Full Length Research Paper

Smoothing Newton method for absolute value equations based on aggregate function

Longquan Yong[1,2]*, Sanyang Liu[1], Shemin Zhang[2] and Fang'an Deng[2]

[1]Department of Applied Mathematics, Xidian University, Xi'an 710071, China.
[2]Department of Mathematics, Shaanxi University of Technology, Hanzhong 723001, China.

We 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 equations 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 equations. In this paper, we present a smoothing method for the AVE. First, we replace the absolute value function by a smooth one, called aggregate function. With this smoothing technique, the non-smooth AVE is formulated as a smooth nonlinear equations, furthermore, an unconstrained differentiable optimization problem. Then we adopt quasi-Newton method to solve this problem. Numerical results indicate that the method is feasible and effective to absolute value equations.

Key words: Absolute value equation, quasi-Newton method, smoothing method, aggregate function.

INTRODUCTION

We consider the absolute value equation (AVE):

$$Ax - |x| = b \qquad (1)$$

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 was introduced in John (2004) and investigated in a more general context in Mangasarian (2007a).

The finite-dimensional variational inequality (VI), which is a generalization of the nonlinear complementarity problem (NCP), provides a broad unifying setting for the study of optimization and equilibrium problems and serves as the main computational framework for the practical solution of a host of continuum problems in the mathematical sciences. As were shown in Cottle et al. (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) provides a theorem of the alternatives for a more general form of AVE, $Ax + B|x| = b$, and enlightens the relation between the AVE and the interval matrix. Mangasarian (2006), the AVE is shown to be equivalent to the bilinear program, the generalized LCP, and the standard LCP if 1 is not an eigenvalue of A. 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 on A and B, 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, 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 parameter-free piecewise linear concave minimization problem on a polyhedral set,

---

*Corresponding author. E-mail: yonglongquan@sohu.com.

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, 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 absolute value equation 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 et al. (2009). The method utilizes both the semismooth and the smoothing Newton steps, in which the semi-smooth 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 (2009a, b), Yong (2009, 2010), and Noor et al. (2011a, b). Yong (2010) adopted particle swarm optimization (PSO) to AVE based on aggregate function, and Noor (2011a, b) proposed iterative method for solving absolute value equations.

In this paper, we present a new method for solving AVE (1). We replace the absolute value function by a smooth one, called aggregate function. With this smoothing technique, the non-smooth AVE is formulated as a smooth nonlinear equations; furthermore, an unconstrained differentiable optimization problem. Then, we adopt quasi-Newton method to AVE. The numerical experiments show that the proposed algorithm is effective in dealing with the AVE.

In 'A smoothing function and its properties', we give a smoothing function and study its properties which will be used in the next section. Meanwhile, we give some propositions or lemmas for AVE that will be used later. In 'Newton method for AVE' we describe and present smoothing Newton method to AVE. Effectiveness of the method is demonstrated in 'Computational results' by solving some randomly generated AVE problems with singular values of A exceeding 1.

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 $\|x\|$, while $|x|$ 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 (I and e are suitable dimension in context). A vector of zeros in a real space of arbitrary dimension will be denoted by **0**.

For $x = (x_1, x_2, \cdots, x_n)^T \in R^n$, $x_{\min} = \min\{x_1, x_2, \cdots, x_n\}$, that is, the minimal component of x. $X = diag\{x_i\}$ for the diagonal matrix whose elements are the coordinates x_i of $x \in R^n$.

A SMOOTHING FUNCTION AND ITS PROPERTIES

Defining $H := R^n \to R^n$ 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. Here, we give a smoothing function of H and study its properties. We first give some properties of H which will be used in 'Newton method for AVE'.

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

Proposition 1 (Existence of AVE 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 $S = \left\{ x \mid (A+I)x - b \geq 0, (A-I)x - b \geq 0 \right\}$.

(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 (Mangasarian, 2007a).

Proposition 2 (Unique solvability of AVE)

(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 $\|A^{-1}\| < 1$ (Mangasarian, 2006).

Proposition 3 (Existence of nonnegative solution)

Let $A \geq 0$, $\|A\| < 1$ and $b \leq 0$, then a nonnegative

solution to the AVE (1) exists (Mangasarian, 2006).

