Optimization Letters, , 1–8 (2008)

A Generalized Newton Method for Absolute Value Equations

O. L. MANGASARIAN olvi@cs.wisc.edu
Computer Sciences Department
University of Wisconsin
Madison, WI 53706
Department of Mathematics
University of California at San Diego
La Jolla, CA 92093

Received May 2008

Editor: Panos Pardalos

Abstract. A direct generalized Newton method is proposed for solving the NP-hard absolute value equation (AVE) $Ax-|x|=b$ when the singular values of A exceed 1. A simple MATLAB implementation of the method solved 100 randomly generated 1000-dimensional AVEs to an accuracy of 10^{-6} in less than 10 seconds each. Similarly, AVEs corresponding to 100 randomly generated linear complementarity problems with 1000×1000 nonsymmetric positive definite matrices were also solved to the same accuracy in less than 29 seconds each.

Keywords: absolute value equation, generalized Newton, linear complementarity problems

1. Introduction

We consider the absolute value equation (AVE):

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

where $A \in R^{n\times n}$ and $b \in R^n$ are given, and $|\cdot|$ denotes absolute value. A slightly more general form of the AVE, $Ax + B|x| = b$ was introduced in [11] and investigated in a more general context in [6]. The AVE (1) was investigated in detail theoretically in [7] and a bilinear program was prescribed there for the special case when the singular values of A are not less than one. No computational results were given in either [7] or [6]. In contrast in [5], computational results were given for a linear-programming-based successive linearization algorithm utilizing a concave minimization model. As was shown in [7], the general NP-hard linear complementarity problem (LCP) [3, 4, 2], which subsumes many mathematical programming problems, can be formulated as an AVE (1). This implies that (1) is NP-hard in its general form.

In Section 2 of the present work we propose a generalized Newton algorithm that is globally convergent under certain assumptions. Effectiveness of the method is demonstrated in Section 3 by solving 100 randomly generated 1000-dimensional AVEs with singular values of A exceeding 1. Each AVE is solved to an accuracy of 10^{-6} in time that is 26 times faster than that of the successive linearization algorithm of [5]. The generalized Newton method also solved a hundred 1000-dimensional randomly generated positive definite linear complementarity problems to the same accuracy of 10^{-6}. Section 4 concludes the paper with some open questions.

A word about our notation and background material. The scalar product of two vectors x and y in the n-dimensional real space will be denoted by $x'y$. For $x \in R^n$, the norm $\|x\|$ will denote the 2-norm $(x'x)^{\frac{1}{2}}$, $|x|$ will denote the vector in R^n of absolute values of components of x and $sign(x)$ will denote a vector with components equal to 1, 0 or -1 depending on whether the corresponding component of x is positive, zero or negative. In addition, $diag(sign(x))$ will denote a diagonal matrix corresponding to $sign(x)$. The plus function x_+, which replaces the negative components of x by zeros, is a projection operator that projects x onto the nonnegative orthant. Note that the absolute value function can be written as $|x| = x_+ + (-x)_+$. A generalized Jacobian $\partial|x|$ of $|x|$ based on a subgradient [10, 9] of its components is given by the diagonal matrix $D(x)$:

$$D(x) = \partial|x| = diag(sign(x)). \tag{2}$$

For an $m \times n$ matrix A, A_i will denote the ith row of A. The identity matrix in a real space of arbitrary dimension will be denoted by I, while a column vector of ones of arbitrary dimension will be denoted by e. For a solvable matrix equation $By = d$ we shall use the MATLAB backslash notation $B\backslash d$ to denote a solution y. Similarly for $A \in R^{n\times n}$, $svd(A)$ will denote the n singular values of A and $eig(A)$ will denote its n eigenvalues.

