

求解带扰动的线性方程组的贪婪随机 Kaczmarz 方法

巫文婷

(北京理工大学 数学与统计学院, 北京 100081)

摘要: 当相容的线性代数方程组的右端向量发生扰动时, 给出了由贪婪随机 Kaczmarz 方法所产生的迭代解与原线性代数方程组的最小范数解之间的期望误差的上界, 并说明了随着迭代步数的增长, 该期望解误差以线性速率下降至一个给定阈值。数值实验表明, 该阈值能够很好地估计贪婪随机 Kaczmarz 方法的迭代解误差所能达到的最小值。

关键词: Kaczmarz 方法; 线性方程组; 扰动; 随机迭代; 收敛性质

中图分类号: O241.6

文献标志码: A

Greedy Randomized Kaczmarz Method for Solving Noisy Linear Systems

WU Wenting

(School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: When the right-hand side vector of the consistent system of linear equations is disturbed by noise, we give an upper bound for the error in expectation between the iteration vector generated by the greedy randomized Kaczmarz method and the least-norm solution of the noise-free system of linear equations, and illuminate that, as the iteration step increases, this solution error in expectation decreases to a given threshold with a linear rate. Numerical experiments show that this threshold can give a good estimate of minimum that the iterative solution error of the greedy randomized Kaczmarz method can reach.

Key words: Kaczmarz method; system of linear equations; noise; randomized iteration; convergence property

对于系数矩阵为 $A \in \mathbb{C}^{m \times n}$ 且右端向量为 $b \in \mathbb{C}^m$ 的大规模相容线性代数方程组

$$Ax = b \quad (1)$$

的求解, Kaczmarz 方法^[1]是经典的行处理迭代方法, 在信号与图像处理领域有着广泛的应用。其每步迭代只需按照给定的循环顺序选取系数矩阵的某一行, 并将当前迭代向量正交投影至由该行所形成的超平面上。Strohmer 等^[2]提出按照与系数矩阵每一行的欧氏范数平方成比例的概率准则随机选取系数矩阵的行, 得到了收敛更为快速的随机 Kaczmarz 方法。若用 $(\cdot)^*$ 表示相应矩阵或向量的共轭转置, 则当初始迭代向量在 A^* 的列空间中时, 随机 Kaczmarz 方法期望线性收敛^[2-5]到线性代数方程组(1)的最小欧氏范数解 $x_* = A^\dagger b$, 其中 A^\dagger 表示系数矩阵 A 的 Moore-Penrose 伪逆。当线性代数方程组(1)的右端向量发生扰动时, Needell^[6]给出了随机 Kaczmarz 方法的期望解误差的上界, 并说明了随着迭代步数的增长, 随机 Kaczmarz 方法的期望解误差会以线性速率下降至一个误差阈值。之后, Zouzias 等^[7]对 Needell 所提出的期望解误差的上界进行了改进。

影响随机 Kaczmarz 方法收敛速率的关键在于其中所蕴含的用于选取每步迭代所需调用的系数矩阵行的概率准则。为了提高随机 Kaczmarz 方法的收敛速率, Bai 等^[8-9]提出了一个可以获取每步迭代中残向量的模较大分量的概率准则, 并基于该概率准则构造出了贪婪随机 Kaczmarz 方法。当初始迭代向量在 A^* 的列空间时, 贪婪随机 Kaczmarz 方法所产生的迭代序列 $\{x_k\}_{k=0}^\infty$ 收敛到线性代数方程组(1)的最小范数解 $x_* = A^\dagger b$ 且具有期望线性收敛速率

$$E \|x_k - x_*\|_2 \leq \left[1 - \frac{1}{2} \left(\frac{1}{\gamma} \|A\|_F^2 + 1 \right) \frac{\lambda_{\min}(A^*A)}{\|A\|_F^2} \right]^{\frac{k-1}{2}} \left(1 - \frac{\lambda_{\min}(A^*A)}{\|A\|_F^2} \right)^{\frac{1}{2}} \|x_0 - x_*\|_2$$

收稿日期: 2021-05-27

基金项目: 国家自然科学基金(12001043, 12071472)

第一作者: 巫文婷(1992—), 女, 特别副研究员, 理学博士, 主要研究方向为数值分析与科学计算。

E-mail: wuwenting@bit.edu.cn



论文
拓展
介绍

其中 $\gamma = \|A\|_F^2 - \min_{1 \leq i \leq m} \|A^{(i)}\|_2^2$, 且 $\|A^{(i)}\|_2^2, \|A\|_F^2$ 和 $\lambda_{\min}(A^*A)$ 分别表示矩阵 A 第 i 行的欧氏范数平方、矩阵 A 的 Frobenius 范数平方和矩阵 A^*A 的最小非零特征值。

