

参赛学生姓名：王洛彬

中学：南京市金陵中学

省份：江苏省

国家/地区：中国

指导老师姓名：严涛

指导老师单位：南京理工大学

论文题目：求解绝对值方程组的快速算法

求解绝对值方程组的快速算法

南京市金陵中学 王洛彬

摘要:

在本文中, 我们提出了一种求解绝对值方程组的快速数值方法. 该方法基于文献[1]中的Douglas-Rachford分裂法, 增加了一个外推步骤, 该步骤仅需进行一些简单的计算. 我们提供了一些数值例子以验证该算法的有效性.

关键词:

绝对值方程组, Douglas-Rachford分裂法, 线性组合, 快速算法.

1 引言

在本文中, 我们考虑的方程组形式为

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

其中 A 是一个 $n \times n$ 的矩阵, b 是 n 维的向量, $|x| = (|x_1|, |x_2|, \dots, |x_n|)$. 2006 年 Mangasarian and Meyer [5] 称问题(1) 为绝对值方程组问题.

绝对值方程是一类具有广泛应用和深厚理论基础的数学问题. 这类方程的特殊性质和求解方法使其在多个领域中具有重要意义. 例如, 在信号处理领域, 它们被用于噪声降低和信号恢复. 在优化问题中, 绝对值方程组的解解决了带有绝对值约束的问题, 例如最小化绝对值目标函数. 在控制理论中, 绝对值方程组有助于分析和设计系统稳定性[2, 4, 6]. 有关绝对值方程组的详细信息, 请参见文献[3].

绝对值方程组的核心在于方程内的绝对值函数. 绝对值函数 $|x|$ 为求解方程组带来了困难. 对于单变量绝对值方程组, 例如 $|x - a| = b$, 解通常可以通过代数方法直

接获得. 然而, 随着变量数量的增加, 问题变得更加复杂. 例如, 方程组

$$\begin{pmatrix} 0 & 0 \\ -1 & -0.5 \end{pmatrix} x - x = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

有唯一的解 $(1, 0)$. 而绝对值方程组

$$\begin{pmatrix} 0 & 0 \\ -1 & -0.5 \end{pmatrix} x - |x| = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

却有三个解: $(-1, -4)$, $(-1, 4/3)$ 和 $(1, 0)$.

令

$$Q(x) = Ax + x - b, \quad F(x) = Ax - x - b.$$

Mangasarian and Meyer 在文献[5] 中证明了绝对值方程组(1) 和如下的问题等价: 寻找 x , 使得

$$Q(x) \geq 0, \quad F(x) \geq 0, \quad \langle F(x), Q(x) \rangle = 0. \quad (2)$$

其中 $\langle \cdot \rangle$ 表示两个向量的内积, 即将两个向量对应的分量做乘积并将所有的乘积作和. (2)被称为广义线性互补问题(GLCP). 基于(2) 和(1) 的等价性, 文献[1] 的作者提出了求解绝对值方程组(1) 的Douglas-Rachford 分裂方法.

令 $e(x) = Ax - |x| - b$, 显然地, $e(x) = 0$ 当且仅当 x 是绝对值方程组(1) 的解. 因此我们可以应用 $\|e(x)\|$ 来衡量 x 和绝对值方程组的解之间的距离. 不难看出 $e(x) = 0$ 等价于

$$Q(x) + F(x) = Q(x) + F(x) - \gamma e(x), \quad (3)$$

其中 $\gamma > 0$. 基于(3)求解(1) 的不动点迭代方法是:

$$Q(x_{k+1}) + F(x_{k+1}) = Q(x_k) + F(x_k) - \gamma e(x_k).$$

用 $Ax + x - b$ 和 $Ax - x - b$ 分别替换 $Q(x), F(x)$, 上述方法可以写成如下形式[1]. 给定 x_k, x_{k+1} 可通过下列方式产生:

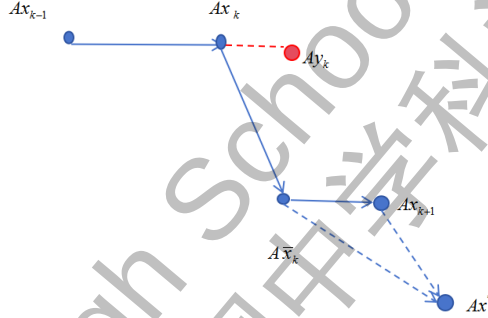
$$2Ax_{k+1} = 2Ax_k - \gamma(Ax_k - |x_k| - b), \quad (4)$$

其中 γ 是一个正的常数.

