基于极大熵牛顿法求解绝对值方程*

邓永坤, 王海军, 张萍

(中国矿业大学理学院, 江苏徐州 221000)

摘要: 针对绝对值方程 $Ax - |x| = b$ 的求解问题, 在假设矩阵 A 的奇异值大于 1 时, 给出了求解绝对值方程的一个新方法。通过引入一种极大熵函数将绝对值方程进行光滑化处理, 进而把绝对值方程转换为光滑非线性方程组, 然后利用极大熵牛顿法对其进行求解。数值实验结果表明了该方法的正确性和有效性。

关键词: 绝对值方程; 极大熵函数; 光滑非线性方程组; 极大熵牛顿法

中图分类号: O242.2 **文献标志码:** A **文章编号:** 1001-3695(2012)12-4546-03

doi:10.3969/j.issn.1001-3695.2012.12.037

Maximum entropy Newton method for absolute value equations

DENG Yong-kun, WANG Hai-jun, ZHANG Ping

(School of Sciences, China University of Mining & Technology, Xuzhou Jiangsu 221000, China)

Abstract: Aimed at the solution to the absolute value equations(AVE) $Ax - |x| = b$, this paper proposed a new method for solving the AVE under the condition that all singular values of A exceed one. First, using the maximum entropy function, transformed the AVE into smooth nonlinear equations. Then used the maximum entropy Newton method to solve this problem. Numerical results indicate that the method is feasible and effective to AVE.

Key words: absolute value equations(AVE); maximum entropy function; smooth nonlinear equations; maximum entropy Newton method

0 引言

考虑以下形式的绝对值方程:

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

其中: $A \in \mathbb{R}^{n \times n}$ 为对称矩阵; $b \in \mathbb{R}^n$; 绝对值 $|x|$ 依分量形式选取。

绝对值方程问题首先由 Rohn^[1] 提出, 之后许多学者对其理论及算法进行了深入的研究。文献[2]对式(1)进行了非常详尽的理论分析, 指出式(1)等价于双线性规划、广义线性互补问题, 并且证明了当 1 不是矩阵 A 的特征值时, 该方程可以转换为线性互补问题, 同时文中给出了该方程有解、无解、唯一解、非负解及 2^n 个解的充分条件。文献[3]给出了求解该形式绝对值方程的广义牛顿法, 并且证明了该方法在适当的条件下具有线性收敛速度。

在近期的研究成果中, 文献[4] 用一个光滑函数代替绝对值函数后给出了求解该绝对值方程的一个光滑牛顿法, 并证明了该算法具有二次收敛性, 该方法是目前求解式(1)较好的成果之一; 文献[5,6]利用区间算法对式(1)进行求解, 并且分析了算法的收敛速度; 文献[7~9]基于极大熵函数光滑处理技术, 给出了求解式(1)的极大熵自适应微粒群混合算法、和声搜索算法及牛顿法; 文献[10]结合优化技术给出了一种迭代算法; 文献[11]提出改进的 Gauss-Seidel 算法对其求解, 并证明了该算法计算速度明显快于文献[10]中迭代算法。

本文约定以下符号: I 表示单位矩阵; $\|x\|$ 表示 x 的 2 范

数; A^T 表示矩阵 A 的转置; $x^T y$ 为向量 x 和 y 的内积; e 为元素均为 1 的 $n \times 1$ 的列向量; $N(x, r)$ 表示 x 的 r 邻域。

1 问题转换

定义函数:

$$H(x) := Ax - |x| - b \quad H: \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (2)$$

很明显, 假设 x 是式(1)的解, 当且仅当 $H(x) = 0$ 。

引理 1^[2] 若矩阵 A 的奇异值大于 1, 则对任意的 $b \in \mathbb{R}^n$, 方程 $Ax - |x| = b$ 存在唯一解。

注: 该引理给出了式(1)存在唯一解的条件。在本文中总是假设矩阵 A 的奇异值是大于 1 的。

接下来给出极大熵函数的定义及其性质。

定义 1 设 $f(x) = \max\{f_i(x)\} (i = 1, 2, \dots, m)$ 。其中 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, 定义函数

$$F_p(x) = p \ln \left(\sum_{i=1}^m \exp\left(\frac{1}{p} f_i(x)\right) \right) \quad (3)$$