当线性代数方程组(1)的右端向量 b 被加上一个不为零的扰动向量 $r \in \mathbb{C}^m$ 时, 实际求解的线性代数方程组问题将变为

$$Ax = y \tag{2}$$

其中, $y = b + r$ 。若对任意矩阵 $G \in \mathbb{C}^{m \times n}$ 和任意向量 $u \in \mathbb{C}^m$, 用 $G^{(i)}$ 和 $u^{(i)}$ 分别表示矩阵 G 的第 i 行和向量 u 的第 i 个分量, 由于贪婪随机 Kaczmarz 方法的第 k 步迭代将当前迭代向量 x_k 投影至超平面 $A^{(i_k)}x = y^{(i_k)}$ 上, 而线性代数方程组(1)的最小范数解 $x_* = A^\dagger b$ 满足 $Ax_* = y - r$, 故其并不能收敛到 x_* 。在这种情况下, 本文对贪婪随机 Kaczmarz 方法的期望解误差进行分析, 得到了其上界, 证明了随着迭代步数的增长, 由贪婪随机 Kaczmarz 方法所产生的迭代解与原线性代数方程组(1)的最小范数解 x_* 之间的误差以线性速率下降至一个给定阈值。数值实验表明本文所给出的阈值能够很好地估计贪婪随机 Kaczmarz 方法的迭代解误差所能达到的最小值。

1 贪婪随机 Kaczmarz 方法

不失一般性, 在本文的讨论中, 总是假设系数矩阵 A 不存在零行, 即 $A^{(i)} \neq 0, i = 1, 2, \dots, m$ 。当线性代数方程组(1)的右端向量发生扰动时, 求解线性代数方程组(2)的贪婪随机 Kaczmarz 方法^[8]如下:

输入: A, y, ℓ 与 x_0

输出: x_ℓ

1: for $k = 0, 1, \dots, \ell - 1$ do

2: 计算

$$\epsilon_k = \frac{1}{2} \left(\frac{1}{\|y - Ax_k\|_2^2} \max_{1 \leq i \leq m} \left\{ \frac{|y^{(i)} - A^{(i)}x_k|^2}{\|A^{(i)}\|_2^2} \right\} + \frac{1}{\|A\|_F^2} \right)$$

3: 确定正整数指标集

$$U_k = \left\{ i \mid |y^{(i)} - A^{(i)}x_k|^2 \geq \epsilon_k \|y - Ax_k\|_2^2 \|A^{(i)}\|_2^2 \right\}$$

4: 计算向量 $\tilde{r}_k \in \mathbb{C}^m$ 的第 i 个分量

$$\tilde{r}_k^{(i)} = \begin{cases} y^{(i)} - A^{(i)}x_k, & \text{如果 } i \in U_k \\ 0, & \text{否则} \end{cases}$$

5: 按照概率 $P_r(\text{row} = i_k) = \frac{|\tilde{r}_k^{(i_k)}|^2}{\|\tilde{r}_k\|_2^2}$ 选取

$i_k \in U_k$

$$6: \text{ 令 } x_{k+1} = x_k + \frac{(y^{(i_k)} - A^{(i_k)}x_k)}{\|A^{(i_k)}\|_2^2} (A^{(i_k)})^*$$

7: endfor

令 $R(A)$ 和 $R(A)^\perp$ 分别为系数矩阵 A 的像空间和其像空间的正交补子空间, 则有 $r = r_{R(A)} + r_{R(A)^\perp}$, 其中 $r_{R(A)}$ 和 $r_{R(A)^\perp}$ 分别表示向量 r 在 $R(A)$ 和 $R(A)^\perp$ 上的正交投影。若线性代数方程组(1)右端向量的扰动 r 在系数矩阵 A 的像空间 $R(A)$ 中, 则线性代数方程组(2)等价于线性代数方程组

$$Ax = b + r_{R(A)} \tag{3}$$

易知线性代数方程组(3)是相容的且其最小范数解为 $\tilde{x}_* = A^\dagger (b + r_{R(A)})$ 。此时, 若初始迭代向量在 A^* 的列空间中, 则求解线性代数方程组(2)的贪婪随机 Kaczmarz 方法所产生的迭代序列 $\{x_k\}_{k=0}^\infty$ 期望线性收敛到 \tilde{x}_* 。

2 误差分析

与求解原线性代数方程组(1)的贪婪随机 Kaczmarz 方法的分析^[8]类似, 根据求解线性代数方程组(2)的贪婪随机 Kaczmarz 方法中 ϵ_k 的定义可知, 对于 $k = 1, 2, \dots$, 由于

$$y^{(i_{k-1})} - A^{(i_{k-1})}x_k = y^{(i_{k-1})} - A^{(i_{k-1})} \left(x_{k-1} + \frac{(y^{(i_{k-1})} - A^{(i_{k-1})}x_{k-1})}{\|A^{(i_{k-1})}\|_2^2} (A^{(i_{k-1})})^* \right) = 0$$

有

$$\epsilon_k \|A\|_F^2 = \frac{1}{2} \frac{\max_{1 \leq i \leq m} \left(\frac{|y^{(i)} - A^{(i)}x_k|^2}{\|A^{(i)}\|_2^2} \right)}{\sum_{i=1}^m \frac{\|A^{(i)}\|_2^2}{\|A\|_F^2} \frac{|y^{(i)} - A^{(i)}x_k|^2}{\|A^{(i)}\|_2^2}} + \frac{1}{2} = \frac{1}{2} \frac{\max_{1 \leq i \leq m} \left(\frac{|y^{(i)} - A^{(i)}x_k|^2}{\|A^{(i)}\|_2^2} \right)}{\sum_{i \neq i_{k-1}} \frac{\|A^{(i)}\|_2^2}{\|A\|_F^2} \frac{|y^{(i)} - A^{(i)}x_k|^2}{\|A^{(i)}\|_2^2}} + \frac{1}{2} \geq$$

$$\frac{1}{2} \left(\frac{\|A\|_F^2}{\sum_{i=1, i \neq k-1}^m \|A^{(i)}\|_2^2} + 1 \right)$$

而对于 $k=0$, 有

$$\epsilon_0 \|A\|_F^2 = \frac{1}{2} \frac{\max_{1 \leq i \leq m} \left(\frac{|\mathbf{y}^{(i)} - A^{(i)} \mathbf{x}_0|^2}{\|A^{(i)}\|_2^2} \right)}{\sum_{i=1}^m \frac{\|A^{(i)}\|_2^2}{\|A\|_F^2} \frac{|\mathbf{y}^{(i)} - A^{(i)} \mathbf{x}_0|^2}{\|A^{(i)}\|_2^2}} + \frac{1}{2} \geq 1$$

因此, 可得引理 1.

