppose that $\lim_{k\to\infty} \alpha_k = 0$. Then, for all sufficiently large k, $\hat{\alpha}_k := \alpha_k/\delta$ does not satisfy Eq. 7, i.e., $\|F(z^k + \hat{\alpha}_k \triangle z^k)\| > [1 - \sigma(1 - 1/\beta)\hat{\alpha}_k]\|F(z^k)\|$. So, for all sufficiently large k, it follows that $(\|F(z^k + \hat{\alpha}_k \Delta z^k)\| - \|F(z^k)\|)/\hat{\alpha}_k > -\sigma(1 - 1/\beta)\|F(z^k)\|$. Since $\mu_* > 0$, it follows that F is continuously differentiable at z^*. Let $k \to \infty$, then the above inequality gives

$$\langle F(z^*), F'(z^*)\Delta z^*\rangle / \|F(z^*)\| \geq -\sigma(1 - 1/\beta)\|F(z^*)\|. \tag{10}$$

In addition, by Eq. 6 we have

$$\begin{aligned}\frac{1}{\|F(z^*)\|}\langle F(z^*), F'(z^*)\Delta z^*\rangle &= -\|F(z^*)\| + \frac{\tau_*^2}{\beta\|F(z^*)\|}\langle F(z^*), e^0\rangle \\ &\leq -\|F(z^*)\| + \frac{\tau_*^2\|F(z^*)\|}{\beta\|F(z^*)\|} \leq -\|F(z^*)\| + \tau_*/\beta = (-1 + 1/\beta)\|F(z^*)\|.\end{aligned}$$

Published by Atlantis Press, Paris, France.

This, together with Eq. 10, implies that $-1 + 1/\beta \geq -\sigma(1 - 1/\beta)$, which contradicts the fact that $\sigma \in (0, 1)$ and $\beta > 1$.

By combining **Case 1** with **Case 2**, we obtain $F(z^*) = 0$. Thus, a simple continuation discussion yields that x^* is a solution of Eq. 1. The proof is complete. □

Now, we discuss the local superlinear convergence of Algorithm 1. Generally, in the proof of the local superlinear convergence of the smoothing-type algorithm, it needs the assumption that all generalized Jacobian matrix of the function F at the solution point are nonsingular. The following lemma demonstrates that such an assumption holds trivially for the problem concerned in this paper.

Lemma 6 *Suppose that Assumption 1 holds and $z^* := (\mu_*, x^*)$ is an accumulation point of $\{z^k\}$ generated by Algorithm 1. (i) Define $JF(z^*) := \{\lim F'(z^k) : z^k \to z^*\}$. Then, $JF(z^*) \subseteq \left\{ V \mid V = \begin{pmatrix} 1 & 0 \\ 0 & A + B \cdot diag(d_i) \end{pmatrix}, \ d_i \in [-1, 1], \ i = 1, 2, \ldots, n \right\}$. (ii) All $V \in JF(z^*)$ are nonsingular. (iii) There exist a neighborhood $\mathcal{N}(z^*)$ of z^* and a constant η such that for any $z := (\mu, x) \in \mathcal{N}(z^*)$ with $\mu > 0$, $F'(z)$ is nonsingular and $\|(F'(z))^{-1}\| \leq \eta$.*

Proof. A direct computation yields the result (i). Since Assumption 1 holds, similar to the proof of Lemma 4, it is easy to obtain the result (ii). By [12, Lemma 2.6], we can obtain the result (iii). □

By using Lemma 2 (iii) and Lemma 6, in a similar way as those in [13, Theorem 8], we can obtain the local quadratic convergence of Algorithm 1 as follows.

Theorem 3 *Suppose that Assumption 1 holds and $z^* := (\mu_*, x^*)$ is an accumulation point of the sequence $\{z^k\}$ generated by Algorithm 1. Then the whole sequence $\{z^k\}$ converges to z^* with $\|z^{k+1} - z^k\| = (\|z^k - z^*\|^2)$ and $\mu_{k+1} = \mu_k^2$.*

3 Numerical Experiments

In this section, we report some numerical results of Algorithm 1 for solving Eq. 1. All experiments are done by using a PC with CPU of 2.93GHz and RAM of 2.0GB, and all codes are finished in MATLAB.

Problem: Take $A = 2 * randn(n, n) - 2 * randn(n, n)$ and $B = 2 * randn(n, n) - 2 * randn(n, n)$. In order to ensure that Assumption 1 holds, we update the matrix A by the following: Let $[U \ S \ V] = svd(A)$ and $I \in \Re^{n \times n}$ denote the identity matrix. If $\min\{S(i, i) : i = 1, 2, \ldots, n\} = 0$, we set $A := U(S + 0.01I)V$. Furthermore, set $A := \frac{\lambda_{\max}(B^T B) + 0.01}{\lambda_{\min}(A^T A)} A$. Take $u := randn(n, 1)$ and set $b := Au + B|u|$ so that the problem is feasible.