Proposition 4

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 (Jiri, 2009b).

Lemma 1

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.
(iii) $\|A^{-1}\| < 1$.

Lemma 2

Suppose that A is nonsingular and $\|A^{-1}B\| < 1$. Then, $A + B$ is nonsingular.(Stewart, 1973).

Proof

We first show that $I + A^{-1}B$ is nonsingular. For, if not, then for some non-zero vector $x \in R^n$ we have that $(I + A^{-1}B)x = 0$, which shows $\|x\| \le \|A^{-1}Bx\| \le \|A^{-1}B\|\|x\|$, so $\|A^{-1}B\| \ge 1$, too, which gives the contradiction. Since $(A+B) = A(I + A^{-1}B)$ is nonsingular, we have $A + B$ is nonsingular.

Lemma 3

Let $D = diag(d)$ with $d_i \in [-1, 1], i = 1, 2, \cdots, n$. Suppose $\|A^{-1}\| < 1$. Then, $A + D$ is nonsingular.

Proof

Since $\|A^{-1}D\| \le \|A^{-1}\|\|D\| < \|D\| \le 1$, by Lemma 2, we have $A + D$ is nonsingular.

Definition 1

A function $H_\mu := R^n \to R^n, \mu > 0$ is called a uniformly smoothing approximation function of a non-smooth function $H := R^n \to R^n$ if, for any $x \in R^n$, H_μ is continuously differentiable, and there exists a constant κ such that (Qi, 2002):

$$\|H_\mu(x) - H(x)\| \le \kappa \mu, \quad \forall \mu > 0.$$

Where $\kappa > 0$ is constant that does not depend on x.

Obviously, absolute value function $|x|$ is non-differentiable. Let $\varphi(x) := |x|$. Since $\varphi(x_i) := |x_i| = \max\{x_i, -x_i\}, i = 1, 2, \cdots, n$, we can adopt the aggregate function introduced in LI (1994) to smooth the max function. The smoothing approximation function to the function $\varphi(x_i)$ is derived as:

$$\varphi_\mu(x_i) = \mu \ln\left\{ \exp\left(\frac{x_i}{\mu}\right) + \exp\left(-\frac{x_i}{\mu}\right) \right\}, i = 1, 2, \cdots, n.$$

According to Theorem 3 of LI (1994), we have:

$$0 \le \varphi_\mu(x_i) - \varphi(x_i) \le \ln 2 \cdot \mu, \quad i = 1, 2, \cdots, n.$$

Thus $\varphi_\mu(x_i)$ is a uniformly smoothing approximation function of $\varphi(x_i)$.

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

Defining $H_\mu := R^n \to R^n$ by:

$$H_\mu(x) = Ax - \varphi_\mu(x) - b. \tag{3}$$

Clearly, H_μ is a smoothing function of H. Now we give some properties of H_μ, which will be used in 'Newton method for AVE'.

By simple computation, for any $\mu > 0$, the Jacobian of H_μ at $x \in R^n$ is

$$H_\mu'(x) = A - diag\left(\frac{\exp\left(\frac{x_i}{\mu}\right) - \exp\left(-\frac{x_i}{\mu}\right)}{\exp\left(\frac{x_i}{\mu}\right) + \exp\left(-\frac{x_i}{\mu}\right)}, i = 1, 2, \cdots, n \right).$$

Theorem 1

Suppose that $\|A^{-1}\| < 1$. Then $H_\mu'(x)$ is nonsingular for

any $\mu > 0$.

Proof

Note that for any $\mu > 0$,

$$\left|\frac{\exp\left(\frac{x_i}{\mu}\right)-\exp\left(-\frac{x_i}{\mu}\right)}{\exp\left(\frac{x_i}{\mu}\right)+\exp\left(-\frac{x_i}{\mu}\right)}\right|=\left|1-2\frac{\exp\left(-\frac{x_i}{\mu}\right)}{\exp\left(\frac{x_i}{\mu}\right)+\exp\left(-\frac{x_i}{\mu}\right)}\right|<1,$$

$i=1,2,\cdots,n$.