令 x^* 是绝对值方程组(1) 的解. 根据文献[1] 的分析, 我们可知

$$\|Ax_k - Ax^*\|^2 \leq \|Ax_{k-1} - Ax^*\|^2 - \frac{\gamma(2-\gamma)}{4} \|e(x_{k-1})\|^2 \quad (5)$$

也就是说, Ax_k 比 Ax_{k-1} 更靠近 Ax^* . 通常情况下为了保证 $\{Ax_k\}$ 收敛到 Ax^* , γ 的选取比较保守, 我们猜测存在一个 β_k 使得 $Ay_k = Ax_k + \beta_k(Ax_k - Ax_{k-1})$ 比 Ax_k 更靠近 Ax^* [如图所示]. 这意味着从 Ax_k 出发, 沿着方向 $Ax_k - Ax_{k-1}$ 移动, $\|Ax - Ax^*\|^2$ 的函数值可能会下降. 如果真的下降的话, 我们称方向 $Ax_k - Ax_{k-1}$ 为 $\|Ax - Ax^*\|^2$ 在 Ax_k 处的下降方向. 令 $A\bar{x}_k = Ax_k - \frac{\gamma}{2}(Ax_k - |x_k| - b)$. 根据(5), $A\bar{x}_k$ 比 Ax_k 更靠近 Ax^* . 因此 $-(Ax_k - |x_k| - b)$ 也是 $\|Ax - Ax^*\|^2$ 在 Ax_k 处的下降方向.



2 问题提出

如果 $-(Ax_k - |x_k| - b)$ 和 $Ax_k - Ax_{k-1}$ 都是函数 $\|Ax - Ax^*\|^2$ 在 Ax_k 处的下降方向, 由这两个方向的组合构成的新方向是否也是 $\|Ax - Ax^*\|^2$ 在 Ax_k 处的下降方向呢? 如果是的话, 如何选择合适的组合系数? 我们能否仅仅通过简单的计算就能获得一个比 $A\bar{x}_k$ 更好的点(我们称更靠近 Ax^* 的点为更好的点)? 基于这个观察、以文献[1] 的方法为基础, 我们对此问题进行了探究.

3 探究过程

首先我们给出本文用到的记号. 本文用 R^n 表示所有 n 维向量构成的集合, $|\cdot|$ 表示实数的绝对值, $\|x\|$ 定义为 $\sqrt{\langle x, x \rangle}$, $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ 对任意的 $x, y \in R^n$. 绝对值方程组(1) 的解集记为 S .

3.1 快速算法

本节首先给出我们的新方法.

-
- 1: 任意给出 $x_0 = x_1 \in R^n$, $k = 1, 2, \dots$, $\gamma \in (0, 2)$,
 - 2: 计算

$$2A\bar{x}_k = 2Ax_k - \gamma(Ax_k - |x_k| - b), \quad (6)$$

和

$$Ax_{k+1} = A\bar{x}_k + \beta_k(Ax_k - Ax_{k-1}). \quad (7)$$

- 3: $k = k + 1$, 返回第2步.
-

3.2 理论分析

接下来我们分析这个方法的收敛性质. 整篇文章我们假设矩阵 A 是可逆的. 首先我们讨论 $k = 1$ 和 $k = 2$ 的情况.

$$\begin{aligned} 2A\bar{x}_1 &= 2Ax_1 - \gamma(Ax_1 - |x_1| - b), \\ Ax_2 &= A\bar{x}_1 + \beta_1(Ax_1 - Ax_0) = A\bar{x}_1 \\ 2A\bar{x}_2 &= 2Ax_2 - \gamma(Ax_2 - |x_2| - b), \\ Ax_3 &= A\bar{x}_2 + \beta_2(Ax_2 - Ax_1) \end{aligned} \quad (8)$$

