NUMERICAL COMPARISONS OF SMOOTHING FUNCTIONS FOR OPTIMAL CORRECTION OF AN INFEASIBLE SYSTEM OF ABSOLUTE VALUE EQUATIONS

Fakhrodin Hashemi and Saeed Ketabchi*

Department of Applied Mathematics
Faculty of Mathematical Sciences
University of Guilan, Rasht, Iran

(Communicated by Zhong Wan)

ABSTRACT. Optimal correction of an infeasible system of absolute value equations (AVEs), leads into a nonconvex and nonsmooth fractional problem. Using Dinkelbach’s approach, this problem can be reformulated to form a single variable equation. In this paper, first, we have smoothed the equation by considering four important and famous smoothing functions (see [2, 23]) and thus, to solve it, a smoothing-type algorithm based on the Difference of Convex (DC) algorithm-Newton methods is proposed. Finally, the randomly generated AVEs were compared to find the best smoothing function.

1. Introduction. The system of absolute value equations (AVEs) are considered in the form of

\[ Ax + B|x| = b, \]  

(1)

where \( A, B \in \mathbb{R}^{n \times n}, \ B \neq 0 \) and \( b \in \mathbb{R}^{n} \). Here, \( |x| \) is the absolute value of each component of the vector \( x \in \mathbb{R}^{n} \).

System (1) was first introduced by Rohn in [22] and recently has been investigated in the literature [6,7,9,12–17,21,23].

However, in many models, we often encounter problems which are presented as systems of infeasible AVEs. Optimal correction of this given system in the special case, when \( B = -I \) is considered by Ketabchi and et al in [8,10,11,18], which leads into a nonconvex and nonsmooth problem as follows :

\[ \min_{x \in \mathbb{R}^{n}} \frac{\|Ax - |x| - b\|^2}{1 + \|x\|^2}. \]  

(2)

Based on the results obtained for the optimal correction of the system (1) for a special case in which \( B = -I \), it can be concluded that the general form for the optimal correction of the system(1) is as follows:

\[ \min_{x \in \mathbb{R}^{n}} \frac{\|Ax + B|x| - b\|^2}{1 + \|x\|^2}. \]  

(3)

2010 Mathematics Subject Classification. Primary: 65D10; Secondary: 34M03, 90C25, 90C26.

Key words and phrases. Smoothing function, infeasible absolute value equations, difference of convex function, nonconvex and nonsmooth.

* Corresponding author: Saeed Ketabchi.
To solve problem (2), a few algorithms were applied without considering nonconvexity property of the objective function [10,11,18].

Furthermore, it should be noted that since the solutions of problems (2) and (3) are with a large norm to control the solutions of the Tikhonov regularization method was proposed by Moosaei and et al [18].

Based on the Dinkelbach’s approach [4], Problem (3) can be reformulated to form a single variable equations. This means that we deal with the solution of the unconstrained optimization problem with a nonconvex and nonsmooth objective function.

By considering nonconvexity and nonsmoothness properties of the objective function, we rewrite it as a smoothed Difference of Convex (DC) function based on the proposed four smoothing functions, \( \phi_1, \phi_2, \phi_3, \phi_4 \) for \( |t| \) (see [23]). Then, we present a smoothing-type algorithm (DC-Newton methods), to solve the smoothed problem whose objective function is the DC function.

The rest of this paper is organized as follows. Section 2 presents smoothing approximations for optimal correction of AVEs. Section 3 and Section 4 discuss algorithms and numerical experiments. We conclude the paper in Section 5.

Now, we describe our notations: all vectors will be column, \( \mathbb{R}^{n \times n} \) denotes the set of all \( n \times n \) real matrices. \( x, y \) are given vectors in the \( n \)-dimensional real space, \( x^\top y \) is the scalar (inner) product of \( x^\top \) and \( y \) in which \( x^\top \) is the transpose of vector \( x \). For \( x \in \mathbb{R}^n \), \( \|x\| \) denotes 2-norm. The values of \( \lambda > 0 \) and \( \mu > 0 \) are parameters of regularization and smoothing, respectively. \( \text{dom} \) is denoted to show domain of a function.