Hence, by Lemma 3, we have $H_\mu^{'}(x)$ is nonsingular.

Theorem 2

Let $H(x)$ and $H_\mu(x)$ be defined as (2) and (3), respectively. Then, $H_\mu(x)$ is a uniformly smoothing approximation function of $H(x)$.

Proof

For any $\mu > 0$,

$$\left\|H_\mu(x)-H(x)\right\|=\left\|\varphi_\mu(x)-\varphi(x)\right\|=\sqrt{\sum_{i=1}^{n}\left|\varphi_\mu(x_i)-\varphi(x_i)\right|^2}\le\sqrt{n}\ln 2\cdot\mu .$$

If we denote $x(\mu)$ is the solution of (3), then $x(\mu)$ converges to the solution of (1) as μ goes to zero.

Define $\theta := R^n \to R$ by $\theta(x) = \frac{1}{2}\|H(x)\|^2$.

For any $\mu > 0$, Define $\theta_\mu := R^n \to R$ by $\theta_\mu(x) = \frac{1}{2}\|H_\mu(x)\|^2$.

We can get the following theorem.

Theorem 3

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

Proof

For any $\mu > 0$ and $x\in R^n$, $\nabla\theta_\mu(x) = [H_\mu^{'}(x)]^T H_\mu(x)$.

By Theorem 1, $H_\mu^{'}(x)$ is nonsingular. Hence, if $\nabla\theta_\mu(x) = 0$, then $H_\mu(x)$ and $\theta_\mu(x) = 0$. Following, we give quasi-Newton method for solving $H_\mu(x) = 0$.

NEWTON METHOD FOR AVE

Here, we give a smoothing Newton method for solving $H_\mu(x) = 0$. First we state this algorithm as the following.

Algorithm 1 Quasi-Newton method for AVE

Step 1

Given $\mu_0 > 0, k = 0$. Establish the objective function $\theta_{\mu_k}(x) = \frac{1}{2}\|H_{\mu_k}(x)\|^2$.

Step 2

Apply quasi-Newton method to solve $\min_x \theta_{\mu_k}(x)$. Let $x_k = \arg\min_x \theta_{\mu_k}(x)$.

Step 3

Check whether the stopping rule is satisfied. If satisfied, stop.

Step 4

Let $\mu_{k+1} = \mu_k +(1-e^{\mu_k})/e^{\mu_k}$, $k := k+1$. Return to step 2.

Remark

(i) The most popular quasi-Newton algorithm is the BFGS method, named for its discoverers Broyden, Fletcher, Goldfarb, and Shanno. Quasi-Newton algorithm needs to be given initial point. In the algorithm, we can set the initial point or randomly select initial point.
(ii) Unconstrained optimization algorithms often use BFGS or DFP method proposed by Jorge Nocedal et al. (1999). These algorithms are simple and easy. If we use Matlab's function *fminunc* for solving the optimization problem, the user simply provides a objective function $\theta_{\mu_k}(x)$ for each subroutine. The default of function *fminunc* is the BFGS method. If you want to use the DFP algorithm in the program, add Optimset.Hessupdate = 'dfp'.
(iii) The origin of formula $\mu_{k+1} = \mu_k +(1-e^{\mu_k})/e^{\mu_k}$ is the