其中: p 称为极大熵因子, 称 $F_p(x)$ 为 $f(x)$ 的极大熵函数。

引理 2^[12] 对于任意的 $x \in \mathbb{R}^n$, 极大熵函数有以下性质:

a) $F_p(x)$ 随参数 p 的减少而单调减少, 且当 $p \rightarrow 0$ 时, 函数 $F_p(x)$ 以 $f(x)$ 为极限, 即

$$F_s(x) \leq F_r(x), s \leq r, \lim_{p \rightarrow 0} F_p(x) = f(x) \quad (4)$$

b) $F_p(x)$ 与 $f(x)$ 间有以下不等式关系:

$$0 \leq F_p(x) - f(x) \leq p \ln(m) \quad (5)$$

其中: m 代表分量函数的个数。

收稿日期: 2012-05-13; 修回日期: 2012-06-14 基金项目: 中央高校基本科研业务费专项基金资助项目(2010LKXS01)

作者简介: 邓永坤(1988-), 男, 山东青州人, 硕士研究生, 主要研究方向为计算数学及其应用(dengyongkun@126.com); 王海军(1972-), 男, 山西长治人, 教授, 主要研究方向为计算数学及其应用; 张萍(1986-), 女, 山西运城人, 硕士研究生, 主要研究方向为数理统计。

在式(2)中,令 $\varphi(x_i) = |x_i| = \max\{x_i, -x_i\}$ ($i = 1, 2, \dots, n$),用极大熵函数对 $\varphi(x_i)$ 进行光滑化处理,可得 $\varphi(x_i)$ 的一致光滑近似函数如下:

$$\varphi_p(x_i) = p \ln \left\{ \exp\left(\frac{x_i}{p}\right) + \exp\left(-\frac{x_i}{p}\right) \right\} \quad i = 1, 2, \dots, n \quad (6)$$

由引理 2^[12] 易知:

定理 1 对于任意的 $x \in R^n$, 令 $\varphi_p(x_i) = p \ln \left\{ \exp\left(\frac{x_i}{p}\right) + \exp\left(-\frac{x_i}{p}\right) \right\}$, $\varphi(x_i) = |x_i|$, 则

a) $\varphi_p(x_i)$ 随参数 p 的减少而单调减少, 当 $p \rightarrow 0$ 时, $\varphi_p(x_i)$ 以 $\varphi(x_i)$ 为极限;

b) 存在不等式关系:

$$0 \leq \varphi_p(x_i) - \varphi(x_i) \leq p \ln 2 \quad (7)$$

对任意的 $p > 0$, 令 $\varphi_p(x) = (\varphi_p(x_1), \varphi_p(x_2), \dots, \varphi_p(x_n))^T$, 则式(2)可以转换为

$$H_p(x) := Ax - \varphi_p(x) - b \quad (8)$$

令 $E_i = \frac{\exp(\frac{x_i}{p}) - \exp(-\frac{x_i}{p})}{\exp(\frac{x_i}{p}) + \exp(-\frac{x_i}{p})}$ ($i = 1, 2, \dots, n$), 则 H_p 的雅可比矩阵为

$$H'_p(x) := A - \text{diag}(E_i) \quad (9)$$

引理 3^[9] 假设若矩阵 A 的奇异值大于 1, 则 $H'_p(x)$ 对于任意 $p > 0$ 是非奇异的。

定义 2^[13] 光滑逼近函数。给定函数 $H: R^n \rightarrow R^n$, 称光滑函数 $H_p(\cdot): R^n \rightarrow R^n$ ($p > 0$) 为 H 的光滑逼近函数, 如果对任意的 $x \in R^n$, 存在 $\kappa > 0$, 使得

$$\|H_p(x) - H(x)\| \leq \kappa p, \forall p > 0 \quad (10)$$

如果 κ 不依赖于 x , 则称 $H_p(\cdot)$ 为 H 的一致光滑逼近函数。

定理 2 式(8)为式(2)的一致光滑逼近函数, 即 $H_p(x)$ 为 $H(x)$ 的一致光滑逼近函数。

证明 对于任意 $p > 0$, 由定理 1 可知,

$$\|H_p(x) - H(x)\| = \|\varphi_p(x) - \varphi(x)\| = \sqrt{\sum_{i=1}^n |\varphi_p(x_i) - \varphi(x_i)|} \leq \sqrt{n} \ln 2 \cdot p$$