引理3.1 令 $x_2, \bar{x}_2$ 由(8) 计算产生, x^* 是绝对值方程组(1) 的解, 则

$$2\langle Ax_2 - Ax_1, Ax_2 - Ax^* \rangle \leq -\frac{\gamma(1-\gamma)}{2} \|e(x_1)\|^2. \quad (9)$$

证明: 容易看出

$$\begin{aligned} \|Ax_2 - Ax^*\|^2 &= \|Ax_2 - Ax_1 + Ax_1 - Ax^*\|^2 \\ &= \|Ax_2 - Ax_1\|^2 + \|Ax_1 - Ax^*\|^2 + 2\langle Ax_2 - Ax_1, Ax_1 - Ax^* \rangle \end{aligned}$$

以及

$$\|Ax_2 - Ax^*\|^2 \leq \|Ax_1 - Ax^*\|^2 - \frac{\gamma(2-\gamma)}{4} \|e(x_1)\|^2,$$

我们可以得到

$$2\langle Ax_2 - Ax_1, Ax_1 - Ax^* \rangle \leq -\frac{\gamma(2-\gamma)}{4}\|e(x_1)\|^2 - \|Ax_2 - Ax_1\|^2.$$

因此

$$2\langle Ax_2 - Ax_1, Ax_1 - Ax_2 + Ax_2 - Ax^* \rangle \leq -\frac{\gamma(2-\gamma)}{4}\|e(x_1)\|^2 - \|Ax_2 - Ax_1\|^2.$$

所以

$$\begin{aligned} 2\langle Ax_2 - Ax_1, Ax_2 - Ax^* \rangle &\leq -\frac{\gamma(2-\gamma)}{4}\|e(x_1)\|^2 + \|Ax_2 - Ax_1\|^2 \\ &= -\frac{\gamma(2-\gamma)}{4}\|e(x_1)\|^2 + \frac{\gamma^2}{4}\|e(x_1)\|^2 \\ &= -\frac{\gamma(1-\gamma)}{2}\|e(x_1)\|^2. \end{aligned}$$

引理3.2 令 $x_2, \bar{x}_2$ 由(8) 计算产生, x^* 是绝对值方程组(1) 的解, 则

$$2\langle A\bar{x}_2 - Ax_2, Ax_2 - Ax^* \rangle \leq -\frac{\gamma}{2}\|e(x_2)\|^2 \quad (10)$$

证明: 由于

$$\begin{aligned} \|A\bar{x}_2 - Ax^*\|^2 &= \|A\bar{x}_2 - Ax_2 + Ax_2 - Ax^*\|^2 \\ &= \|A\bar{x}_2 - Ax_2\|^2 + \|Ax_2 - Ax^*\|^2 + 2\langle A\bar{x}_2 - Ax_2, Ax_2 - Ax^* \rangle \\ &\leq \|Ax_2 - Ax^*\|^2 - \frac{\gamma(2-\gamma)}{4}\|e(x_2)\|^2, \end{aligned}$$

我们可以得到

$$\begin{aligned} 2\langle A\bar{x}_2 - Ax_2, Ax_2 - Ax^* \rangle &\leq -\frac{\gamma(2-\gamma)}{4}\|e(x_2)\|^2 - \|A\bar{x}_2 - Ax_2\|^2 \\ &= -\frac{\gamma}{2}\|e(x_2)\|^2 \end{aligned}$$

基于以上引理, 我们可以得到下面的结论:

定理 令 $x_2, \bar{x}_2, x_3$ 通过(8) 计算产生, x^* 是绝对值方程组(1) 的解. 取 $\beta_2 = \frac{(1-\gamma)^2}{\gamma(1+\gamma)}$,

则

$$\|Ax_3 - Ax^*\|^2 \leq \|Ax_2 - Ax^*\|^2 - \frac{(1-\gamma)^3}{4} \|e(x_1)\|^2 - \frac{\gamma - \gamma^2}{8} \|e(x_2)\|^2.$$