2. Smoothing approximations for optimal correction of AVEs. In order to control the optimal solution of (3), we used the Tikhonov regularization method as following:

\[
\min_{x \in \mathbb{R}^n} \omega(x) = \frac{\|Ax + B|x|-b\|^2}{1 + \|x\|^2} + \lambda\|x\|^2,
\]

where \( \lambda \) is a positive parameter.

Now, using Dinkelbach’s approach, we reformulate problem (4) to form a univariate equation. that is

\[
\varphi(t) = \min_{x \in \mathbb{R}^n} \{\|Ax + B|x|-b\|^2 + \lambda\|x\|^2(1 + \|x\|^2) - t(1 + \|x\|^2)\} = 0.
\]

It is easily proved (see [4]) that the root of Eq. 5 is equivalent to the global optimal value of problem (4).

We bring a theorem and a corollary from [18] without proof, that describe some properties of \( \varphi(t) \).

**Theorem 2.1.** The function \( \varphi \) is a strictly decreasing and concave function, and Eq. 5 has a unique root in the interval \([0, \|b\|^2]\).

**Corollary 1.** The function \( \Theta(t) = -\varphi(t) \) is subdifferentiable and \( 1 + \|x_t\|^2 \) is its subgradient at point \( t \) where \( x_t \) is a solution of the following problem

\[
\min_{x \in \mathbb{R}^n} \{\|Ax + B|x|-b\|^2 + \lambda\|x\|^2(1 + \|x\|^2) - t(1 + \|x\|^2)\}.
\]

We note that \( t^* \) is the root of \( \Theta(t) = 0 \) if and only if \( t^* = \min_{x \in \mathbb{R}^n} \omega(x) \). Hence, to find \( t^* \), the generalized Newtown method is applied. For implementing this method, we need to compute the subgradient value \( 1 + \|x_t\|^2 \). On the other hand,
to compute the mentioned subgradient, a solution for a non-convex and non-smooth inner problem of Eq. 5 is required, i.e. problem (6).

Here, we use the following four important approximations for $|t|$ (see [2, 23]):

$$
\phi_1(t, \mu) = \mu \left[ \ln(1 + e^{-\frac{t}{\mu}}) + \ln(1 + e^{\frac{t}{\mu}}) \right].
$$

$$
\phi_2(t, \mu) = \begin{cases} 
  t, & \text{if } t \geq \frac{\mu}{2}, \\
  \left( \frac{t^2}{\mu} + \frac{\mu^2}{4} \right), & \text{if } -\frac{\mu}{2} < t < \frac{\mu}{2}, \\
  -t, & \text{if } t \leq -\frac{\mu}{2}.
\end{cases}
$$

$$
\phi_3(t, \mu) = \sqrt{4\mu^2 + t^2}.
$$

$$
\phi_4(t, \mu) = \begin{cases} 
  \frac{t^2}{2\mu}, & \text{if } |t| \leq \mu, \\
  |t| - \frac{\mu^2}{2}, & \text{if } |t| > \mu.
\end{cases}
$$

In fact, for fixed $\mu > 0$, these functions have the following local behavior (see [23]):

$$
\phi_3 > \phi_1 > \phi_2 > |t| > \phi_4.
$$

In the numerical results section, we examine the behavior of these functions from the practical point of view.

These approximations are continuously differentiable and on the other hand, for fixed $\mu > 0$, their first order derivative is continuously Lipschitz (for more detail see [2, 23]). Therefore, according to the Radmacher theorem [19], it can be said that their second derivative is almost everywhere.

For computing the generalized second derivative, it is necessary to express the following definition.