2. The Generalized Newton Method

We begin by defining the piece-wise linear vector function $g(x)$ specified by the AVE (1) as follows:

$$g(x) = Ax - |x| - b. \tag{3}$$

A generalized Jacobian $\partial g(x)$ of $g(x)$ is given by:

$$\partial g(x) = A - D(x), \tag{4}$$

where $D(x) = diag(sign(x))$ as defined in (2). The generalized Newton method for finding a zero of the equation $g(x) = 0$ consists then of the following iteration:

$$g(x^i) + \partial g(x^i)(x^{i+1} - x^i) = 0. \tag{5}$$

Replacing $g(x^i)$ by its definition (3) and setting $\partial g(x^i) = A - D(x^i)$ gives:

$$Ax^i - |x^i| - b + (A - D(x^i))(x^{i+1} - x^i) = 0. \tag{6}$$

Noting that $D(x^i)x^i = |x^i|$, the generalized Newton iteration (6) simplifies to the following:

$$(A - D(x^i))x^{i+1} = b. \tag{7}$$

Solving for x^{i+1} gives:

$$x^{i+1} = (A - D(x^i))\backslash b, \tag{8}$$

which is our final simple **generalized Newton iteration** for solving the AVE (1).

We shall need a few theoretical results to establish convergence of the iteration (8). We first quote the following result from [7, Lemma 1].

Lemma 1 *The singular values of the matrix $A \in R^{n\times n}$ exceed 1 if and only if the minimum eigenvalue of $A'A$ exceeds 1.*

The following useful consequence of the above lemma gives sufficient conditions that the Newton iteration (8) is well defined.

Lemma 2 *If the singular values of $A \in R^{n\times n}$ exceed 1 then $(A-D)^{-1}$ exists for any diagonal matrix D whose diagonal elements equal ± 1 or 0.*

Proof: If $(A - D)$ is singular then:

$$(A - D)x = 0 \text{ for some } x \neq 0. \tag{9}$$

We then have the following contradiction where the first inequality follows from Lemma 1:

$$x'x < x'A'Ax = x'DAx = x'DDx \leq x'x. \tag{10}$$

Hence $(A - D)$ is nonsingular. ■

We now establish boundedness of the Newton iterates of (8) and hence the existence of an accumulation for these iterates.

Proposition 3 Boundedness of Newton Iterates *Let the singular values of A exceed 1. Then, the iterates $x^{i+1} = (A-D(x^i))\backslash b$ of the generalized Newton method (8) are well defined and bounded. Consequently, there exists an accumulation point $\bar{x}$ such that $(A-\bar{D})\bar{x} = b$ for some diagonal matrix $\bar{D}$ with diagonal elements of ± 1 or 0.*

Proof: By Lemma 2 $(A - D(x^i))^{-1}$ exists. Hence, the generalized Newton iteration $x^{i+1} = (A - D(x^i))\backslash b$ is well defined. Suppose now that the sequence $\{x^i\}$ is unbounded. Then there exists a subsequence $\{x^{i_j+1}\} \to \infty$ with nonzero x^{i_j+1} such that $D(x^{i_j}) = \tilde{D}$, where $\tilde{D}$ is a fixed diagonal matrix with diagonal elements equal to ± 1 or 0 extracted from the finite number of possible configurations for $D(x^i)$ in the sequence $\{D(x^i)\}$, and such that the bounded subsequence $\{\frac{x^{i_j+1}}{\|x^{i_j+1}\|}\}$ converges to $\tilde{x}$. Hence,

$$(A - \tilde{D})\frac{x^{i_j+1}}{\|x^{i_j+1}\|} = \frac{b}{\|x^{i_j+1}\|}. \tag{11}$$

Letting $j \to \infty$ gives:

$$(A - \tilde{D})\tilde{x} = 0, \quad \|\tilde{x}\| = 1, \tag{12}$$

since $x^{i_j+1} \to \infty$. This however contradicts the nonsingularity of $(A - \tilde{D})$ which follows from Lemma 2. Consequently, the sequence $\{x^i\}$ is bounded and there exists an accumulation point $(\bar{D}, \bar{x})$ of $\{D(x^i), x^{i+1}\}$ such that $\bar{x} = (A - \bar{D})\backslash b$, or equivalently $(A - \bar{D})\bar{x} = b$. ■

Under a somewhat restrictive assumption we can establish finite termination of the generalized Newton iteration at an AVE solution as follows.