引理 1 求解系数矩阵为 $A \in \mathbb{C}^{m \times n}$ 且右端向量为 $\mathbf{y} \in \mathbb{C}^m$ 的线性代数方程组 (2) 的贪婪随机 Kaczmarz 方法的概率准则中的量 $\epsilon_k, k=0, 1, 2, \dots$, 满足

$$\epsilon_0 \geq \frac{1}{\|A\|_F^2}$$

和

$$\epsilon_k \geq \frac{1}{2} \left(\frac{1}{\gamma} + \frac{1}{\|A\|_F^2} \right), k=1, 2, \dots$$

其中 $\gamma = \|A\|_F^2 - \min_{1 \leq i \leq m} \|A^{(i)}\|_2^2$.

当线性代数方程组 (1) 右端向量的扰动 \mathbf{r} 在 $R(A)$ 中时, 若初始迭代向量在 A^* 的列空间中, 则求解线性代数方程组 (2) 的贪婪随机 Kaczmarz 方法所产生的迭代向量期望线性收敛到线性代数方程组 (3) 的最小范数解 $\tilde{\mathbf{x}}_*$. 对于一般的扰动情形, 求解线性代数方程组 (2) 的贪婪随机 Kaczmarz 方法所产生的迭代向量与 $\tilde{\mathbf{x}}_*$ 之间的期望误差的上界可以由引理 2 给出.

引理 2 如果初始迭代向量 $\mathbf{x}_0 \in \mathbb{C}^n$ 在 A^* 的列空间中, 则贪婪随机 Kaczmarz 方法求解扰动后的线性代数方程组 (2) 所产生的迭代序列 $\{\mathbf{x}_k\}_{k=0}^\infty$ 与线性代数方程组 (3) 的最小范数解 $\tilde{\mathbf{x}}_*$ 之间的期望误差满足

$$E \|\mathbf{x}_k - \tilde{\mathbf{x}}_*\|_2 \leq \sqrt{\alpha^{k-1} \alpha_0} \|\mathbf{x}_0 - \tilde{\mathbf{x}}_*\|_2 + \sqrt{\frac{\beta}{1-\alpha}}$$

其中

$$\alpha = 1 - \frac{\lambda_{\min}(A^*A)}{4} \left(\frac{1}{\gamma} + \frac{1}{\|A\|_F^2} \right),$$

$$\alpha_0 = 1 - \frac{\lambda_{\min}(A^*A)}{2\|A\|_F^2},$$

$$\beta = 2 \max_{1 \leq i \leq m} \left(\frac{|\mathbf{r}_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) - \frac{\|\mathbf{r}_{R(A)^\perp}\|_2^2}{2\|A\|_F^2},$$

$$\gamma = \|A\|_F^2 - \min_{1 \leq i \leq m} \|A^{(i)}\|_2^2 \tag{4}$$

证明: 由求解线性代数方程组 (2) 的贪婪随机 Kaczmarz 方法的定义和 $A\tilde{\mathbf{x}}_* = \mathbf{b} + \mathbf{r}_{R(A)}$ 可知

$$\begin{aligned} \mathbf{x}_{k+1} - \tilde{\mathbf{x}}_* &= \mathbf{x}_k - \tilde{\mathbf{x}}_* + \frac{\mathbf{y}^{(i_k)} - A^{(i_k)} \mathbf{x}_k}{\|A^{(i_k)}\|_2^2} (A^{(i_k)})^* = \\ \mathbf{x}_k - \tilde{\mathbf{x}}_* &+ \frac{\mathbf{b}^{(i_k)} + \mathbf{r}_{R(A)}^{(i_k)} - A^{(i_k)} \mathbf{x}_k + \mathbf{r}_{R(A)^\perp}^{(i_k)}}{\|A^{(i_k)}\|_2^2} (A^{(i_k)})^* = \end{aligned}$$

$$\left(I - \frac{(A^{(i_k)})^* A^{(i_k)}}{\|A^{(i_k)}\|_2^2} \right) (\mathbf{x}_k - \tilde{\mathbf{x}}_*) + \frac{\mathbf{r}_{R(A)^\perp}^{(i_k)}}{\|A^{(i_k)}\|_2^2} (A^{(i_k)})^*$$

其中 I 表示具有合适阶数的单位矩阵, 则有

$$\begin{aligned} \|\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_*\|_2^2 &= \left\| \left(I - \frac{(A^{(i_k)})^* A^{(i_k)}}{\|A^{(i_k)}\|_2^2} \right) (\mathbf{x}_k - \tilde{\mathbf{x}}_*) \right\|_2^2 + \\ &\frac{|\mathbf{r}_{R(A)^\perp}^{(i_k)}|^2}{\|A^{(i_k)}\|_2^2} \end{aligned} \tag{5}$$