**Definition 2.2.** Let $f \in C^{1,1}(\mathbb{R})$ and let $x_0 \in \mathbb{R}$. The generalized second derivative of $f$ at $x_0$, denoted by $\partial^2 f(x_0)$, is the set defined as below

$$
\partial^2 f(x_0) = \text{conv}\{\alpha | \exists x_i \rightarrow x_0 \text{ with } f \text{ twice differentiable at } x_i \text{ and } f''(x_i) \rightarrow \alpha\}.
$$

According to Definition 2.2, the generalized second derivative is computed as following:

$$
\partial^2 \phi_1(t, \mu) = \frac{1}{\mu} \left[ \frac{e^{-\frac{t}{\mu}}}{(1 + e^{-\frac{t}{\mu}})^2} + \frac{e^{\frac{t}{\mu}}}{(1 + e^{\frac{t}{\mu}})^2} \right].
$$

$$
\partial^2 \phi_2(t, \mu) = \begin{cases} 
  0, & \text{if } |t| > \frac{\mu}{2}, \\
  \frac{2}{\mu}, & \text{if } -\frac{\mu}{2} < t < \frac{\mu}{2}, \\
  \left[0, \frac{\mu}{2}\right], & \text{if } t = -\frac{\mu}{2}, \frac{\mu}{2}.
\end{cases}
$$

$$
\partial^2 \phi_3(t, \mu) = \frac{4\mu^2}{(4\mu^2 + t^2)^{\frac{3}{2}}}.
$$

$$
\partial^2 \phi_4(t, \mu) = \begin{cases} 
  0, & \text{if } |t| > \mu, \\
  \frac{1}{\mu}, & \text{if } |t| < \mu, \\
  \left[0, \frac{1}{\mu}\right], & \text{if } t = -\mu, \mu.
\end{cases}
$$

The above calculations are used in the numerical results section.
Based on smoothing functions of \( \phi_i \) for \( i = 1, 2, 3, 4 \), problem (6) is smoothed in the form

\[
\min_{x \in \mathbb{R}^n} \{ \|Ax + B\Phi_i(x) - b\|^2 + \lambda\|x\|^2(1 + \|x\|^2) - t(1 + \|x\|^2) \},
\]

(7)

where \( \Phi_i(x) = (\phi_1(x_1, \mu), \phi_2(x_2, \mu), \ldots, \phi_n(x_n, \mu)) \).

To find the solution of problem (7), in the next section, we present a smoothing-type algorithm based on DC algorithm and Newton method.

3. Algorithms and their convergence. As you know, problem (7) can be reformulated as a DC programming problem. To get the end, we state the following proposition.

**Proposition 1.** There exists \( \beta > 0 \) such that the function \( \|Ax + B\Phi_i(x) - b\|^2 + \beta\|x\|^2 \) is convex. Then problem (7) is rewritten in form of a DC programming problem as below

\[
\min_{x \in \mathbb{R}^n} f(x) = g(x) - h(x),
\]

(8)

where \( g, h \) are the defined convex functions as following

\[
\begin{align*}
g(x) &= \|Ax + B\Phi_i(x) - b\|^2 + \beta\|x\|^2 + \lambda\|x\|^2(1 + \|x\|^2), \\
h(x) &= t(1 + \|x\|^2) + \beta\|x\|^2.
\end{align*}
\]

**Proof.** It is obvious. \( \Box \)

Now, to solve problem (8), by choosing the constant parameter \( \mu_0 > 0 \) and the initial vector \( x_0 \in \text{dom} \ g \), in each iteration, we solve the following problem using the algorithm outlined below in order to obtain a local solution of problem (8).

\[
\min_{x \in \mathbb{R}^n} \tilde{f}(x) = g(x) - (\nabla h(x_0))^\top(x - x_0).
\]

(9)

The objective function of problem (9) is convex and continuously differentiable, and its gradient is continuously Lipschitz. Therefore, there is its generalized Hessian which the following proposition shows how to calculate its generalized Hessian and its gradient.

**Proposition 2.** For fixed \( \mu > 0 \) and \( i = 1, 2, 3, 4 \), set \( d_i = (\phi'_1(x_1, \mu), \phi'_2(x_2, \mu), \ldots, \phi'_n(x_n, \mu), \) \( \bar{d}_i = (\partial^2 \phi_i(x_1, \mu), \partial^2 \phi_i(x_2, \mu), \ldots, \partial^2 \phi_i(x_n, \mu)) \) \( \top, \) \( v_i = Ax + B\Phi_i(x) - b, \) \( M_i = Ax + B\text{diag}(d_i) \) and \( C_i = \text{diag}(B^\top v_i), \) Then the gradient and generalized Hessian of problem (9) are computed as following

\[
\begin{align*}
\nabla \tilde{f}(x) &= 2M_i v_i + (2\beta + 2\lambda)x + 4\lambda x\|x\|^2 - (2\beta + 2t)x_0, \\
\partial^2 \tilde{f}(x) &= 2(\text{diag}(C_i d_i) + M_i^\top M_i) + (2\beta + 2\lambda + 4\lambda\|x\|^2)I + 8\lambda xx^\top.
\end{align*}
\]

