

一类不可微方程组的区间算法*

刘风华^{1,2}

(1. 中国矿业大学 信息与电气工程学院, 江苏 徐州 221008; 2. 徐州工程学院 数学与物理科学学院, 江苏 徐州 221008)

摘要: 针对不可微方程组—绝对值方程 $Ax + B|x| = b$ 的数值解问题进行研究, 提出了通过构造极大熵函数和新的区间算子对方程进行求解的区间极大熵算法。该算法能同时求出绝对值方程的近似解和估算其近似解的误差限, 并在 A 的奇异值全部大于 $|B|$ 的奇异值时, 证明了算法的收敛性且收敛速度至少是线性的。理论分析和数值结果均表明提出的算法是有效的。

关键词: 绝对值方程; 极大熵; 区间算子; 区间算法

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 1001-3695(2013)12-3540-03

doi: 10.3969/j.issn.1001-3695.2013.12.005

Interval algorithm for one class of non-differentiable equations

LIU Feng-hua^{1,2}

(1. School of Information & Electrical Engineering, China University of Mining & Technology, Xuzhou Jiangsu 221008, China; 2. School of Mathematics & Physical Science, Xuzhou Institute of Technology, Xuzhou Jiangsu 221008, China)

Abstract: This paper concerned with the non-differentiable equations and the absolute value equations. Based on maximum entropy and a new interval operator, this paper proposed a interval maximum entropy algorithm which could solve the absolute value equations and estimate error between real solution and approximate solution. It proved the convergence and linear convergent rate when the singular values of A exceeded the singular value of $|B|$. Theoretic analysis and numerical results show the method is effective.

Key words: absolute value equations; maximum entropy; interval operator; interval algorithm

0 引言

考虑如下形式的绝对值方程

$$Ax + B|x| = b \quad (1)$$

其中: $A, B \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ 。

绝对值方程(1)最先是由 Rohn^[1] 在研究区间线性方程组时提出, 随后 Mangasarian 对绝对值方程的理论和数值求解进行了深入研究。文献[2,3]给出了绝对值方程(1)解的分类, 讨论了解的存在唯一性。文献[4,5]证明了绝对值方程等价于线性互补问题且其求解是 NP-难题。文献[6]研究了绝对值方程和背包问题的联系。文献[7]将绝对值方程转换为凹函数极小化问题, 提出了连续线性算法(SLA)。文献[8]进一步提出效率更高的半光滑牛顿法, 文献[9]改进了这种方法。文献[10]基于磨光函数提出了一个求解绝对值方程的二次收敛算法。文献[11]基于区间分析理论给出了求解绝对值方程的区间验证迭代方法。

在实际计算中, 如何保证数值解的有效性是一类重要问题。目前研究绝对值方程数值解有效性的文献很少, 本文的主要工作就是基于这一点, 利用区间数学和极大熵的方法建立了新的区间算子, 构造了求解绝对值方程(1)的区间算法。新算法的显著特点是不仅能够求出绝对值方程(1)的近似解, 还

可以同步估算其近似解和准确解的误差。

1 区间数学基本概念

$I(\mathbb{R}^n)$ 表示所有 n 维区间向量的集合, $I(X^{(0)})$ 表示 $X^{(0)}$ 上所有区间向量的集合。

对于 $X = [\underline{X}, \bar{X}] \in I(\mathbb{R})$, 中点、宽度、绝对值分别定义为

$$\text{mid}(X) = [\underline{X}, \bar{X}] / 2$$

$$\text{wid}(X) = \bar{X} - \underline{X}$$

$$|X| = \max(|\underline{X}|, |\bar{X}|)$$

任给区间 $X = [\underline{X}, \bar{X}]$, $Y = [\underline{Y}, \bar{Y}]$, 其四则运算定义为

$$X + Y = [\underline{X} + \underline{Y}, \bar{X} + \bar{Y}]$$

$$X - Y = [\underline{X} - \bar{Y}, \bar{X} - \underline{Y}]$$

$$X \cdot Y = [\min(\underline{X}\underline{Y}, \underline{X}\bar{Y}, \bar{X}\underline{Y}, \bar{X}\bar{Y}), \max(\underline{X}\underline{Y}, \underline{X}\bar{Y}, \bar{X}\underline{Y}, \bar{X}\bar{Y})]$$

$$X/Y = [\underline{X}, \bar{X}] \cdot [1/\bar{Y}, 1/\underline{Y}], 0 \notin Y \quad (2)$$