用 E_k 表示固定前 k 步迭代的条件期望, 即 $E_k[\cdot] = E[\cdot | i_0, i_1, \dots, i_{k-1}]$, 其中 $i_l (l=0, 1, \dots, k-1)$ 为第 l 步迭代所选取的行指标, 则对等式 (5) 两边取条件期望可得

$$\begin{aligned} E_k \|\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_*\|_2^2 &= \|\mathbf{x}_k - \tilde{\mathbf{x}}_*\|_2^2 - \\ &\sum_{i_k \in U_k} \frac{|\mathbf{y}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2}{\sum_{i \in U_k} |\mathbf{y}^{(i)} - A^{(i)} \mathbf{x}_k|^2} \frac{|\mathbf{b}^{(i_k)} + \mathbf{r}_{R(A)}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2}{\|A^{(i_k)}\|_2^2} + \\ &\sum_{i_k \in U_k} \frac{|\mathbf{y}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2}{\sum_{i \in U_k} |\mathbf{y}^{(i)} - A^{(i)} \mathbf{x}_k|^2} \frac{|\mathbf{r}_{R(A)^\perp}^{(i_k)}|^2}{\|A^{(i_k)}\|_2^2} \end{aligned}$$

由于

$$\begin{aligned} |\mathbf{y}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2 &= |\mathbf{b}^{(i_k)} + \mathbf{r}_{R(A)}^{(i_k)} + \mathbf{r}_{R(A)^\perp}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2 \leq \\ &\left(|\mathbf{b}^{(i_k)} + \mathbf{r}_{R(A)}^{(i_k)} - A^{(i_k)} \mathbf{x}_k| + |\mathbf{r}_{R(A)^\perp}^{(i_k)}| \right)^2 \leq \\ &2 |\mathbf{b}^{(i_k)} + \mathbf{r}_{R(A)}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2 + 2 |\mathbf{r}_{R(A)^\perp}^{(i_k)}|^2 \end{aligned}$$

可知

$$|\mathbf{b}^{(i_k)} + \mathbf{r}_{R(A)}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2 \geq \frac{1}{2} |\mathbf{y}^{(i_k)} - A^{(i_k)} \mathbf{x}_k|^2 - |\mathbf{r}_{R(A)^\perp}^{(i_k)}|^2$$

则根据正整数指标集 U_k 的定义可得

$$\begin{aligned} E_k \|x_{k+1} - \tilde{x}_*\|_2^2 &\leq \|x_k - \tilde{x}_*\|_2^2 - \\ &\frac{1}{2} \sum_{i_k \in U_k} \frac{|y^{(i_k)} - A^{(i_k)} x_k|^2}{\sum_{i \in U_k} |y^{(i)} - A^{(i)} x_k|^2} \frac{|y^{(i_k)} - A^{(i_k)} x_k|^2}{\|A^{(i_k)}\|_2^2} + \\ &2 \sum_{i_k \in U_k} \frac{|y^{(i_k)} - A^{(i_k)} x_k|^2}{\sum_{i \in U_k} |y^{(i)} - A^{(i)} x_k|^2} \frac{|r_{R(A)^\perp}^{(i_k)}|^2}{\|A^{(i_k)}\|_2^2} \leq \\ &\|x_k - \tilde{x}_*\|_2^2 - \frac{1}{2} \varepsilon_k \|y - Ax_k\|_2^2 + \\ &2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) \end{aligned}$$

对于在矩阵 A^* 的像空间 $R(A^*)$ 中的任意向量 u , 有不等式

$$\|Au\|_2^2 \geq \lambda_{\min}(A^*A) \|u\|_2^2 \quad (6)$$

成立。利用该不等式, 通过引理 1、向量 $b + r_{R(A)} - Ax_k = A(\tilde{x}_* - x_k) \in R(A)$ 和向量 $r_{R(A)^\perp}$ 的正交性、 $x_k - \tilde{x}_* \in R(A^*)$ 以及 $\gamma \leq \|A\|_F^2$ 可知, 对于 $k = 1, 2, \dots$, 有

$$\begin{aligned} E_k \|x_{k+1} - \tilde{x}_*\|_2^2 &\leq \|x_k - \tilde{x}_*\|_2^2 - \frac{1}{4} \left(\frac{1}{\gamma} + \right. \\ &\left. \frac{1}{\|A\|_F^2} \right) \|y - Ax_k\|_2^2 + 2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) = \\ &\|x_k - \tilde{x}_*\|_2^2 - \frac{1}{4} \left(\frac{1}{\gamma} + \frac{1}{\|A\|_F^2} \right) (\|b + r_{R(A)} - Ax_k\|_2^2 + \\ &\|r_{R(A)^\perp}\|_2^2) + 2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) \leq \\ &\left[1 - \frac{\lambda_{\min}(A^*A)}{4} \left(\frac{1}{\gamma} + \frac{1}{\|A\|_F^2} \right) \right] \|x_k - \tilde{x}_*\|_2^2 + \\ &2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) - \frac{1}{4} \left(\frac{1}{\gamma} + \frac{1}{\|A\|_F^2} \right) \|r_{R(A)^\perp}\|_2^2 \leq \\ &\left[1 - \frac{\lambda_{\min}(A^*A)}{4} \left(\frac{1}{\gamma} + \frac{1}{\|A\|_F^2} \right) \right] \|x_k - \tilde{x}_*\|_2^2 + \\ &2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) - \frac{\|r_{R(A)^\perp}\|_2^2}{2\|A\|_F^2} \end{aligned}$$