证明：由于

$$\begin{aligned} \|Ax_3 - Ax^*\|^2 &= \|A\bar{x}_2 + \beta_2(Ax_2 - Ax_1) - Ax^*\|^2 \\ &= \|Ax_2 - Ax^* + \beta_2(Ax_2 - Ax_1) - \frac{\gamma}{2}e(x_2)\|^2 \\ &= \|Ax_2 - Ax^* + \beta_2(Ax_2 - Ax_1)\|^2 + \frac{\gamma^2}{4} \|e(x_2)\|^2 \\ &\quad - \gamma \langle e(x_2), Ax_2 - Ax^* + \beta_2(Ax_2 - Ax_1) \rangle \\ &= \|Ax_2 - Ax^* + \beta_2(Ax_2 - Ax_1)\|^2 + \frac{\gamma^2}{4} \|e(x_2)\|^2 \\ &\quad - \gamma \langle e(x_2), Ax_2 - Ax^* \rangle - \gamma \beta_2 \langle e(x_2), (Ax_2 - Ax_1) \rangle \\ &\leq \|Ax_2 - Ax^* + \beta_2(Ax_2 - Ax_1)\|^2 + \frac{\gamma^2 - \gamma}{4} \|e(x_2)\|^2 \\ &\quad - \gamma \beta_2 \langle e(x_2), (Ax_2 - Ax_1) \rangle, \end{aligned} \quad (11)$$

上述最后一个不等式成立的原因是因为(10). 另一方面,

$$\begin{aligned} &\|Ax_2 - Ax^* + \beta_2(Ax_2 - Ax_1)\|^2 \\ &= \|Ax_2 - Ax^*\|^2 + \beta_2^2 \|Ax_2 - Ax_1\|^2 + 2\beta_2 \langle Ax_2 - Ax^*, Ax_2 - Ax_1 \rangle \\ &\leq \|Ax_2 - Ax^*\|^2 + \beta_2^2 \|Ax_2 - Ax_1\|^2 - \frac{\beta_2 \gamma (1-\gamma)}{2} \|e(x_1)\|^2. \end{aligned} \quad (12)$$

用(12) 替代(11), 我们可得:

$$\begin{aligned} \|Ax_3 - Ax^*\|^2 &\leq \|Ax_2 - Ax^*\|^2 + \beta_2^2 \|Ax_2 - Ax_1\|^2 - \frac{\beta_2 \gamma (1-\gamma)}{2} \|e(x_1)\|^2 \\ &\quad + \frac{\gamma^2 - \gamma}{4} \|e(x_2)\|^2 - \gamma \beta_2 \langle e(x_2), (Ax_2 - Ax_1) \rangle \\ &= \|Ax_2 - Ax^*\|^2 + \frac{\beta_2^2 \gamma^2}{4} \|e(x_1)\|^2 - \frac{\beta_2 \gamma (1-\gamma)}{2} \|e(x_1)\|^2 \\ &\quad + \frac{\gamma^2 - \gamma}{4} \|e(x_2)\|^2 - \gamma \beta_2 \langle e(x_2), (Ax_2 - Ax_1) \rangle \\ &= \|Ax_2 - Ax^*\|^2 - \frac{\beta_2 \gamma [2 - (2 + \beta_2) \gamma]}{4} \|e(x_1)\|^2 \end{aligned}$$

$$\begin{aligned}
& -\frac{\gamma-\gamma^2}{4}\|e(x_2)\|^2 + \frac{\beta_2\gamma^2}{2}\langle e(x_2), e(x_1) \rangle \\
& \leq \|Ax_2 - Ax^*\|^2 - \frac{\beta_2\gamma[2-(2+\beta_2)\gamma]}{4}\|e(x_1)\|^2 \\
& \quad -\frac{\gamma-\gamma^2}{4}\|e(x_2)\|^2 + \frac{\beta_2\gamma^2}{4}(\alpha\|e(x_2)\|^2 + \frac{1}{\alpha}\|e(x_1)\|^2) \\
& = \|Ax_2 - Ax^*\|^2 - \left(\frac{\beta_2\gamma[2-(2+\beta_2)\gamma]}{4} - \frac{\beta_2\gamma^2}{4\alpha}\right)\|e(x_1)\|^2 \\
& \quad - \left(\frac{\gamma-\gamma^2}{4} - \frac{\alpha\beta_2\gamma^2}{4}\right)\|e(x_2)\|^2
\end{aligned}$$