Proposition 4 Finite Termination of the Newton Iteration *Let the singular values of A exceed 1. If $D(x^{i+1}) = D(x^i)$ for some i for the well defined generalized Newton iteration (8), then x^{i+1} solves the AVE (1).*

Proof: By Lemma 2 the generalized Newton iteration $x^{i+1} = (A - D(x^i))\backslash b$ is well defined, and if $D(x^{i+1}) = D(x^i)$ then:

$$0 = (A - D(x^i))x^{i+1} - b = Ax^{i+1} - D(x^{i+1})x^{i+1} - b = Ax^{i+1} - |x^{i+1}| - b. \quad (13)$$

Hence x^{i+1} solves the AVE (1) ■

For our final result, Proposition 7 below, which establishes linear convergence of the generalized Newton iteration (8), we first establish the following two lemmas.

Lemma 5 Lipschitz Continuity of the Absolute Value *Let x and y be points in R^n. Then:*

$$\| \, |x| \, - \, |y| \, \| \leq 2\|x - y\|. \quad (14)$$

Proof: The result follows from the following string of equalities and inequalities.

$$\begin{array}{rcl} \| \, |x| \, - \, |y| \, \| & = & \|x_+ + (-x)_+ - y_+ - (-y)_+\| \\ & \leq & \|x_+ - y_+\| + \|(-x)_+ - (-y)_+\| \\ & \leq & \|x - y\| + \|(-x) - (-y)\| \\ & = & 2\|x - y\|. \end{array} \quad (15)$$

The first inequality above follows from from the triangle inequality, and the second inequality from the nonexpansive property of the projection operator $(\cdot)_+$ [1, Proposition 2.1.3]. ■

Lemma 6 Linear Convergence of the Newton Iteration *Under the assumption that $\|(A - D)^{-1}\| < \frac{1}{3}$ for any diagonal matrix D with diagonal elements of ± 1 or 0, the generalized Newton iteration (8) converges linearly from any starting point to a solution $\bar{x}$ for any solvable AVE (1).*

Proof: Let $\bar{x}$ be a solution of the AVE (1). To simplify notation, let $\bar{D} = D(\bar{x}) = diag(sign(\bar{x}))$ and $D^i = D(x^i) = diag(sign(x^i))$. Noting that $|\bar{x}| = \bar{D}\bar{x}$ and $|x^i| = D^i x^i$ we have the following upon subtracting $(A - \bar{D})\bar{x} = b$ from $(A - D^i)x^{i+1} = b$:

$$\begin{array}{rcl} A(x^{i+1} - \bar{x}) = D^i x^{i+1} - \bar{D}\bar{x} & = & D^i(x^{i+1} + x^i - x^i) - \bar{D}\bar{x} \\ & = & |x^i| - |\bar{x}| + D^i(x^{i+1} - x^i) \\ & = & |x^i| - |\bar{x}| + D^i(x^{i+1} - \bar{x} + \bar{x} - x^i) \end{array} \quad (16)$$

Hence,

$$(A - D^i)(x^{i+1} - \bar{x}) = |x^i| - |\bar{x}| - D^i(x^i - \bar{x}). \tag{17}$$

Consequently:

$$(x^{i+1} - \bar{x}) = (A - D^i)^{-1}(|x^i| - |\bar{x}| - D^i(x^i - \bar{x})), \tag{18}$$

and by Lemma 5:

$$\|x^{i+1} - \bar{x})\| \leq \|(A - D^i)^{-1}\|(2\|x^i - \bar{x}\| + \|x^i - \bar{x}\|). \tag{19}$$

Hence

$$\|x^{i+1} - \bar{x})\| \leq 3\|(A - D^i)^{-1}\|(\|x^i - \bar{x}\|) < \|x^i - \bar{x}\|, \tag{20}$$

where the last inequality of (20) follows from $\|(A-D^i)^{-1}\| < \frac{1}{3}$. Hence the sequence $\{\|x^i - \bar{x}\|\}$ converges linearly to zero and $\{x^i\}$ converges linearly to $\bar{x}$. ■

We are now ready to prove our final result.