定义 1 如果函数 $f(x)$ 和区间函数 $F(X)$ 满足 $F(x) = f(x)$, $f(x) \in F(X)$, $\forall x \in X$, 则称 $F(X)$ 为 $f(x)$ 的区间扩张。

有关区间数学的其他概念和内容可参见文献[12,13]。给定矩阵 $A = (a_{ij})_{n \times n}$, 则 $|A|$ 表示 $(|a_{ij}|)_{n \times n}$ 。 I 表示相应的单位矩阵, $Y = (1, 1, \dots, 1)^T$; $\sigma_{\min}(A)$ 、 $\sigma_{\max}(A)$ 、 $\|A\|$ 分别表示 A 的最小、最大的奇异值、 A 的列和范数。

收稿日期: 2013-03-14; **修回日期:** 2013-05-21 **基金项目:** 国家自然科学基金资助项目(31270577); 徐州工程学院校科研基金资助项目(XKY2011101)

作者简介: 刘风华(1976-), 女, 讲师, 博士研究生, 主要研究方向为数值计算、数据挖掘(liufenghua@xzit.edu.cn)。

2 区间算子及其收敛性分析

2.1 极大熵函数

设函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$f(x) = Ax + B|x| - b \quad (3)$$

显然, 式(2)的零点 $\tilde{x}$ 满足绝对值方程(1)。

令 $g(x) = |x|$, 则

$$g(x) = \max\{x, -x\} \quad (4)$$

其中: $\max$ 是按分量取的。构造 $g(x)$ 的极大熵函数 $g_p: \mathbb{R}^n \rightarrow \mathbb{R}^n$:

$$g_p(x) = (g_{p1}(x), g_{p2}(x), \dots, g_{pn}(x)) \quad (5)$$

其中: $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$, $g_{pi}(x) = \frac{\ln(e^{px_i} + e^{-px_i})}{p}$, $p > 0$ 。

定理 1 设 $g(x)$ 、 $g_p(x)$ 分别如式(4)(5)所示, $\forall x \in \mathbb{R}^n$ 成立。

a) $g_p(x) < g_q(x)$, $\forall p > q > 0$ 。

b) $g(x) \leq g_p(x) \leq g(x) + \frac{\ln 2}{p} Y$ 。

证明 a) $x \in \mathbb{R}$, 对 p 求导,

$$g'_{pi} = \frac{e^{px} (px - \ln(e^{px} + e^{-px})) + e^{-px} (-px - \ln(e^{px} + e^{-px}))}{p^2 (e^{px} + e^{-px})} < 0$$

故 $g_{pi}(x) < g_{qi}(x)$, $\forall p > q > 0$ 。得证。

b) $\frac{\ln 2 e^{pg_i(x)}}{p} \leq g_{pi}(x) \leq \frac{\ln 2 e^{pg_i(x)}}{p}$, $g_i(x) \leq g_{pi}(x) \leq g_i(x) + \frac{\ln 2}{p}$

其中: $x \in \mathbb{R}^n$, $g_i(x)$ 是 $g(x)$ 的第 i 个分量, 所以

$$g(x) \leq g_p(x) \leq g(x) + \frac{\ln 2}{p} Y, \text{证毕。}$$

显然, 令 $p \rightarrow +\infty$, $g_p(x) \rightarrow g(x)$ 。

令 $f_p(x) = Ax + Bg_p(x) - b$ (6)