再由定义 2^[13] 可知 $H_p(x)$ 为 $H(x)$ 的一致光滑逼近函数, 证毕。

所以可以通过求解光滑非线性方程组(式(8))的解来近似求解方程组(式(2))的解, 即求得绝对值式(1)的近似解。

2 算法及收敛性分析

算法步骤如下:

a) 给定一个初始向量 $x^0 \in R^n$, 允许误差 $\varepsilon = 10^{-9}$ 及参数 $p > 0$, 构造光滑非线性方程组 $H_p(x) = Ax - \varphi_p(x) - b$, 置 $k = 0$ 。

b) 计算 x^{k+1} , 通过 $\begin{cases} x^{k+1} = x^k + \Delta x^k \\ H'_p(x^k) \Delta x^k + H_p(x^k) = 0 \end{cases}, k = 0, 1, 2, \dots$ 。

c) 若 $\|x^{k+1} - x^k\| \leq \varepsilon$, 则停止, 得到解 $x^* = x^k$; 否则转入步骤 b)。

接下来分析算法的收敛性及收敛速度。

引理 4^[14] 若 $F: D \subset R^n \rightarrow R^n$ 在 x^* 处 F 可导, x^* 是非线性方程组 $F(x) = 0$ 的解, $S_0 \subset D$ 是 x^* 的邻域, 又设 $M: S_0 \subset D \rightarrow$

$L(R^n)$ 在 x^* 处连续, $M(x^*)$ 非奇异, 则存在闭球 $S = \bar{S}(x^*, \delta) \subset S_0$, 使映象 $G(x) = x - [M(x)]^{-1}F(x)$ 在 $x \in S$ 上有定义, 且在 x^* 处有 F 导数:

$$G'(x^*) = I - [M(x^*)]^{-1}F'(x^*) \quad (11)$$

若还有 $\rho(I - [M(x^*)]^{-1}F'(x^*)) < 1$, 则 x^* 是迭代序列 $x^{k+1} = x^k - [M(x^k)]^{-1}F(x^k)$ 的吸引点。

引理 5^[14] 假定映象 $F: D \subset R^n \rightarrow R^n$ 及 $D_0 \subset D$, 若存在常数 $c \geq 0$ 及 $p \in [0, 1]$, 对所有的 $x, y \in D_0$ 有 $\|F'(x) - F'(y)\| \leq c \|x - y\|^p$ 成立, 则有以下不等式成立:

$$\|F(y) - F(x) - F'(x)(y - x)\| \leq \frac{c}{p+1} \|y - x\|^{1+p} \quad (12)$$

引理的证明在文献[14]中有详细描述, 此处从略。

定理 3 若矩阵 A 的奇异值大于 1 且 x^* 是非线性光滑方程组式(8)的解, 则存在闭球 $S = \bar{S}(x^*, \delta)$ ($\delta > 0$), 使映象 $G(x) = x - [H'_p(x)]^{-1}H_p(x)$ 对所有 $x \in S$ 有定义, 且以上算法生成的迭代序列 $\{x^k\}$ 超线性收敛到 x^* , 如果还假定

$$\|H'_p(x) - H'_p(x^*)\| \leq \alpha \|x - x^*\| \quad \forall x \in S \quad (13)$$

其中: $\alpha > 0$ 为常数, 则迭代序列 $\{x^k\}$ 至少二阶收敛。

证明 因为 $H_p(x)$ 连续可微且 $H'_p(x)$ 非奇异, 在引理 4^[14] 中取 $M(x) = H'_p(x)$, 则映象 $G(x) = x - [H'_p(x)]^{-1}H_p(x)$ 在球 $S = \bar{S}(x^*, \delta)$ 上有定义且 $G'(x^*) = I - [H'_p(x^*)]^{-1}H'_p(x^*) = 0$, 于是 $\rho(G'(x^*)) = 0 < 1$, 由引理 4^[14] 得 $\{x^k\}$ 收敛到 x^* , 且当 $x^k \neq x^*$ 时, 有 $\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0$, 所以迭代序列 $\{x^k\}$ 是超线性收敛。

假定式(13)成立, 则由引理 5^[14] 可知: $\|H_p(x) - H_p(x^*) - H'_p(x^*)(x - x^*)\| \leq \frac{1}{2} \alpha \|x - x^*\|^2$ 于是: $\|G(x) - x^*\| = \|x - [H'_p(x)]^{-1}H_p(x) - x^*\| \leq \|[H'_p(x)]^{-1}\| (\|H'_p(x) - H'_p(x^*)\| \|x - x^*\| + \|H_p(x) - H_p(x^*) - H'_p(x^*)(x - x^*)\|) \leq 2 \|[H'_p(x^*)]^{-1}\| \left(\frac{3}{2} \alpha \|x - x^*\|^2\right) = \eta \|x - x^*\|^2$

若取 $x = x^k, G(x) = x^{k+1}$, 则有 $\|x^{k+1} - x^*\| \leq \eta \|x^k - x^*\|^2$, 即证得迭代序列 $\{x^k\}$ 的收敛阶至少为 2, 证毕。