Now, we state smoothing-type algorithm based on DC algorithm and Newton method.

**Theorem 3.1.** If \( \{x_j\} \) be a generated sequence by Algorithm 1, then it is convergent to a local optimal solution for problem (8).

**Proof.** Since the objective function of problem (8) is coercive and also \( \{g(x_j) - h(x_j)\} \) is a decreasing sequence; therefore, convergence is obviously proved (for more details see [1, 24]). \( \Box \)

As it was mentioned, to find the root of Eq. 5 \( (t^*) \), we need to solve Problem (6), which we have given Algorithm 1 to solve the problem. Now, by considering the results of this algorithm, to obtain \( t^* \), Algorithm 2 is presented (see [3, 20]).
Algorithm 1 DC-Newton method
1. Choose the preliminary parameters $\delta \in (0, 1)$, $\lambda > 0$, $\mu_0 > 0$, $\theta_1, \theta_2 > 0$, $\beta > 0$, and the initial points $x_0, y_0 \in \text{dom } g$ and set $j = 0$, $\epsilon = 1e-6$.
2. While $\|\nabla f(y_{j+1})\|_\infty > \epsilon$ do.
   A) Calculate the following value
   $$y_{j+1} = y_j - \alpha (N_j + \delta I)^{-1} \nabla f(y_j),$$
   Where $\alpha$ is determined by the Armijo line search and $N_j \in \partial^2 \bar{f}(y_j)$.
   B) Update the value of $\mu$ as: If $\|\nabla f(y_{j+1})\| > \theta_1 \mu_j$, then set $\mu_{j+1} = \mu_j$; otherwise, choose $\mu_{j+1} = \theta_2 \mu_j$.
3. If $y_{j+1} = x_j$ then stop and denote $x_j$ as a solution. Otherwise, set $x_{j+1} = y_{j+1}$ and go to Step 2.
4. End

Algorithm 2 Generalized Newton method
1. Choose the values $A, B \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, accuracy parameter $\epsilon = 1e-6$ and a starting point $t_k \in [0, \|b\|^2]$ and set $k = 0$.
2. While $|t_{k+1} - t_k| \geq \epsilon$
   Solve problem (6) using Algorithm 1
   $$\min_{x \in \mathbb{R}^n} \{\|Ax + B|x| - b\|^2 + \lambda \|x\|^2(1 + \|x\|^2) - t(1 + \|x\|^2)\}. \quad (10)$$
   Calculate
   $$\Theta(t_k) = t_k - f(x(t_k)),$$
   where $f(x(t_k))$ is the value of objective function in problem (10). Set
   $$t_{k+1} = t_k - \frac{\Theta(t_k)}{1 + \|x(t_k)\|^2}.$$
3. If the condition of step 2 is held, continue; otherwise, the algorithm will be stopped.
4. Numerical experiments. In this section, the numerical results obtained from Algorithm 2 for optimal correction system (1) are reported. The infeasible system (1) in both examples 1 and 2 is randomly generated using the MATLAB code listed.

Example 1. In this example, we considered the infeasible general AVEs which is randomly generated by the following MATLAB code.

```matlab
u=10*(rand(n,1)-rand(n,1)); u=u/norm(u); k=null(u'); k=[k, zeros(n,1)]; x=spdiags(rand(n,1),0,n,n)*(rand(n,1)-rand(n,1)); x=spdiags(ones(n,1)-sign(x),0,n,n)*10*(rand(n,1)-rand(n,1)); B=spdiags(x,0,n,n); A=k+B; b=u;
```
All numerical experiments are implemented in MATLAB 2017 on a personal computer with the specification (CPU: Core i7, 12GB RAM, OS: Windows 8).