上述最后一个不等式成立源于Young's 不等式 $2\langle x, y \rangle \leq \alpha\|x\|^2 + \frac{1}{\alpha}\|y\|^2$ 对任意的 $\alpha > 0$ 都成立. 如果

$$\frac{\beta_2\gamma[2-(2+\beta_2)\gamma]}{4} - \frac{\beta_2\gamma^2}{4\alpha} > 0$$

和

$$\frac{\gamma-\gamma^2}{4} - \frac{\alpha\beta_2\gamma^2}{4} > 0$$

成立, 我们可以得到

$$\|Ax_3 - Ax^*\|^2 \leq \|Ax_2 - Ax^*\|^2.$$

令

$$\alpha = \frac{1+\gamma}{2(1-\gamma)}, \quad \beta_2 = \frac{(1-\gamma)^2}{\gamma(1+\gamma)}.$$

通过简单的计算可得

$$\frac{\beta_2\gamma[2-(2+\beta_2)\gamma]}{4} - \frac{\beta_2\gamma^2}{4\alpha} = \frac{(1-\gamma)^3}{4} \quad \frac{\gamma-\gamma^2}{4} - \frac{\alpha\beta_2\gamma^2}{4} = \frac{\gamma-\gamma^2}{8}.$$

所以

$$\|Ax_3 - Ax^*\|^2 \leq \|Ax_2 - Ax^*\|^2 - \frac{(1-\gamma)^3}{4}\|e(x_1)\|^2 - \frac{\gamma-\gamma^2}{8}\|e(x_2)\|^2.$$

证毕. 根据上述定理, 当 $\gamma \in (0, 1)$ 时, $\|Ax_3 - Ax^*\|^2 \leq \|Ax_2 - Ax^*\|^2$ 成立.

用同样的方法, 我们可以证明

$$\|Ax_{k+1} - Ax^*\|^2 \leq \|Ax_k - Ax^*\|^2 - \eta\|e(x_k)\|^2 - \rho\|e(x_{k-1})\|^2, \quad \forall k = 1, 2, \dots, \quad (13)$$

其中 $\eta, \rho > 0$. 将(13) 中的不等式从 $k = 1$ 加到 $k = N$, 我们可得

$$\|Ax_{N+1} - Ax^*\|^2 \leq \|Ax_1 - Ax^*\|^2 - \eta\|e(x_N)\|^2 - (\eta + \rho) \sum_{i=1}^{N-1} \|e(x_i)\|^2 - \rho\|e(x_1)\|^2.$$

因此 $\lim_{k \rightarrow \infty} \|e(x_k)\| = 0$. 由于 A 是可逆的, 所以有

$$\lim_{k \rightarrow \infty} x_k = x^*$$

成立.

4 数值结果

本节我们将通过几个实例来验证我们新方法的有效性.

实例 考虑绝对值方程组, 其中 $A = \text{tridiag}(-1, 8, -1) \in R^{n \times n}$, $b = Ax^* - |x^*|$, 最优解为 $x^* = (-1, 1, -1, 1, \dots, -1, 1) \in R^n$ [1]. “n”表示方程组未知数的个数. 我们对比了加速前的方法(记为“DR”)和加速后的方法(记为“ADR”), 记录了这两种方法求解同一个问题所用的迭代次数(Iter.)和计算机的计算时间(Time), “residual”表示停机时得到的 x 对应的 $\|Ax - |x| - b\|$ 的值. 这里我们取 $\gamma = 0.99$.

n	ADR			DR		
	Iter.	Time	residual	Iter.	Time	residual
100	33	0.24761	6.6702e-09	47	0.40978	6.6609e-09
200	34	0.52491	4.657e-09	47	0.7258	9.8529e-09
300	34	0.92201	5.2544e-09	48	1.2935	6.7719e-09
500	34	1.7018	7.2121e-09	48	2.2034	7.4042e-09
1000	35	4.3369	4.8706e-09	49	5.9619	6.1535e-09
2000	35	13.7619	6.5807e-09	49	18.8238	9.1118e-09
3000	35	43.9206	7.6964e-09	50	47.6439	6.5539e-09

和文献[1] 中的算法相比, 我们的新方法只增加了简单的计算量, 表格中的数据表明新方法在计算时间和迭代次数方面有一定的优势, 体现了我们新方法确实提升了计算效率.

