Bull. Malays. Math. Sci. Soc. (2024) 47:148
https://doi.org/10.1007/s40840-024-01739-8

BULLETIN of the
MALAYSIAN MATHEMATICAL
SCIENCES SOCIETY
http://www.springer.com/mathematics/journal/40840

Greedy Block Extended Kaczmarz Method for Solving the Least Squares Problems

Ni-Hong Ke[1,2] · Rui Li[3] · Jun-Feng Yin[1,2]

Received: 18 January 2024 / Revised: 14 May 2024 / Accepted: 25 June 2024 /
Published online: 29 July 2024

Abstract

A greedy block extended Kaczmarz method is introduced for solving the least squares problem where the greedy rule combines the maximum-distances with relaxation parameters. In order to save the computational cost of Moore–Penrose inverse, an average projection technique is used. The convergence theory of the greedy block extended Kaczmarz method is established and an upper bound for the convergence rate is also derived. Numerical experiments show that the proposed method is efficient and better than the randomized block extended Kaczmarz methods in terms of the number of iteration steps and computational time.

Keywords Least squares · Randomized extended Kaczmarz method · Greedy block · Convergence

Mathematics Subject Classification 65F10 · 65F20

Communicated by Rosihan M. Ali.

✉ Jun-Feng Yin
yinjf@tongji.edu.cn

Ni-Hong Ke
kenihong@tongji.edu.cn

Rui Li
zuqiuziye@163.com

[1] School of Mathematical Sciences, Tongji University, Shanghai 200092, China

[2] Key Laboratory of Intelligent Computing and Applications (Ministry of Education), Shanghai 200092, China

[3] College of Data Science, Jiaxing University, Jiaxing 314001, China

1 Introduction

Consider the solution of the least squares problem

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|_2^2, \tag{1.1}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, which widely arises from many scientific and engineering computing fields, such as image reconstruction [15], big data analysis[6] and optimization [16].

Iterative methods, particularly stochastic iterative methods, recently attract much attention in solving the least squares problem, as direct methods such as QR decomposition and singular value decomposition are usually expensive due to the memory and computational cost. One stochastic iterative method is the randomized extended Kaczmarz (REK) method [26], which was proved to have an exponential convergence in expectation towards the least squares solution $x_{LS} = A^{\dagger}b$ of (1.1). To accelerate the randomized extended Kaczmarz method, the randomized double block Kaczmarz (RDBK) method was introduced in [20] by selecting multiple rows and columns for projection. In order to save the computational cost of the Moore-Penrose inverse, the randomized extended average block Kaczmarz method [10] and the extended randomized multiple rows method [23] were presented and well studied. For more research on the randomized extended Kaczmarz method, we refer the readers to [3, 4, 9, 24].

Greedy techniques including maximizing the distance and the residual were firstly considered for Schwarz method in [18] and were proved to improve the efficiency of Kaczmarz methods [11]. A greedy randomized Kaczmarz method was proposed in [2] by using the combination of the maximum distance and the average distance to construct a novel greedy strategy. Further, a different greedy randomized Kaczmarz method was presented and studied in [21] with the maximum distance and a relaxation parameter. From a geometric point of view, a geometric probability randomized Kaczmarz method and its greedy version were established in [25].

In order to improve the performance of the randomized block extended Kaczmarz methods, a greedy block extended Kaczmarz method is proposed for solving the least squares problem, and the average block projection is used to save the computational cost. The convergence theory of the greedy block extended Kaczmarz method is established and an upper bound for the convergence rate is derived and analyzed in details. Numerical experiments show that the proposed method is efficient and better than the existing randomized block extended Kaczmarz methods.

The rest of the paper is organized as follows. In Sect. 2, the greedy block extended Kaczmarz method is presented and its convergence theory is established. Numerical experiments are provided in Sect. 3 to illustrate the efficiency and excellent performance of the proposed method. Finally we conclude the paper with a brief summary in Sect. 4.

2 The Greedy Block Extended Kaczmarz Method

This section introduces the greedy block extended Kaczmarz method for solving the least squares problem and establishes its convergence theory.

In the past decade, a number of greedy rules are proposed and studied, for instance, maximizing the distance, the residual and the geometric angles, and different greedy strategies usually lead to different block iterative methods [2, 21, 25].