The MATLAB code given in example 1 is generated using the following lemma.

Lemma 4.1. If \( \{ x \in \mathbb{R}^n \mid (A+B)x - b \geq 0 \} = \emptyset \) or \( \{ x \in \mathbb{R}^n \mid (A-B)x - b \geq 0 \} = \emptyset \) then system (1) is infeasible.

Example 2. In this example, a special case of AVEs with \( B = -I \) is considered. This case is randomly generated using the mentioned MATLAB code.

We implemented examples 1 and 2 using \( \phi_i, i = 1, 2, 3, 4 \) and \( n = 10, 50, 100, \ldots, 5000 \), respectively. Each infeasible system, with the same dimensions is randomly generated 10 times, for testing. Then, we gave the average numerical results for some dimensions in tables 1 and 2. In these tables values \( M_{time}, M_e, M_{error} \) are computed as below

\[
M_{time} = \frac{1}{10} \left( \sum_{l=1}^{10} \text{time}_l \right), \quad M_e = \frac{1}{10} \left( \sum_{l=1}^{10} (\| (A + E^*)x^* + B|x^*| - (b + r^*) \|) \right),
\]

\[
M_{error} = \frac{1}{10} \left( \sum_{l=1}^{10} (|t_{k+1} - t_k|) \right).
\]

In the formulas above, \( \text{time} \) denotes the running time of Algorithm 2; \( x^* \) is a solution to problem (4). Finally, \( E^* \) and \( r^* \) are values which are computed in [18] as below

\[
E^* = \frac{-(Ax^* + B|x^*| - b)x^*}{1 + \| x^* \|^2}, \quad r^* = \frac{Ax^* + B|x^*| - b}{1 + \| x^* \|^2}.
\]

In tables 1 and 2, respectively, the reported values for the columns associated with \( \phi_i, i = 1, 2, 3, 4 \) indicate an error of optimal correction for the infeasible systems generated in examples 1 and 2, and the convergence error of Algorithm 2 with high accuracy. It can also be concluded that the numerical values of the mentioned columns are almost identical. But there is a huge difference in the \( M_{time} \) columns.

This difference is compared with the Dolan and Moré method [5]. In this method, we consider each \( \phi_i \) along Algorithm 2 as a solver and infeasible systems with different dimensions as test problems from set \( P \) which is randomly generated in examples 1 and 2. For each problem \( p \) and solver \( s \), set

\[
t_{p,s} = \text{time required to optimal correction problem } p \text{ by solver } s
\]
We utilize the performance ratio

\[ r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in S\}}, \]

where S is the mentioned solvers set. We assume that \( r_{p,s} \leq r_M \) is true for all selected \( p \) and \( s \). In addition, \( r_{p,s} = r_M \) if and only if the optimal correction problem \( p \) by solver \( s \) cannot be done. In order to establish a general evaluation for each solver, we define

\[ d_s(\tau) = \frac{1}{n_p} \text{size}\{p \in P : r_{p,s} \leq \tau\}, \]

which is called the performance profile of time for solver \( s \). Then, \( d_s(\tau) \) is the probability for solver \( s \in S \) that a performance ratio \( t_{p,s} \) is within a factor \( \tau \in [1,5] \) of the best possible ratio. Results of the computations are shown in figures 1 and 2.

![Figure 1. Performance profile of computing time of Algorithm 2 for the example 1.](image1.png)
Figures 1 and 2 demonstrate that the computation time of algorithm 2 is the best for smoothing function $\phi_4$, while $\phi_1$ has a lower performance with respect to $\phi_2$ and $\phi_3$.

5. **Conclusion.** In this paper, we studied the optimal correction of an infeasible system of absolute value equations. We converted the leading fractional problem to a univariate equation, using Dinkelbach’s approach. By considering four smoothing functions, $\phi_1, \phi_2, \phi_3, \phi_4$ for $|t|$, we presented a smoothing-type algorithm (DC-Newtont method), to solve it. Numerical results in tables 1 and 2 show good results for the smoothing function $\phi_4$ in small problems as well as large problems. Therefore, from the numerical results, we concluded that $\phi_4$ is the best choice of smoothing functions and the computation time of Algorithm 2 for the smoothing function $\phi_1$ is weaker than $\phi_2$ and $\phi_3$.