可以通过计算光滑函数 f_p 的零点来构造算法近似求解绝对值方程(1)的解。

2.2 区间扩张

定义区间函数

$$(IE1) G_p(X) = [G_{p1}(X), G_{p2}(X), \dots, G_{pn}(X)]^T$$

$$(IE2) G'_p(X) = \text{diag}(G'_{p1}(X), G'_{p2}(X), \dots, G'_{pn}(X))$$

其中: $G_{pi}(X) = \frac{\ln(1 + e^{2pX_i})}{p} - X_i$, $G'_{pi}(X) = 1 - \frac{2}{1 + e^{2pX_i}}$ ($i = 1, 2, \dots, n$)。

定理 2 区间函数(IE1)、(IE2)分别是 $g_p(x)$ 、 $g'_p(x)$ 的包含单调性的区间扩张。

证明 由式(5)可得,

$$g_{pi}(x) = \frac{\ln(e^{px_i} + e^{-px_i})}{p} = \frac{1}{p} \ln(e^{2px_i} + 1) - x_i \frac{\partial g_{pi}(x)}{\partial x_i} = (0, 0, \dots, \frac{e^{px_i} - e^{-px_i}}{e^{px_i} + e^{-px_i}}, 0)^T = (0, 0, \dots, 1 - \frac{2}{1 + e^{2px_i}}, 0)^T \quad (7)$$

所以 $\forall x \in X \in I(\mathbb{R}^n)$, $g_p(x) = G_p(x)$, 同时 $g_p(x) \in G_p(X)$, 则 $G_p(X)$ 是 $g_p(x)$ 的包含单调性的区间扩张。

同理, $G'_p(X)$ 是雅克比矩阵 $g'_p(x)$ 的包含单调性的区间扩张。证毕。

根据式(6), $F'_p(X) = A + BG'_p(X)$ 是 f'_p 的区间扩张。

下面构造区间算子:

$$K(X) = A^{-1}(b - BG_p(X) - BG'_p(X)(X - \text{mid}(X)))$$

算法 1

a) 给定初始区间 $X^{(0)}$, $k=0$, 精度 ω 。

b) 令 $p=10$, 计算:

$$K(X^{(k)}) = A^{-1}(b - BG_p(X) - BG'_p(X)(X - \text{mid}(X)))$$

以及新的区间 $X^{(k+1)} = X^{(k)} \cap K(X^{(k)})$ 。

c) 若 $\text{wid}(X) \leq \omega$, 则取近似解 $\tilde{x} = \text{mid}(X^{(k)})$; 否则 $k=k+1$, 转步骤 b)。

2.3 算法的收敛性

引理 1^[3] 若 $\sigma_{\max}(|B|) < \sigma_{\min}(A)$, 则对于任意 $b \in \mathbb{R}^n$, 式(1)有唯一解。

引理 2 若 $\sigma_{\max}(|B|) < \sigma_{\min}(A)$, 那么 $\|A^{-1}|B|\| < 1$ 。

证明 设 $A = US_A V$ 是 A 的奇异值分解, 则 $A^{-1} = V'S_A^{-1}U'$ 。

$$\|A\|^2 = \max_{\|x\|=1} \|Ax\|^2 = \max_{\|x\|=1} x^T A^T A x = \max_{\|x\|=1} x^T V^T S_A^2 V x = \max_{\|y\|=1} y^T S_A^2 y = \|S_A\|^2$$

同理 $\| |B| \|^2 = \|S_{|B|}\|^2$, $\|A^{-1}\|^2 = \|S_A^{-1}\|^2$ 。其中 $S_{|B|}$ 满足 $|B|$ 的奇异值分解 $|B| = U_{|B|} S_{|B|} V_{|B|}$ 。

$$\|A^{-1}|B|\|^2 \leq \|A^{-1}\|^2 \| |B| \|^2 =$$

$$\|S_A^{-1}\|^2 \|S_{|B|}\|^2 = \left(\frac{\sigma_{\max}(|B|)}{\sigma_{\min}(A)}\right)^2 < 1$$