In this paper, the greedy rule of maximum-distances with relaxation parameter is utilized for both row and column projections. At the $(k+1)$-th iteration, the column block $\mathcal{J}_k$ and row block $\mathcal{I}_k$ are selected as follows:

$$\mathcal{J}_k = \{j \in [n] : |A_{(j)}^T z^{(k)}|^2 \geq \tilde{\varepsilon}_k \|A_{(j)}\|_2^2\},$$
$$\mathcal{I}_k = \{i \in [m] : |b_i - z_i^{(k+1)} - (A^{(i)})^T x^{(k)}|^2 \geq \varepsilon_k \|A^{(i)}\|_2^2\},$$

where $A^{(i)}$, $A_{(j)}$ denote the i-th row and the j-th column of A respectively, $[m]$ represents the set $\{1, 2, \ldots, m\}$, and

$$\tilde{\varepsilon}_k = \rho_y \max_{j \in [n]} \left\{ \frac{|A_{(j)}^T z^{(k)}|^2}{\|A_{(j)}\|_2^2} \right\}, \quad \varepsilon_k = \rho_x \max_{i \in [m]} \left\{ \frac{|b(i) - z^{(k+1)}(i) - (A^{(i)})^T x^{(k)}|^2}{\|A^{(i)}\|_2^2} \right\},$$
$$\rho_y, \rho_x \in (0, 1].$$

The condition $\rho_y, \rho_x \in (0, 1]$ guarantees that $\mathcal{J}_k$ and $\mathcal{I}_k$ are non-empty sets. Without pre-partitioning the rows and columns of A, the blocks $\mathcal{J}_k$ and $\mathcal{I}_k$ are adaptive and made up of the larger entries of the distance vectors at each iteration.

By combining the above greedy selection rule with the average block projection technique, the greedy block extended Kaczmarz method is proposed and described in detail in Algorithm 1.

In Algorithm 1, $\tilde{\eta}_k$ and η_k are sparse residual vectors used to create two linear combinations of rows in $A_{:,\mathcal{J}_k}^T$ and $A_{\mathcal{I}_k,:}$ respectively. These linear combinations serves as the direction of row and column projections, thus the computaion of Moore-Penrose inverse is not required.

Before discussion of the convergence property of the greedy block extended Kaczmarz method, the following useful lemma is introduced.

Lemma 1 *([10]). Let* $A \in \mathbb{R}^{m \times n}$ *and* $\operatorname{rank}(A) = r$. *For any* $u \in R(A)$, *it holds that*

$$\sigma_1^2(A)\|u\|_2^2 \geq \|A^T u\|_2^2 \geq \sigma_r^2(A)\|u\|_2^2,$$

where $\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_r(A) > 0$ *denote all the nonzero singular values of* A.

Denote $R(A)^{\perp}$ as the orthogonal complement of the column space of A and $b_{R(A)^{\perp}}$ as the orthogonal projection of b onto $R(A)^{\perp}$. The convergence theory of the sequence $\{z^{(k)}\}_{k=0}^{\infty}$ generated by Algorithm 1 is established as follows.

Algorithm 1 (The greedy block extended Kaczmarz method)

Require: $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, the maximum number of iteration steps ℓ, , initial guess $x^{(0)} = 0, z^{(0)} = b$ and relaxation parameters $\rho_x, \rho_z \in (0, 1]$
Ensure: $x^{(k+1)}$
1: **for** $k = 0, 1, 2, \dots, \ell - 1$ **do**
2: Compute $\tilde{\varepsilon}_k = \rho_z \max\limits_{j\in[n]} \left\{ \frac{|A_{(j)}^T z^{(k)}|^2}{\|A_{(j)}\|_2^2} \right\}$
3: Determine the block $\mathcal{J}_k = \{j : |A_{(j)}^T z^{(k)}|^2 \geq \tilde{\varepsilon}_k \|A_{(j)}\|_2^2\}$
4: Set $\tilde{\eta}_k = \sum\limits_{j\in\mathcal{J}_k} (-A_{(j)}^T z^{(k)}) e_j$
5: Update $z^{(k+1)} = z^{(k)} - \frac{\tilde{\eta}_k^T A^T z^{(k)}}{\|A\tilde{\eta}_k\|_2^2} A\tilde{\eta}_k$
6: Compute $\varepsilon_k = \rho_x \max\limits_{i\in[m]} \left\{ \frac{|b_i - z_i^{(k+1)} - (A^{(i)})^T x^{(k)}|^2}{\|A^{(i)}\|_2^2} \right\}$
7: Determine the block $\mathcal{I}_k = \{i : |b_i - z_i^{(k+1)} - (A^{(i)})^T x^{(k)}|^2 \geq \varepsilon_k \|A^{(i)}\|_2^2\}$
8: Set $\eta_k = \sum\limits_{i\in\mathcal{I}_k} (b_i - z_i^{(k+1)} - (A^{(i)})^T x^{(k)}) e_i$
9: Update $x^{(k+1)} = x^{(k)} + \frac{\eta_k^T (b - z^{(k+1)} - Ax^{(k)})}{\|A^T \eta_k\|_2^2} A^T \eta_k$
10: **end for**