**REFERENCES**

[1] F. J. A. Artacho, R. M. Fleming and P. T. Vuong, Accelerating the dc algorithm for smooth functions, *Mathematical Programming*, 169 (2018), 95–118.
[2] X. Chen, Smoothing methods for nonsmooth, nonconvex minimization, *Mathematical Programming*, 134 (2012), 71–99.
[3] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Vol. 5, SIAM, 1990.
[4] W. Dinkelbach, On nonlinear fractional programming, *Management Science*, 13 (1967), 492–498.
[5] E. D. Dolan and J. J. Moré, Benchmarking optimization software with performance profiles, *Mathematical Programming*, 91 (2002), 201–213.
[6] S.-L. Hu and Z.-H. Huang, A note on absolute value equations, *Optimization Letters*, 4 (2010), 417–424.
[7] X. Jiang and Y. Zhang, A smoothing-type algorithm for absolute value equations, *Journal of Industrial and Management Optimization*, 9 (2013), 789–798.
[8] S. Ketabchi and H. Moosaei, An efficient method for optimal correcting of absolute value equations by minimal changes in the right hand side, Computers and Mathematics with Applications, 64 (2012), 1882–1885.

[9] S. Ketabchi and H. Moosaei, Minimum norm solution to the absolute value equation in the convex case, Journal of Optimization Theory and Applications, 154 (2012), 1080–1087.

[10] S. Ketabchi and H. Moosaei, Optimal error correction and methods of feasible directions, Journal of Optimization Theory and Applications, 154 (2012), 209–216.

[11] S. Ketabchi, H. Moosaei and S. Fallahi, Optimal error correction of the absolute value equation using a genetic algorithm, Mathematical and Computer Modelling, 57 (2013), 2339–2342.

[12] O. L. Mangasarian, Absolute value equation solution via concave minimization, Optimization Letters, 1 (2007), 3–8.

[13] O. L. Mangasarian, Absolute value programming, Computational Optimization and Applications, 36 (2007), 43–53.

[14] O. L. Mangasarian, A generalized newton method for absolute value equations, Optimization Letters, 3 (2009), 101–108.

[15] O. L. Mangasarian, Primal-dual bilinear programming solution of the absolute value equation, Optimization Letters, 6 (2012), 1527–1533.

[16] O. L. Mangasarian, Absolute value equation solution via dual complementarity, Optimization Letters, 7 (2013), 625–630.

[17] O. L. Mangasarian and R. R. Meyer, Absolute value equations, Linear Algebra and Its Applications, 419 (2006), 359–367.

[18] H. Moosaei, S. Ketabchi and P. M. Pardalos, Tikhonov regularization for infeasible absolute value equations, Optimization, 65 (2016), 1531–1537.

[19] A. Nekvinda and L. Zajíček, A simple proof of the rademacher theorem, Časopis pro pěstování matematiky, 113 (1988), 337–341.

[20] J. Nocedal and S. J. Wright, Numerical Optimization, 2nd edition, Springer, 2006.

[21] O. Prokopyev, On equivalent reformulations for absolute value equations, Computational Optimization and Applications, 44 (2009), 363–372.

[22] J. Rohn, A theorem of the alternatives for the equation $Ax + B|x| = b$, Linear and Multilinear Algebra, 52 (2004), 421–426.

[23] B. Saheya, C.-H. Yu and J.-S. Chen, Numerical comparisons based on four smoothing functions for absolute value equation, Journal of Applied Mathematics and Computing, 56 (2018), 131–149.

[24] P. D. Tao and L. T. H. An, Convex analysis approach to dc programming: Theory, algorithms and applications, Acta Mathematica Vietnamica, 22 (1997), 289–355.

Received June 2018; 1st revision December 2018; Final revision April 2019.

E-mail address: fhashemi@phd.guilan.ac.ir
E-mail address: sketabchi@guilan.ac.ir