3 数值实验

运用 MATLAB 7.0 进行编程计算, 参数的设定和符号说明如下: 计算精度 $\varepsilon = 1.0e - 9$; k 表示迭代次数; n 表示问题的维数; T 表示算法运行的 CPU 时间(计算单位为秒); x^* 表示广义绝对值方程的近似解; x^0 表示迭代初始点; MENM 表示极大熵牛顿法; MGSM 表示文献[11]中改进 Gauss-Seidel 算法; IM 表示文献[10]中迭代算法。

算例 1 矩阵 A 记为: $A = (a_{ij})_{n \times n}$, 矩阵 A 按以下要求构造: $a_{ii} = 4n, a_{i,i+1} = a_{i+1,i}, a_{ij} = 0.5 (i = 1, 2, \dots, n)$; 令 $b = (A - I)e$, 则 $x = (1, 1, \dots, 1)^T$ 为绝对值方程 $Ax - |x| = b$ 的准确解。

a) 选取不同初始点及 p , 利用本文算法进行计算, 其结果如表 1~4 所示。

b) 选取文献[10, 11]中初始点 $x^0 = (x_1, x_2, \dots, x_n)^T, x_i = 0.001i$, 令 $p = 0.1$ (或 $p = 0.01$), 利用本文算法进行计算, 其结果如表 5 所示。

表 1 计算结果 ($n=3, p=0.1$)

x^0	T/s	k	x^*
$(0,0,0)^T$	0.078000	3	$(1.0000,1.0000,1.0000)^T$
$(0.9,0.9,0.9)^T$	0.063000	2	$(1.0000,1.0000,1.0000)^T$
$(3,3,3)^T$	0.063000	2	$(1.0000,1.0000,1.0000)^T$
$(-2,2,-2)^T$	0.110000	4	$(1.0000,1.0000,1.0000)^T$

表 2 计算结果 ($n=3, p=0.01$)

x^0	T/s	k	x^*
$(0,0,0)^T$	0.078000	3	$(1.0000,1.0000,1.0000)^T$
$(0.9,0.9,0.9)^T$	0.062000	2	$(1,1,1)^T$
$(3,3,3)^T$	0.063000	2	$(1.0000,1.0000,1.0000)^T$
$(-2,2,-2)^T$	0.093000	3	$(1,1,1)^T$

表 3 计算结果 ($n=5, p=0.1$)

x^0	T/s	k	x^*
$(0,0,0,0,0)^T$	0.125000	3	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(0.9,0.9,0.9,0.9,0.9)^T$	0.094000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(3,3,3,3,3)^T$	0.078000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(-2,2,-2,2,-2)^T$	0.141000	4	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$

表 4 计算结果 ($n=5, p=0.01$)

x^0	T/s	k	x^*
$(0,0,0,0,0)^T$	0.125000	3	$(1,1,1,1,1)^T$
$(0.9,0.9,0.9,0.9,0.9)^T$	0.079000	2	$(1,1,1,1,1)^T$
$(3,3,3,3,3)^T$	0.078000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(-2,2,-2,2,-2)^T$	0.109000	3	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$

表 5 计算结果

n	T/s			
	MENM	MENM	MGSM	IM
10	0.152508	3	4	7
50	3.112601	3	4	9
100	19.279655	3	4	8

算例 2 矩阵 A 的对角线为 500, 非对角线元素从区间 $[1, 2]$ 中随机选取, 且矩阵 A 为对称矩阵, 令 $b = (A - I)e$, 则 $x = (1, 1, \dots, 1)^T$ 为绝对值方程 $Ax - |x| = b$ 的准确解。

a) 选取不同初始点及 p , 利用本文算法进行计算, 其结果如表 6~9 所示。

b) 选取文献[9, 10, 11]中初始点 $x^0 = (0, 0, \dots, 0)^T$, 令 $p = 0.1$ (或 $p = 0.01$), 利用本文算法进行计算, 其结果如表 10 所示。

表 6 计算结果 ($n=3, p=0.1$)

x^0	T/s	k	x^*
$(0,0,0)^T$	0.093000	3	$(1.0000,1.0000,1.0000)^T$
$(0.9,0.9,0.9)^T$	0.063000	2	$(1.0000,1.0000,1.0000)^T$
$(3,3,3)^T$	0.063000	2	$(1.0000,1.0000,1.0000)^T$
$(-2,2,-2)^T$	0.094000	3	$(1.0000,1.0000,1.0000)^T$