Proposition 7 Sufficient Conditions for the Linear Convergence of the Newton Iteration *Let $\|A^{-1}\| < \frac{1}{4}$ and $\|D(x^i)\| \neq 0$. Then, the AVE (1) is uniquely solvable for any b and the generalized Newton iteration (8) is well defined and converges linearly to the unique solution of AVE from any starting point x^0.*

Proof: The unique solvability of the AVE (1) for any b follows from [7, Proposition 4] which requires that $\|A^{-1}\| < 1$. By the Banach perturbation lemma [8, Page 45], $\|(A - D(x^i))^{-1}\|$ exists for any x^i since A^{-1} exists and $\|A^{-1}\| \cdot \|D(x^i)\| < 1$. We also have by the same lemma that:

$$\|(A - D(x^i))^{-1}\| \leq \frac{\|A^{-1}\| \cdot \|D(x^i)\|}{1 - \|A^{-1}\| \cdot \|D(x^i)\|} < \frac{\frac{1}{4} \cdot 1}{1 - \frac{1}{4} \cdot 1} = \frac{1}{3}. \tag{21}$$

Hence by Lemma 6 above the sequence $\{x^i\}$ converges linearly to the unique solution of AVE from any starting point x^0. ■

Remark 8 *We note that, just like the ordinary Newton method for solving nonlinear equations, our generalized Newton iteration (8) for solving the nondifferentiable AVE (1) requires at the very least that the AVE (1) is solvable. Without such a solvability assumption, the generalized Newton iteration (8) will not converge. As a trivial example consider the AVE in R^1: $x - |x| = 1$ which has no solution. For if it did, it would lead to the contradiction: $-1 = |x| - x \geq 0$. For this example, our generalized Newton iteration (8) degenerates to $x^{i+1} = \frac{1}{1-sign(x^i)}$. We thus have $x^{i+1} = \infty$ for $x^i > 0$. For $x^i = 0$ we have $x^{i+1} = 1$ and $x^{i+2} = \infty$, while for $x^i < 0$ we have $x^{i+1} = \frac{1}{2}$ and $x^{i+2} = \infty$.*

We turn now to our computational testing.

3. Computational Results

We used MATLAB 7.1 to test our generalized Newton iteration (8) on 100 consecutively generated solvable random AVEs (1) with fully dense matrices $A \in R^{1000\times 1000}$, as well as on 100 AVEs based on consecutively generated random solvable 1000-dimensional linear complementarity problems.

For the first group of explicitly generated AVEs we first chose a random A from a uniform distribution on [-10,10], then we chose a random x from a uniform distribution on [-1,1]. Finally we computed $b = Ax - |x|$. We ensured that the singular values of each A exceeded 1 by actually computing the minimum singular value and rescaling A by dividing it by the minimum singular value multiplied by a random number in the interval [0,1]. Our method solved all 100 AVEs to an accuracy of 10^{-6} in total time of 940.30 seconds on 3.00Ghz Pentium 4 processor running Linux Release 5.1. In comparison the successive linear programming algorithm of [5] solved 95 out of 100 similarly sized AVEs on the same machine in $24,773$ seconds to the same accuracy. We also note that our method converged in very few iterations, typically 5, under the mere condition that the singular values of A exceed 1 which ensures that the generalized Newton iteration (8) is well defined. Note that we do not impose the more stringent conditions of Proposition 7 on our examples.