而对于 $k=0$, 有

$$\begin{aligned} E \|x_1 - \tilde{x}_*\|_2^2 &\leq \|x_0 - \tilde{x}_*\|_2^2 - \frac{1}{2\|A\|_F^2} \|y - Ax_0\|_2^2 + \\ &2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) = \\ &\|x_0 - \tilde{x}_*\|_2^2 - \frac{1}{2\|A\|_F^2} (\|b + r_{R(A)} - Ax_0\|_2^2 + \\ &\|r_{R(A)^\perp}\|_2^2) + 2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) \leq \\ &\left[1 - \frac{\lambda_{\min}(A^*A)}{2\|A\|_F^2} \right] \|x_0 - \tilde{x}_*\|_2^2 + \\ &2 \max_{1 \leq i \leq m} \left(\frac{|r_{R(A)^\perp}^{(i)}|^2}{\|A^{(i)}\|_2^2} \right) - \frac{\|r_{R(A)^\perp}\|_2^2}{2\|A\|_F^2} \end{aligned}$$

对不等式两边取期望可得对于 $k = 1, 2, \dots$, 有

$$E \|x_{k+1} - \tilde{x}_*\|_2^2 \leq \alpha E \|x_k - \tilde{x}_*\|_2^2 + \beta$$

而对于 $k=0$, 有

$$E \|x_1 - \tilde{x}_*\|_2^2 \leq \alpha_0 \|x_0 - \tilde{x}_*\|_2^2 + \beta$$

因此, 由于 $0 < \alpha < 1, \beta \geq 0$, 对于 $k = 1, 2, \dots$, 有

$$\begin{aligned} E \|x_k - \tilde{x}_*\|_2^2 &\leq \alpha E \|x_{k-1} - \tilde{x}_*\|_2^2 + \beta \leq \\ &\alpha^2 E \|x_{k-2} - \tilde{x}_*\|_2^2 + \alpha\beta + \beta \leq \dots \leq \\ &\alpha^{k-1} E \|x_1 - \tilde{x}_*\|_2^2 + \left(\sum_{i=0}^{k-2} \alpha^i \right) \beta \leq \\ &\alpha^{k-1} \alpha_0 \|x_0 - \tilde{x}_*\|_2^2 + \left(\sum_{i=0}^{k-1} \alpha^i \right) \beta \leq \\ &\alpha^{k-1} \alpha_0 \|x_0 - \tilde{x}_*\|_2^2 + \frac{\beta}{1-\alpha} \end{aligned}$$

由 Jensen 不等式可知

$$\begin{aligned} E \|x_k - \tilde{x}_*\|_2 &\leq \left(E \|x_k - \tilde{x}_*\|_2^2 \right)^{\frac{1}{2}} \leq \\ &\left(\alpha^{k-1} \alpha_0 \|x_0 - \tilde{x}_*\|_2^2 + \frac{\beta}{1-\alpha} \right)^{\frac{1}{2}} \leq \\ &\sqrt{\alpha^{k-1} \alpha_0} \|x_0 - \tilde{x}_*\|_2 + \sqrt{\frac{\beta}{1-\alpha}} \end{aligned}$$

基于引理 2, 关于求解扰动后的线性代数方程组 (2) 的贪婪随机 Kaczmarz 方法所产生的迭代近似解与原线性代数方程组 (1) 的最小范数解 x_* 之间的期望误差的上界, 可以给出如下定理。

定理 1 如果初始迭代向量 $x_0 \in \mathbb{C}^n$ 在 A^* 的列空

间中且 \tilde{x}_* 为线性代数方程组(3)的最小范数解,则贪婪随机 Kaczmarz 方法求解扰动后的线性代数方程组(2)所产生的迭代序列 $\{x_k\}_{k=0}^{\infty}$ 与原线性代数方程组(1)的最小范数解 $x_* = A^\dagger b$ 之间的期望误差满足

$$E\|x_k - x_*\|_2 \leq \sqrt{\alpha^{k-1}\alpha_0}\|x_0 - \tilde{x}_*\|_2 + \sqrt{\frac{\beta}{1-\alpha}} + \frac{\|r_{R(A)}\|_2}{\sqrt{\lambda_{\min}(A^*A)}}$$

其中, α 、 α_0 和 β 如(4)式中定义。

证明: 因为 $\tilde{x}_* - x_* \in R(A^*)$,则通过不等式(6)可知

$$\|r_{R(A)}\|_2 = \|A(\tilde{x}_* - x_*)\|_2 \geq \sqrt{\lambda_{\min}(A^*A)}\|\tilde{x}_* - x_*\|_2$$

又因为有

$$\|x_k - x_*\|_2 = \|x_k - \tilde{x}_* + \tilde{x}_* - x_*\|_2 \leq \|x_k - \tilde{x}_*\|_2 + \|\tilde{x}_* - x_*\|_2$$

成立,则由引理2可得

$$E\|x_k - x_*\|_2 \leq E\|x_k - \tilde{x}_*\|_2 + \|\tilde{x}_* - x_*\|_2 \leq \sqrt{\alpha^{k-1}\alpha_0}\|x_0 - \tilde{x}_*\|_2 + \sqrt{\frac{\beta}{1-\alpha}} + \frac{\|r_{R(A)}\|_2}{\sqrt{\lambda_{\min}(A^*A)}}$$