Throughout our experiments, the following parameters are used: $\delta = 0.6; \sigma = 0.00001; \mu_0 = 0.1$. We choose the starting point $x^0 := randn(n, 1)$. Set $\beta := max\{1, 1.01 * \tau_0^2/\mu\}$. The maximum number of iterations is set as 50. We stop the iteration when $\|F(z^k))\| < 10^{-6}$. We will test **Problem** for $n = 100, 200, \cdots, 1000$, respectively; and every case is randomly generated ten times for testing. The numerical results are listed in Table 1, where n denotes the dimension of the testing problem; for the problem of every size, among ten testings, AIt, ACPU, MaxIt, MinIt, ARes, and AF denote the average value of the iteration numbers, the average value of the CPU time in seconds, the maximal value of the iteration numbers, the minimal value of the iteration numbers, the average value of those of $\|F(z^k))\|$ when Algorithm 1 stops, and the number of testings where the algorithm fails, respectively.

Published by Atlantis Press, Paris, France.

Table 1: The numerical results of Algorithm 1 for **Problem**

n	AIt	ACPU(s)	MaxIt	MinIt	ARes	AF
100	2.200	0.009	3	2	1.493e-008	0
200	2.300	0.037	3	2	4.606e-010	0
300	2.400	0.089	3	2	5.979e-010	0
400	2.200	0.133	3	2	3.687e-009	0
500	2.300	0.266	3	2	7.642e-009	0
600	2.500	0.478	3	2	1.323e-008	0
700	2.400	0.677	3	2	8.452e-009	0
800	2.600	0.944	3	2	6.831e-009	0
900	2.300	1.212	3	2	1.018e-008	0
1000	2.300	1.720	3	2	1.620e-008	0

From Table 1, it is easy to see that all randomly generated problems can be solved with few number of iterations and short CPU time; and the numerical results are very stable such that the iteration number does not change when the size of the problem varies. We have also tested some other problems, and the computation effect is similar. Thus, Algorithm 1 is effective for solving Eq. 1.

4 Final Remarks

In this paper, we have reformulated Eq. 1 as a family of parameterized smooth equations and proposed a smoothing Newton method to solve this class of problems. In particular, we showed that the method is globally and locally quadratically convergent under suitable assumptions. We have also reported the preliminary numerical results, which show the effectiveness of the method. Recently, Hu, Huang, and Zhang [5] extended Eq. 1 to the case of second-order cones, i.e., they introduced the system of absolute value equations associated with second order cones. They proposed a generalized Newton method for solving this class of problems. We believe that the method proposed in this paper can be extended to this class of problems effectively.

References

[1] L. Caccetta, B. Qu, G.L. Zhou, A globally and quadratically convergent method for absolute value equations, Comput. Optim. Appl. 48 (2011) 45–58.
[2] N.J. Higham, Estimating the matrix p-norm, Numer. Math. 62 (1992) 539–55.
[3] S.L. Hu, Z.H. Huang, A note on absolute value equations, Optim. Lett. 4 (2010) 417–424.
[4] S.L. Hu, Z.H. Huang, J.-S. Chen, Properties of a family of generalized NCP-functions and a derivative free algorithm for complementarity problems, J. Comput. Appl. Math. 230 (2009) 69–82.

Published by Atlantis Press, Paris, France.

0488

[5] S.L. Hu, Z.H. Huang, Q. Zhang, A generalized Newton method for absolute value equations associated with second order cones, J. Comput. Appl. Math. 235 (2011) 1490–1501.
[6] Z.H. Huang, Y. Zhang, W. Wu, A smoothing-type algorithm for solving system of inequalities, J. Comput. Appl. Math. 220 (2008) 355–363.
[7] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[8] O.L. Mangasarian, Absolute value equation solution via concave minmization, Optim. Lett. 1 (2007) 3–8.
[9] O.L. Mangasarian, A generalized Newton method for absolute value equations, Optim. Lett. 3 (2009) 101–108.
[10] O.L. Mangasarian, R.R. Meyer, Absolute value equations, Linear Algebra Appl. 419 (2006) 359–367.
[11] O.A. Prokopyev, On equivalent reformulations for absolute value equations, Comput. Optim. Appl. 44 (2009) 363–372.
[12] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res. 18 (1993) 227–244.
[13] L. Qi, D. Sun, G.L. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequality problems, Math. Program. 87 (2000) 1–35.
[14] J. Rohn, A theorem of the alternatives for the equation $Ax + B|x| = b$, Linear and Multilinear Algebra, 52 (2004) 421–426.
[15] C. Zhang, Q.J. Wei, Global and finite convergence of a generalized Newton method for absolute value equations, J. Optim. Theory Appl. 143 (2009) 391–403.

Published by Atlantis Press, Paris, France.