证毕。

引理 3 $f_p(\tilde{x}) = 0$ 且 $\tilde{x} \in X$ 则 $\tilde{x} \in K(X)$ 。

证明 令 $\varphi_p(x) = x - A^{-1}f_p(x)$, 则 $\varphi(\tilde{x}) = \tilde{x}$ 。

$\varphi'_p(x) = I - A^{-1}g'_p(x)$, 故 $\varphi'_p(x)$ 的包含单调性的区间扩张为

$$\Phi'_p(X) = I - A^{-1}(A + BG'_p(X)) = -A^{-1}BG'_p(X)$$

当 $\tilde{x} \in X$ 时, $\varphi_p(\tilde{x}) = \varphi_p(\text{mid}(X)) + \varphi'_p(\text{mid}(X) + \theta(\tilde{x} - \text{mid}(X)))(X - \text{mid}(X))$ 。其中 $0 < \theta < 1$ 。

显然

$$\varphi'_p(\text{mid}(X) + \theta(\tilde{x} - \text{mid}(X))) \in \Phi'_p(X)$$

$$\varphi_p(\tilde{x}) \in \varphi_p(\text{mid}(X)) + \Phi'_p(X)(X - \text{mid}(X))$$

即 $\varphi_p(\tilde{x}) \in A^{-1}(-Bg_p(\text{mid}(X)) + b) - A^{-1}BG'_p(X)(X - \text{mid}(X))$ 。

从而 $\tilde{x} = \varphi(\tilde{x}) \in K(X)$ 。证毕。

由引理 3 容易看出, 若 $X \cap K(X) = \emptyset$, 则 X 中无方程 $f_p(\tilde{x}) = 0$ 的解。

定理 2 设 f_p 在 $X^{(0)}$ 上有唯一解 $\tilde{x}$ 且 $\sigma_{\max}(|B|) < \sigma_{\min}(A)$, 则算法 1 产生的区间向量序列 $\{X^{(k)}\}_{k \in \mathbb{N}}$ 满足

a) $\lim_{k \rightarrow +\infty} \text{wid}(X^{(k)}) = 0$;

b) $\lim_{k \rightarrow +\infty} \text{mid}(X^{(k)}) = \tilde{x}$, 且存在 $0 < q < 1$, $\|\tilde{x} - \text{mid}(X^{(k)})\| \leq \frac{1}{2} \frac{q^k}{1-q} \text{wid}(X^{(0)})$ 。

证明 a) 由引理 3 知, $\tilde{x} \in X^{(0)} \cap K(X^{(0)})$, 即 $\tilde{x} \in X^{(1)}$ 。假设 $\tilde{x} \in X^{(k)}$, 则 $\tilde{x} \in K(X^{(k)})$, 故 $\tilde{x} \in X^{(k+1)}$ 。由归纳法, 对一切的正整数 k , 成立 $\tilde{x} \in X^{(k)}$, 且 $X^{(k)} \subseteq X^{(k-1)} \subseteq \dots \subseteq X^{(0)}$ 。于是算法能无限地继续下去, 得到区间向量序列 $\{X^{(k)}\}_{k \in \mathbb{N}}$, 且 $\text{wid}(X^{(k+1)}) \leq \text{wid}(X^{(k)})$ 。

由引理 2 得

$$\|A^{-1}B\| \leq \|A^{-1}\| \|B\| = \|A^{-1}\| \| |B| \| < 1$$

$$\|G'_p(X^{(0)})\| = \|\text{diag}(1 - \frac{2}{e_i^{2pX_i(0)} + 1})\| < 1$$

再令 $q = \|A^{-1}BG'_p(X^{(0)})\| \leq \|A^{-1}B\| \|G'_p(X^{(0)})\| < 1$, 所以

$$\begin{aligned} \text{wid}(K(X^{(k)})) &\leq \text{wid}(A^{-1}BG'_p(X^{(k)})(X^{(k)} - \text{mid}(X^{(k)}))) \leq \\ &\|A^{-1}BG'_p(X^{(k)})\| \text{wid}(X^{(k)}) \leq \|A^{-1}BG'_p(X^{(0)})\| \text{wid}(X^{(k)}) = \\ &q \cdot \text{wid}(X^{(k)}) \end{aligned}$$

于是

$$\text{wid}(X^{(k+1)}) \leq q \cdot \text{wid}(X^{(k)}) \leq q^{(k+1)} \cdot \text{wid}(X^{(0)}) \quad (8)$$