定理1说明了当迭代步数 k 趋于无穷时,贪婪随机 Kaczmarz 方法求解扰动后的线性代数方程组(2)所产生的迭代序列 $\{x_k\}_{k=0}^{\infty}$ 的期望相对解误差 $E\|x_k - x_*\|_2/\|x_*\|_2$ 以 $\sqrt{\alpha}$ 的线性速率下降至某个阈值,并给出了该阈值的估计

$$\tau = \frac{1}{\|x_*\|_2} \left(\sqrt{\frac{\beta}{1-\alpha}} + \frac{\|r_{R(A)}\|_2}{\sqrt{\lambda_{\min}(A^*A)}} \right)$$

若 $r_{R(A)^\perp}$ 为零,则扰动后的线性代数方程组(2)即为相容的线性代数方程组(3)。因此,求解线性代数方程组(2)的贪婪随机 Kaczmarz 方法的迭代序列 $\{x_k\}_{k=0}^{\infty}$ 收敛到线性代数方程组(3)的最小范数解 \tilde{x}_* ,且其与原线性代数方程组(1)的最小范数解 x_* 的差趋于 $\tilde{x}_* - x_*$ 。此时,如果初始迭代向量在 A^* 的列空间中,则求解线性代数方程组(2)的贪婪随机 Kaczmarz 方法所产生的迭代序列 $\{x_k\}_{k=0}^{\infty}$ 满足

$$\frac{E\|x_k - x_*\|_2}{\|x_*\|_2} \leq \sqrt{\alpha^{k-1}\alpha_0} \frac{\|x_0 - \tilde{x}_*\|_2}{\|x_*\|_2} +$$

$$\frac{\|\tilde{x}_* - x_*\|_2}{\|x_*\|_2} \leq \sqrt{\alpha^{k-1}\alpha_0} \frac{\|x_0 - \tilde{x}_*\|_2}{\|x_*\|_2} + \frac{\|r_{R(A)}\|_2}{\|x_*\|_2 \sqrt{\lambda_{\min}(A^*A)}}$$

若 $r_{R(A)}$ 为零,则线性代数方程组(3)即为线性代数方程组(1),且线性代数方程组(3)的最小范数解 \tilde{x}_* 即为 $x_* = A^\dagger b$ 。此时,如果初始迭代向量在 A^* 的列空间中,则求解线性代数方程组(2)的贪婪随机 Kaczmarz 方法所产生的迭代序列 $\{x_k\}_{k=0}^{\infty}$ 满足

$$\frac{E\|x_k - x_*\|_2}{\|x_*\|_2} \leq \sqrt{\alpha^{k-1}\alpha_0} \frac{\|x_0 - x_*\|_2}{\|x_*\|_2} + \frac{1}{\|x_*\|_2} \sqrt{\frac{\beta}{1-\alpha}}$$

3 数值实验

通过数值实验测试贪婪随机 Kaczmarz 方法求解带右端项扰动的线性代数方程组(2)时的数值表现,并将所估计的阈值与贪婪随机 Kaczmarz 方法所产生的真实迭代解误差进行比较。

所测试的系数矩阵 A 分为两类:一类为利用 MATLAB 函数 randn 随机产生的元素服从标准正态分布的矩阵,矩阵维数分别为 $200 \times 100\,000$ 和 $100\,000 \times 200$;另一类为取自稀疏矩阵库^[10]的稀疏矩阵 bibd_16_8 和 ash958,矩阵维数分别为 $120 \times 12\,870$ 和 958×292 。

线性代数方程组(1)的右端向量 b 取为 Ax_* ,其中 $x_* \in \mathbb{R}^n$ 是利用 MATLAB 函数 randn 随机产生的。线性代数方程组(1)的右端项的扰动向量 r 分为三类:一类为随机扰动,一类为在系数矩阵的像空间中的扰动,另一类为在系数矩阵的像空间的正交补空间中的扰动。这三类扰动均由 MATLAB 函数 randn 生成,并且其欧氏范数为原始右端向量 b 的欧氏范数的0.0005倍。由于 $200 \times 100\,000$ 的随机矩阵和稀疏矩阵 bibd_16_8 为行满秩矩阵,故其所对应的右端项扰动 r 一定在其像空间中。所有计算均从初始向量 $x_0 = 0$ 开始。

图1描绘了当线性代数方程组(1)的系数矩阵为随机产生的矩阵且右端项扰动 r 为随机扰动时,贪婪随机 Kaczmarz 方法重复运行50次所产生的相对解误差 E_{RS} 的中值和阈值 τ 分别取以10为底的对数

之后相对于迭代步数的曲线, 其中

$$E_{RS} = \frac{\|x_k - x_*\|_2}{\|x_*\|_2}$$

从图 1 可以看出, 贪婪随机 Kaczmarz 方法所产生的

相对解误差下降至 10^{-3} 左右之后不再减小, 且阈值 τ 能够很好地给出该最小值的估计。特别地, 当系数矩阵为 $200 \times 100\,000$ 的随机矩阵时, τ 与真实相对解误差所能达到的最小值非常接近。