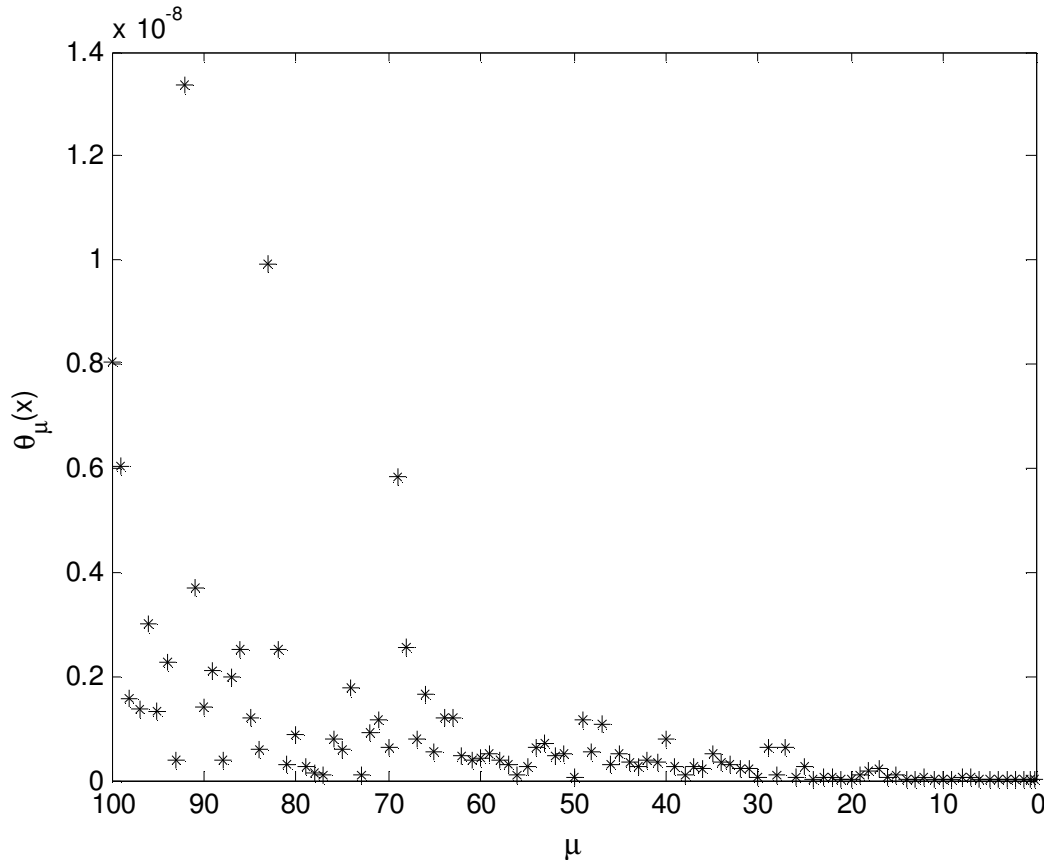**Figure 1.** Reductions of $\min \theta_{\mu_k}(x)$.

Newton iteration of equation $e^{\mu}-1=0$. So μ_k converges to zero quadratically.

COMPUTATIONAL RESULTS

Example 1

First we consider one AVE problem where the data (*A, b*) are:

$$A = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \qquad b = \begin{bmatrix} 4 \\ 5 \\ 5 \\ 4 \end{bmatrix}.$$

Since singular values svd(A)=[5.6180, 4.6180, 3.3820, 2.3820], the AVE is uniquely solvable by Proposition 2. Establish the objective function:

$$\theta_{\mu}(x) = \frac{1}{2} \left\{ \begin{bmatrix} \left[4x_1 + x_2 - \mu \ln\left(\exp\left(\frac{x_1}{\mu}\right) + \exp\left(-\frac{x_1}{\mu}\right) \right) - 4 \right]^2 + \left[x_1 + 4x_2 - x_3 - \mu \ln\left(\exp\left(\frac{x_2}{\mu}\right) + \exp\left(-\frac{x_2}{\mu}\right) \right) - 5 \right]^2 \\ + \left[x_2 + 4x_3 - x_4 - \mu \ln\left(\exp\left(\frac{x_3}{\mu}\right) + \exp\left(-\frac{x_3}{\mu}\right) \right) - 5 \right]^2 + \left[x_3 + 4x_4 - \mu \ln\left(\exp\left(\frac{x_4}{\mu}\right) + \exp\left(-\frac{x_4}{\mu}\right) \right) - 4 \right]^2 \end{bmatrix} \right\}$$

Let $\mu_0 = 100$, and select initial point randomly. Solving by Algorithm 1, the reductions of objective function $\min \theta_{\mu_k}(x)$ are shown in Figure 1.

From Figure 1, we can conclude that the unique solution to this AVE problem can be obtained after 102

iterations, and the unique solution is $x^* = [1,1,1,1]^T$.