表 7 计算结果 ($n=3, p=0.01$)

x^0	T/s	k	x^*
$(0,0,0)^T$	0.172000	3	$(1,1,1)^T$
$(0.9,0.9,0.9)^T$	0.063000	2	$(1,1,1)^T$
$(3,3,3)^T$	0.078000	2	$(1,1,1)^T$
$(-2,2,-2)^T$	0.094000	3	$(1,1,1)^T$

表 8 计算结果 ($n=5, p=0.1$)

x^0	T/s	k	x^*
$(0,0,0,0,0)^T$	0.140000	3	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(0.9,0.9,0.9,0.9,0.9)^T$	0.094000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(3,3,3,3,3)^T$	0.094000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(-2,2,-2,2,-2)^T$	0.125000	3	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$

表 9 计算结果 ($n=5, p=0.01$)

x^0	T/s	k	x^*
$(0,0,0,0,0)^T$	0.125000	3	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(0.9,0.9,0.9,0.9,0.9)^T$	0.109000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(3,3,3,3,3)^T$	0.094000	2	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$
$(-2,2,-2,2,-2)^T$	0.141000	3	$(1.0000,1.0000,1.0000,1.0000,1.0000)^T$

表 10 计算结果

n	T/s				
	MENM	MENM	MGSM	IM	SNM
10	0.142265	3	3	4	2
50	4.287579	3	3	4	3
100	46.600082	3	4	5	3

通过表 1~10 可以看出, 本文算法对于求解式(1)是十分快速且有效的; 表 5、10 数据结果表明, 本文算法明显优于文献[10]中的迭代法及文献[11]中的改进 Gauss-Seidel 算法; 由表 10 还可以看出, 本文算法与文献[9]中的光滑牛顿法具有几乎一样快的计算速度。

4 结束语

本文构造了极大熵牛顿法求解绝对值方程 $Ax - |x| = b$ 。首先通过等价转换将绝对值方程转换为非线性光滑方程组, 然后用牛顿法对其求解, 并且证明了在适当条件下该算法具有较快的收敛速度。大量的数值实验也表明了该算法不仅计算速度快而且十分准确。是否还有其他更好、收敛速度更快的算法来解决绝对值方程 $Ax - |x| = b$ 及其他相关形式的绝对值方程, 将是笔者后续工作的重点。

参考文献:

- [1] ROHN J. Systems of linear interval equations[J]. *Linear Algebra and Its Application*, 1989, 126(11): 39-78.
- [2] MANGASARIAN O L, MEYER R R. Absolute value equations[J]. *Linear Algebra and Its Application*, 2006, 419(5): 359-367.
- [3] MANGASARIAN O L. A generalized Newton method for absolute value equations[J]. *Optimization Letters*, 2009, 3(1): 101-108.
- [4] CACCETTA L, QU B, ZHOU G L. A globally and quadratically convergent method for absolute value equations[J]. *Computation Optimization Application*, 2011, 48(1): 45-58.
- [5] 王爱祥, 王海军. 绝对值方程的区间算法[J]. *贵州大学学报*, 2010, 27(2): 7-10.
- [6] WANG A X, WANG H J, DENG Y K. Interval algorithm for absolute value equations[J]. *Central European Journal Mathematics*, 2011, 9(5): 1171-1184.
- [7] 雍龙泉, 孙培民, 高凯. 极大熵自适应微粒子群混合算法求解绝对值方程[J]. *计算机应用研究*, 2011, 28(7): 2479-2481.
- [8] 雍龙泉. 基于凝聚函数的和声搜索算法求解绝对值方程[J]. *计算机应用研究*, 2011, 28(8): 2922-2926.
- [9] YONG L Q, LIU S Y, ZHANG S M, et al. Smoothing newton method for absolute value equations based on aggregate function[J]. *International Journal of the Physical Sciences*, 2011, 6(23): 5399-5405.
- [10] NOOR M A, IQBAL J, AL-SAID E. On an iterative method for solving absolute value equations[J]. *Optimization Letters*, DOI:10.1007/s11590-011-0332-0.
- [11] NOOR M A, IQBAL J, KHATTRI S, et al. A new iterative method for solving absolute value equations[J]. *International Journal of the Physical Sciences*, 2011, 6(7): 1793-1797.
- [12] LI X S. An efficient method for non-differentiable optimization problem[J]. *Science in China-Series A*, 1994, 24(4): 371-377.
- [13] 韩继业, 修乃华, 戚厚铎. 非线性互补理论与算法[M]. 上海: 上海科学技术出版社, 2006: 24-30.
- [14] 李庆扬, 莫致中, 祁力群. 非线性方程组的数值解法[M]. 北京: 科学出版社, 1987: 38-50.