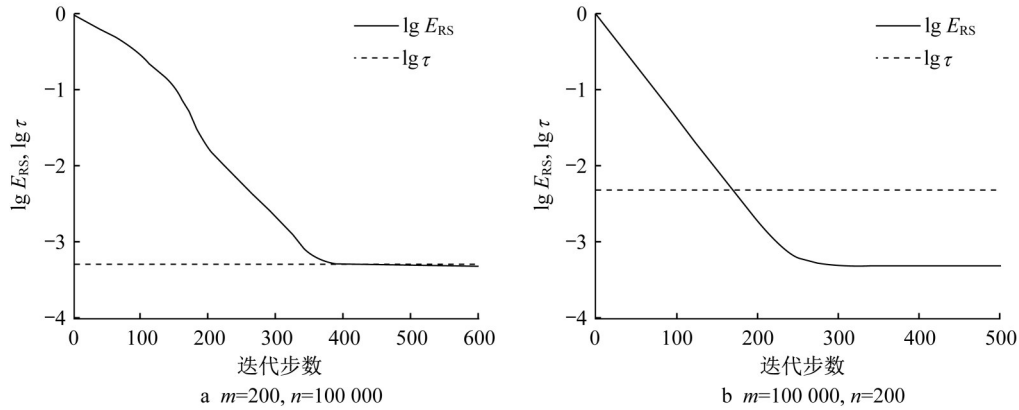


图 1 当 $m = 200, n = 100\,000$ 和 $m = 100\,000, n = 200$ 时, $\lg E_{RS}$ 和 $\lg \tau$ 相对于迭代步数的曲线

Fig.1 Pictures of $\lg E_{RS}$ and $\lg \tau$ versus the iteration step when $m = 200, n = 100\,000$ and $m = 100\,000, n = 200$

图 2 描绘了对于 $100\,000 \times 200$ 的随机产生的系数矩阵, 当线性代数方程组 (1) 的右端项扰动 r 分别在系数矩阵的像空间和系数矩阵的像空间的正交补空间中时, 贪婪随机 Kaczmarz 方法重复运行 50 次所产生的相对解误差的中值和阈值 τ 分别取以 10 为底

的对数之后相对于迭代步数的曲线。从图 2 可以看出, 贪婪随机 Kaczmarz 方法所产生的相对解误差下降至 10^{-3} 左右之后不再减小, 且阈值 τ 能够很好地给出该最小值的估计, 特别当扰动向量 r 在系数矩阵的像空间中时。

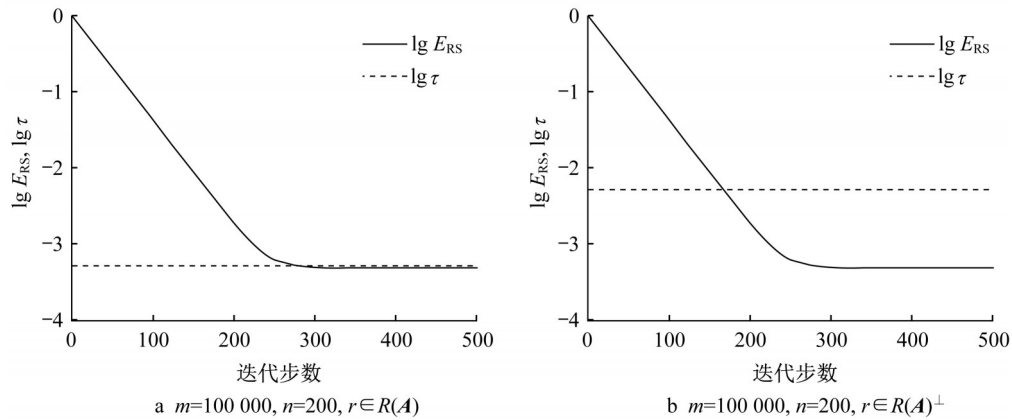


图 2 当 $m = 100\,000, n = 200$ 且 $r \in R(A)$ 和 $r \in R(A)^\perp$ 时, $\lg E_{RS}$ 和 $\lg \tau$ 相对于迭代步数的曲线

Fig.2 Pictures of $\lg E_{RS}$ and $\lg \tau$ versus the iteration step when $m = 100\,000, n = 200$, and $r \in R(A)$ and $r \in R(A)^\perp$

当线性代数方程组 (1) 的系数矩阵为稀疏矩阵 bibd_16_8 和 ash958 时, 图 3 描绘了右端项扰动 r 为随机扰动时, 贪婪随机 Kaczmarz 方法重复运行 50 次所产生的相对解误差的中值和阈值 τ 分别取以 10 为底的对数之后相对于迭代步数的曲线; 图 4 描绘了右端项扰动 r 分别在系数矩阵的像空间和系数矩阵的像空间的正交补空间中时的相应曲线。类似地, 从图 3 和图 4 可以看出, 贪婪随机 Kaczmarz 方法所

产生的相对解误差下降至 10^{-3} 左右之后不再减小, 且阈值 τ 能够很好地给出该最小值的估计, 特别当系数矩阵为 bibd_16_8 和当系数矩阵为 ash958 且扰动向量 r 在系数矩阵的像空间中时。