从而 $\lim_{k \rightarrow +\infty} \text{wid}(X^{(k)}) = 0$ 。

b) 根据算法 1 得 $X^{(k+1)} \subseteq X^{(k)}$, 则

$$\|\text{mid}(X^{(k+1)}) - \text{mid}(X^{(k)})\| \leq \frac{1}{2} \text{wid}(X^{(k)})$$

于是任取 $p > k$, p 是正整数, 有

$$\|\text{mid}(X^{(k+1)}) - \text{mid}(X^{(k)})\| \leq$$

$$\frac{1}{2}(\text{wid}(X^{(k+p+1)}) + \dots + \text{wid}(X^{(k)})) \leq$$

$$\frac{1}{2}(q^{k+p+1} + \dots + q^k) \cdot \text{wid}(X^{(0)}) =$$

$$\frac{q^k}{2(1-q)} \text{wid}(X^{(0)}) \quad (9)$$

于是 $\lim_{k \rightarrow +\infty} \|\text{mid}(X^{(k+1)}) - \text{mid}(X^{(k)})\| = 0$ 。故 $\{X^{(k)}\}_{k \in \mathbb{N}}$

是 Cauchy 列, 必存在一点 $\bar{x} \in X^{(k)}$, 满足 $\lim_{k \rightarrow +\infty} \text{mid}(X^{(k)}) = \bar{x}$ 。

下面证 $\bar{x} = \tilde{x}$:

$$K(X^{(k)}) - \text{mid}(X^{(k)}) =$$

$$-A^{-1}f_p(\text{mid}(X^{(k)})) + A^{-1}BG'_p(X)(X - \text{mid}(X^{(k)}))$$

所以 $-A^{-1}f_p(\text{mid}(X^{(k)})) \in K(X^{(k)}) - \text{mid}(X^{(k)})$

$$\| -A^{-1}f_p(\text{mid}(X^{(k)})) \| \leq \| K(X^{(k)}) - \text{mid}(X^{(k)}) \| \leq$$

$$\| K(X^{(k)}) - \text{mid}(X^{(k+1)}) \| + \| \text{mid}(X^{(k+1)}) - \text{mid}(X^{(k)}) \| \leq$$

$$\text{wid}(K(X^{(k)})) + \| \text{mid}(X^{(k+1)}) - \text{mid}(X^{(k)}) \| \leq$$

$$q^k \cdot \text{wid}(X^{(0)}) + \| \text{mid}(X^{(k+1)}) - \text{mid}(X^{(k)}) \|$$

故 $\lim_{k \rightarrow +\infty} \|A^{-1}f_p(\tilde{x})\| = 0$ 。于是 $f_p(\tilde{x}) = 0$, 又因为 f_p 在 X 上只

有唯一零点, 则 $\tilde{x} = \bar{x}$ 。进一步, 由式(9)可得

$$\|\tilde{x} - \text{mid}(X^{(k)})\| \leq \frac{1}{2} \frac{q^k}{1-q} \text{wid}(X^{(0)})$$

证毕。

定理 2 不仅证明了算法的收敛性, 还给出了每步迭代中近似解和准确解的误差关系。此外, 式(8)还说明了该区间算法收敛速度至少是线性的。

3 数值测试

为了测试算法的有效性, 本文进行如下的测试。所用的程序是借助区间程序包 Intlab 5.4, 用 MATLAB 编程实现的。

例 1

$$A = \begin{pmatrix} -7.22218236086100 & 2.07584958387639 & -9.69452145941927 \\ -5.94469562879454 & -4.55624150060079 & 4.93571353128859 \\ -6.02556514677021 & -6.02371464477876 & -1.09807135424106 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.86362915692333 & 0.69244283564865 & 0.34427493694858 \\ -0.06801131664915 & 0.05030499261034 & 0.67623689010477 \\ -0.16270106454499 & -0.59470528469923 & -0.96072097227037 \end{pmatrix}$$