We also tested our method on AVEs based on randomly generated linear complementarity problems. As pointed out in [7, Proposition 2] every linear complementarity problem (LCP):

$$z \geq 0,\ Mz + q \geq 0,\ z'(Mz + q) = 0, \tag{22}$$

for a given $M \in R^{n\times n}$ and $q \in R^n$, can be reduced to the following AVE:

$$(M - I)^{-1}(M + I)x - |x| = (M - I)^{-1}q, \tag{23}$$

with

$$x = \frac{1}{2}((M - I)z + q), \tag{24}$$

provided 1 is not an eigenvalue of M, which can be easily achieved by rescaling M and q. Using the above relations we solved a hundred 1000-dimensional linear complementarity problems generated as follows. The LCP matrices were nonsymmetric positive definite matrices that were the sum of products of random matrices whose elements were chosen from a uniform distribution on $[-10, 10]$. A random solution z was picked for each LCP whose elements where chosen from a uniform distribution on $[0, 5]$ and such that half the components of z were zero. Finally q was chosen such that the LCP was solvable. The corresponding AVE solved for each LCP was that specified by (23).

Table 1 gives a summary of our computational results. We note the following:

(i) All 100 consecutive instances of both AVEs and LCPs were solved without failure to an accuracy of 10^{-6} in satisfying $\|Ax - |x| - b\| = 0$.

(ii) The average number of Newton iterations was 5 for the 100 AVEs and 8.06 for the 100 LCPs.

(iii) The average time for generating and solving each of the 100 AVEs was 9.40 seconds, and 28.61 seconds for each LCP.

(iv) The overall minimum of the singular values of each A for the 100 AVEs was 288.37 and 1.0 for the LCPs.

Table 1. **Generalized Newton method results for 100 consecutive random AVEs and LCPs each with n=1000.**

Problem	AVE	LCP
Properties	svd(A)>1	M pos. def. non-symm.
Problem size n	1000	1000
Number of problems solved	100	100
$\|Ax - \|x\| - b\| <$ accuracy	10^{-6}	10^{-6}
Total overall Newton iterations	500	806
Total time in seconds	940.30	2860.6
Overall min of singular values of all A's	288.37	1.0

4. Conclusion

We have proposed a fast linearly convergent generalized Newton method for solving the the NP-hard absolute value equation $Ax - |x| = b$ under certain assumptions on A. It turns out that the algorithm works under the less stringent condition that the singular values of A exceed 1. This ensures that the Newton iteration (8) is well defined. It would be very useful to establish convergence under this or even a more general assumption. Another interesting problem to address is that of the most general linear complementarity problem that can be solved as an absolute value equation. These questions as well as enhancement of the Newton iteration by a step size are topics for future research.

Acknowledgments

The research described in this Data Mining Institute Report 08-01, May 2008, was supported by National Science Foundation Grants CCR-0138308 and IIS-0511905.

References

1. D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1995.
2. S.-J. Chung. NP-completeness of the linear complementarity problem. *Journal of Optimization Theory and Applications*, 60:393–399, 1989.
3. R. W. Cottle and G. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications*, 1:103–125, 1968.
4. R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem.* Academic Press, New York, 1992.

5. O. L. Mangasarian. Absolute value equation solution via concave minimization. *Optimization Letters*, 1(1):3–8, 2007. ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/06-02.pdf.
6. O. L. Mangasarian. Absolute value programming. *Computational Optimization and Applications*, 36(1):43–53, 2007. ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/05-04.ps.
7. O. L. Mangasarian and R. R. Meyer. Absolute value equations. *Linear Algebra and Its Applications*, 419:359–367, 2006. ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/05-06.pdf.
8. J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, 1970.
9. B. T. Polyak. *Introduction to Optimization*. Optimization Software, Inc., Publications Division, New York, 1987.
10. R. T. Rockafellar. New applications of duality in convex programming. In *Proceedings Fourth Conference on Probability*, Brasov, Romania, 1971.
11. J. Rohn. A theorem of the alternatives for the equation $Ax+B|x|=b$. *Linear and Multilinear Algebra*, 52(6):421–426, 2004. http://www.cs.cas.cz/ rohn/publist/alternatives.pdf.