4 结论

当线性代数方程组的右端向量发生扰动时, 证

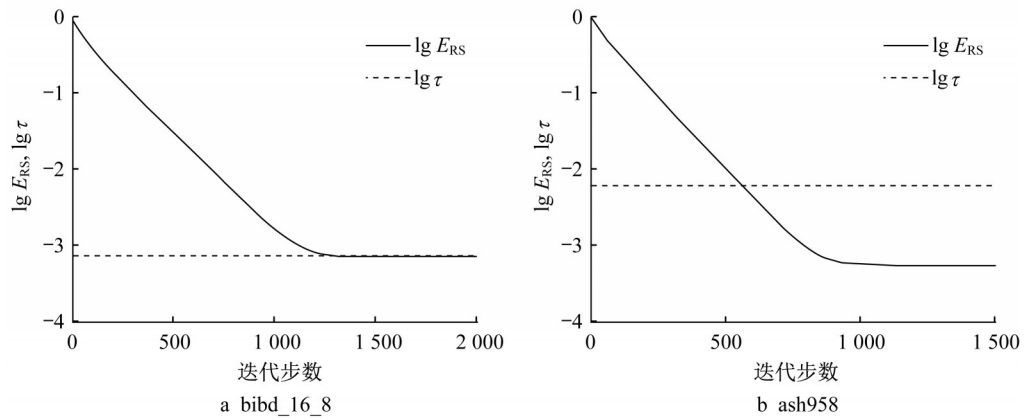


图3 当矩阵为 bibd_16_8 和 ash958 时,lg E_{RS}和lg τ 相对于迭代步数的曲线

Fig.3 Pictures of lg E_{RS} and lg τ versus the iteration step when the matrices are bibd_16_8 and ash958

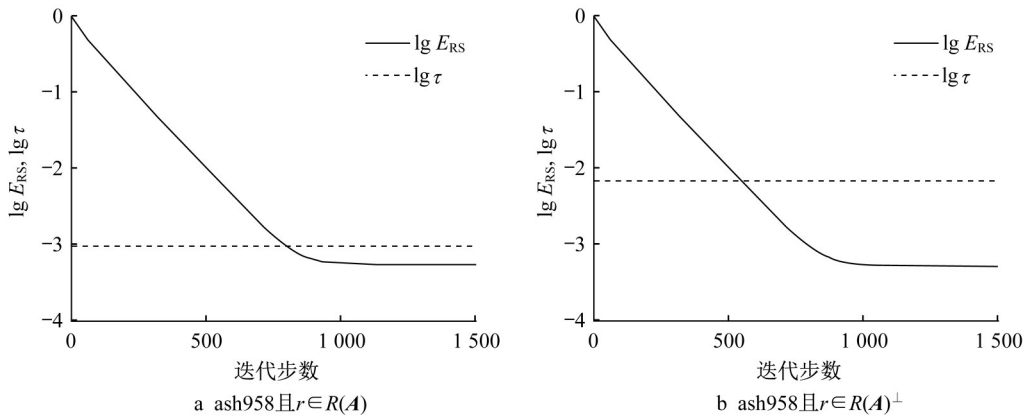


图4 当矩阵为 ash958 且 $r \in R(A)$ 和 $r \in R(A)^\perp$ 时,lg E_{RS}与lg τ 相对于迭代步数的曲线

Fig.4 Pictures of lg E_{RS} and lg τ versus the iteration step when the matrix is ash958, and $r \in R(A)$ and $r \in R(A)^\perp$

明了贪婪随机Kaczmarz方法的期望解误差以线性速率下降至一个给定阈值。数值实验表明该阈值能够很好地估计贪婪随机Kaczmarz方法的迭代解误差所能达到的最小值。可以发现当扰动不是很大且对解的精度要求不是很高时,贪婪随机Kaczmarz方法仍然能够很好地给出原线性代数方程组最小范数解的近似。

参考文献:

[1] KACZMARZ S. Angenäherte Auflösung von Systemen linearer Gleichungen [J]. Bulletin International de l'Academie Polonaise des Sciences et des Lettres, 1937, 35: 355.
 [2] STROHMER T, VERSHYNIN R. A randomized Kaczmarz algorithm with exponential convergence [J]. Journal of Fourier Analysis and Applications, 2009, 15: 262.
 [3] MA A, NEEDELL D, RAMDAS A. Convergence properties of the randomized extended Gauss-Seidel and Kaczmarz methods

[J]. SIAM Journal on Matrix Analysis and Applications, 2015, 36: 1590.
 [4] GOWER R M, RICHTÁRIK P. Stochastic dual ascent for solving linear systems [J]. arXiv, 2015, 1512.06890.
 [5] BAI Z Z, WU W T. On convergence rate of the randomized Kaczmarz method [J]. Linear Algebra and its Applications, 2018, 553: 252.
 [6] NEEDELL D. Randomized Kaczmarz solver for noisy linear systems [J]. BIT Numerical Mathematics, 2010, 50: 395.
 [7] ZOUZIAS A, FRERIS N M. Randomized extended Kaczmarz for solving least squares [J]. SIAM Journal on Matrix Analysis and Applications, 2013, 34: 773.
 [8] BAI Z Z, WU W T. On greedy randomized Kaczmarz method for solving large sparse linear systems [J]. SIAM Journal on Scientific Computing, 2018, 40: A592.
 [9] BAI Z Z, WU W T. On relaxed greedy randomized Kaczmarz methods for solving large sparse linear systems [J]. Applied Mathematics Letters, 2018, 83: 21.
 [10] DAVIS T A, HU Y. The University of Florida sparse matrix collection [J]. ACM Transactions on Mathematical Software, 2011, 38: 1.