$$b = (0.00578336237505 \quad 0.41884035037612 \quad -0.14214940936072)^T$$

经过计算, 每步迭代结束的解区间如表 1 所示。

表 1 迭代过程

k	第 k 次迭代结束的解区间
1	$[-0.47048836862726, 0.37690040891350]$
	$[-0.62108963106578, 0.70176725219446]$
	$[-0.28275544091662, 0.39573277777363]$
4	$[-0.0898985217182, -0.01502224562327]$
	$[-0.00999727889468, 0.10889893241324]$
	$[0.03044808330342, 0.08835233772381]$
8	$[-0.05393480818404, -0.05101901070304]$
	$[0.04716762290398, 0.05179814932206]$
	$[0.05828445969066, 0.06053957481925]$
12	$[-0.05253356322159, -0.05241990728942]$
	$[0.04939235664793, 0.04957285189914]$
	$[0.05936794798892, 0.05945584928689]$
13	$[-0.05250197424116, -0.05245147101261]$
	$[0.04944248209621, 0.04952268551798]$
	$[0.05939236250596, 0.05943142159383]$

得到近似解为

$$\tilde{x} = (-0.05247672262688 \quad 0.04948258380710 \quad 0.05941189204989)^T$$

检验可得 $\|A\tilde{x} + B|\tilde{x}| - b\| < 10^{-7}$ 。

例 2 随机选取阶数 $n = 10, 50, 100, 500$ 的矩阵 A, B , 进行多次实验, 得到表 2 所示结果。

表 2 算法效率测试

n	k	time/s	终止准则
10	13	10.016 000	$\ A\tilde{x} + B \tilde{x} - b\ \leq 10^{-6}$
50	25	21.029 000	$\ A\tilde{x} + B \tilde{x} - b\ \leq 10^{-6}$
100	76	47.404 000	$\ A\tilde{x} + B \tilde{x} - b\ \leq 10^{-6}$
500	106	155.252 000	$\ A\tilde{x} + B \tilde{x} - b\ \leq 10^{-6}$

其中, k 表示迭代次数, time 表示 CPU 运行时间 (以 s 为单位)。

数值结果表明算法是有效的。但是, 对于求解给定的绝对值方程, 初始区间的确定仍然是一个困难, 这是笔者下一步将做的工作。

参考文献:

- [1] ROHN J. Systems of interval linear equations[J]. *Linear Algebra Applications*, 1989, 126(12): 39-78.
- [2] ROHN J. A theorem of the alternatives for the equation $Ax + B|x| = b$ [J]. *Linear and Multilinear Algebra*, 2004, 52(6): 421-426.
- [3] ROHN J. On unique solvability of the absolute value equation [J]. *Optimization Letters*, 2009, 3(4): 603-606.
- [4] MANGASARIAN O L, MEYER R R. Absolute value equations [J]. *Linear Algebra and Applications*, 2006, 419(2-3): 359-367.
- [5] MANGASARIAN O L. Absolute value programming [J]. *Computational Optimization and Applications*, 2007, 36(1): 43-53.
- [6] MANGASARIAN O L. Knapsack feasibility as an absolute value equation solvable by successive linear programming [J]. *Optimization Letters*, 2009, 3(2): 161-170.
- [7] NEMIROVSKII A. Several NP-hard problems arising in robust stability analysis [J]. *Mathematics of Control, Signals, and System*, 1999, 6(2): 99-105.
- [8] MANGASARIAN O L. A generalized Newton method for absolute value equations [J]. *Optimization Letters*, 2009, 3(1): 101-108.
- [9] ZHANG C, WEI Q J. Global and finite convergence of a generalized Newton method for absolute value equations [J]. *Optimization Theory and Applications*, 2009, 143(2): 391-403.
- [10] CACCETTA L, QU Biao, ZHOU Guang-lu. A globally quadratically convergent method for absolute value equations [J]. *Computational Optimization and Applications*, 2011, 48(1): 45-58.
- [11] WANG Hai-jun, LIU Hao, CAO Su-yu. A verification method for enclosing solutions of absolute value equations [J]. *Collectanea Mathematica*, 2013, 64(1): 17-38.