5 结论

尽管目前在绝对值方程组的研究中取得了一些进展，但仍有许多领域值得探索。未来的研究可以集中在以下几个方面：

高效算法的开发：需要更高效的算法来解决大规模和高维度的绝对值方程组，以提高计算性能和解的准确性。

应用扩展：随着技术的进步，不断出现新的应用场景。

绝对值方程组是一个数学上深刻且广泛适用的问题。它的研究涵盖了理论分析、算法开发和实际应用。随着技术和理论不断发展，绝对值方程的研究将在各个领域继续发挥重要作用，并为应对更复杂的问题提供新的见解。

References

- [1] C. R. Chen, D. M. Yu, D. R. Han, Exact and inexact Douglas-Rachford splitting methods for solving large-scale sparse absolute value equations, IMA Journal of Numerical Analysis 44(2), 1036-1060, (2023)
- [2] A. Fakharzadeh, N. N. Shams, An efficient algorithm for solving absolute value equations, Journal of Mathematical Extension 15(3), 1-23 (2021)
- [3] Milan Hladík, Hossein Moosaei, Fakhroddin Hashemi, Saeed Ketabchi, Panos M. Pardalos, An overview of absolute value equations: From theory to solution methods and challenges, <http://arxiv.org/abs/2404.06319v1>
- [4] J. He, Y. Liu, J. Tian, Two numerical iteration methods for solving absolute value equations, Science Asia 44, 40-45 (2018)
- [5] O. L. Mangasarian, R. R. Meyer, Absolute value equations, Linear Algebra and Its Applications 419(2-3), 359-367 (2006)
- [6] J. Rohn, An algorithm for solving the absolute value equation. The Electronic Journal of Linear Algebra 18, 589-599 (2009). DOI 10.13001/1081-3810.1332

致谢

本文主要研究绝对值方程的数值解法。绝对值方程组在优化理论与工程计算中有广泛应用，在指导老师的引导下，我将 Douglas-Rachford 分裂方法作为研究对象，并尝试在算法设计中提出改进。

指导老师在整个研究过程中起到关键作用。我对绝对值方程组的求解感兴趣，他向我推荐了参考文献，在他的帮助下，我明确了研究方向。在理论推导时多次和老师讨论，完成了定理的证明。在数值实验环节，老师提供了部分基础代码框架，并在我遇到运行错误或收敛停滞问题时给出了调试建议。我则在其基础上进行参数修改、数据扩展和实验结果整理。需要说明的是，指导老师的帮助是无偿的，其主要作用是引导与协助，而论文主体内容仍由我独立完成。研究中未引入其他人员参与。

1. 选题来源

由于我对绝对值方程组感兴趣，向老师请教如何求解，老师推荐了相关文献，在阅读了 Mangasarian 与 Meyer 关于绝对值方程组的经典工作后，我意识到 Douglas-Rachford 方法在理论上虽有优势，但仍存在改进空间，因此确立了研究主题。

2. 数据获取与实验设计

研究数据主要为随机生成的测试算例。我与老师共同设计了实验流程：设置不同维度的稀疏矩阵，比较改进前后的收敛速度，并通过改变参数 γ 验证算法稳定性。

3. 分工与贡献

我：负责文献调研、公式推导、实验数据生成、参数调整及论文写作。特别是在论文撰写阶段，我完成了研究背景、理论部分、实验结果与结论。

指导老师：负责研究方向把关，提供文献推荐；理论部分，协助完成定理证明；在实验部分，帮助搭建了基础程序框架并指导调试；在论文写作中，对我的初稿提出修改意见，确保论文逻辑清晰。

4. 遇到的困难与解决

研究过程中我主要遇到三类困难：

理论推导复杂：早期无法证明改进算法的收敛性，在老师建议下我转而采用能量函数法，最终完成了证明。

实验效率低：在高维问题上运行时间过长，老师指导我使用稀疏矩阵存储与向量化方法，我据此修改程序后效率显著提升。

写作逻辑：论文初稿中结构不够紧凑，在老师建议下我调整了章节顺序，使论证更加连贯。