If we let $\mu_0 = 1$, and select initial point randomly, then after 2 iterations, the unique solution $x^* = [1,1,1,1]^T$ can be obtained.

```
n=input('dimension of matrix A=')
    rand('state',0);
A1=zeros(n,n);
for i=1:n
  for j=1:n
    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(n))*ones(n,1)
```

Figure 2. Generating data (A, b) by the Matlab scripts.

Example 2

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 n 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 (Noor et al., 2011a, b).

Here the data (A, b) can be generated by Matlab scripts (Figure 2) and we set the random-number generator to the state of 0 so that the same data can be regenerated. Let $\mu_0 = 1$, and select initial point randomly. Numerical results of this problem are presented in Table 1.

All the experiments were performed on Windows XP system running on a Hp540 laptop with Intel(R) Core(TM) 2×1.8 GHz and 2 GB RAM, and the codes were written in Matlab 7.1. In all instances, the Algorithm 1 performs extremely well, and finally converges to an optimal solution of the AVE after few iterations.

Conclusion

We have proposed a new smooth method for solving the NP-hard absolute value equation $Ax - |x| = b$ under the less stringent condition that the singular values of A exceed 1. The effectiveness of the algorithm is demonstrated by its ability to solve some randomly generated problems. Smoothing Newton method is two

Table 1. Computational results by Algorithm 1.

Dimension	Iterations	Dimension	Iterations
4	2	64	3
8	2	128	3
16	2	256	3
32	2	512	3

time faster than the iterative method proposed by Noor et al. (2011a, b). Possible future work may consist of investigating other optimization algorithm and improvement of the proposed algorithm here.

REFERENCES

Cottle RW, Dantzig G (1968). Complementary pivot theory of mathematical programming. Linear Algebra Appl., 1: 103-125.
Cottle RW, Pang JS, Stone RE (1992). The Linear Complementarity Problems. Academic Press, New York.
Jiri R (2004). A Theorem of the Alternatives for the Equation Ax+B|x|=b. Linear and Multilinear Algebra, 52: 421-426.
Jiri R (2009a). An Algorithm for Solving the Absolute Value Equation. Electron. J. Linear Algebra, 18: 589-599.
Jiri R (2009b). On Unique Solvability of the Absolute Value Equation. Optim. Lett., 4:603-606.
Jorge N, Stephen W (1999). Numerical Optimization. Springer.
LI Xing-si (1994).An efficient method for non-differentiable optimization problem. Science in China-Series A. 24:371-377.
Louis C, Biao Q, Guanglu Z (2011). A globally and quadratically convergent method for absolute value equations. Comput. Optim. Appl., 48(1): 45-58.
Mangasarian OL (2007a). Absolute value programming. Comput. Optim. Appl., 36: 43-53.
Mangasarian OL (2007b). Absolute value equation solution via concave minimization. Optim. Lett., 1:3-8.
Mangasarian OL (2009a). A generalized Newton method for absolute value equations. Optim. Lett., 3:101-108.
Mangasarian OL (2009b). Knapsack feasibility as an absolute value equation solvable by successive linear programming. Optim. Lett., 2: 161-170.
Mangasarian OL, Meyer RR (2006). Absolute value equations. Linear Algebra Appl., 419: 359-367.
Muhammad AN (2011a). Javed Iqbal, Khalida Inayat Noor and Eisa Al-Said. On an iterative method for solving absolute value equations. Optim. Lett. DOI: 10.1007/s11590-011-0332-0.
Muhammad AN, Javed I, Sanjay K, Eisa Al-S (2011b). A new iterative method for solving absolute value equations. Int. J. Phys. Sci., 6(7): 1793-1797.
Oleg P (2009). On equivalent reformulations for absolute value equations. Comput. Optim. Appl., 3: 363-372.
Qi L, Sun D (2002). Smoothing functions and smoothing Newton

method for complementarity and variational inequality problems. J. Optim. Theory Appl., 113:121-147.

Shen-Long H, Zheng-Hai H (2009). A note on absolute value equations. Optim. Lett., 3:417-424.

Stewart GW (1973). Introduction to Matrix Computations. Academic Press, San Diego.

Zhang C, Wei QJ (2009). Global and Finite Convergence of a Generalized Newton Method for Absolute Value Equations. J. Optim. Theory Appl., 143: 391-403.