Theorem 1 *The sequence $\{z^{(k)}\}_{k=0}^{\infty}$ generated by the GBEK method converges to $z^* = b_{R(A)^\perp}$. Moreover, it holds that*

$$\|z^{(k+1)} - z^*\|_2^2 \leq \left(1 - \frac{\rho_z \sigma_r^2(A)}{\|A\|_F^2 - \tilde{\phi}_{\min}}\right)^{k+1} \|z^{(0)} - z^*\|_2^2, \tag{2.1}$$

where $\tilde{\phi}_{\min} = \min\limits_{j\in[n]} \|A_{(j)}\|_2^2$.

Proof By subtracting $z^* = b_{R(A)^\perp}$ from both sides of step 5 in Algorithm 1, we get

$$z^{(k+1)} - z^* = z^{(k)} - z^* - \frac{\tilde{\eta}_k^T A^T z^{(k)}}{\|A\tilde{\eta}_k\|_2^2} A\tilde{\eta}_k.$$

Let $\tilde{P}_k := \frac{A\tilde{\eta}_k \tilde{\eta}_k^T A^T}{\|A\tilde{\eta}_k\|_2^2}$, and $\tilde{P}_k$ is an orthogonal projection. Due to the fact that $\tilde{P}_k z^* = 0$, it holds that

$$z^{(k+1)} - z^* = z^{(k)} - z^* - \tilde{P}_k (z^{(k)} - z^*).$$

Let $\tilde{e}^{(k)} = z^{(k)} - z^*$. Taking the norm of both sides of the above equality yields:

$$\|\tilde{e}^{(k+1)}\|_2^2 = \|\tilde{e}^{(k)}\|_2^2 - \|\tilde{P}_k \tilde{e}^{(k)}\|_2^2. \tag{2.2}$$

Note that $\tilde{e}^{(0)} = z^{(0)} - z^* = AA^\dagger b \in R(A)$ and $\tilde{P}_k \tilde{e}^{(k)} \in R(A)$, then it follows that $\tilde{e}^{(k+1)} \in R(A)$. For the second term on the right side of (2.2), it holds that

$$\|\tilde{P}_k \tilde{e}^{(k)}\|_2^2 = \frac{(\tilde{\eta}_k^T A^T \tilde{e}^{(k)})^2}{\|A\tilde{\eta}_k\|_2^2} \geq \frac{\|(A_{:,\mathcal{J}_k})^T z^{(k)}\|_2^4}{\sigma_1^2(A_{:,\mathcal{J}_k})\|\tilde{\eta}_k\|_2^2} \geq \frac{\tilde{\varepsilon}_k \|A_{:,\mathcal{J}_k}\|_F^2}{\sigma_1^2(A_{:,\mathcal{J}_k})}, \tag{2.3}$$

where the first inequality holds because $\|A\tilde{\eta}_k\|_2^2 = \|(A_{:,\mathcal{J}_k})^T \tilde{\eta}_k\|_2^2 \leq \sigma_1^2(A_{:,\mathcal{J}_k})\|\tilde{\eta}_k\|_2^2$ and the second inequality holds because $\|\tilde{\eta}_k\|_2^2 = \|(A_{:,\mathcal{J}_k})^T z^{(k)}\|_2^2 \geq \tilde{\varepsilon}_k \|A_{:,\mathcal{J}_k}\|_F^2$. Note that

$$\tilde{\eta}_{k-1}^T (A^T z_k) = \tilde{\eta}_{k-1}^T A^T \left(z^{(k-1)} - \frac{\tilde{\eta}_{k-1}^T A^T z^{(k-1)}}{\|A\tilde{\eta}_{k-1}\|_2^2} A\tilde{\eta}_{k-1} \right) = 0,$$

therefore, $\|(A_{:,\mathcal{J}_{k-1}})^T z^{(k)}\|_2^2 = 0$ and then

$$\begin{aligned} \|A^T z^{(k)}\|_2^2 &= \sum_{j \in [n] \setminus \mathcal{J}_{k-1}} \frac{|A_{(j)}^T z^{(k)}|^2}{\|A_{(j)}\|_2^2} \|A_{(j)}\|_2^2 \\ &\leq \max_{j \in [n]} \left\{ \frac{|A_{(j)}^T z^{(k)}|^2}{\|A_{(j)}\|_2^2} \right\} (\|A\|_F^2 - \|A_{\mathcal{J}_{k-1}}\|_F^2). \end{aligned}$$

Thus,

$$\tilde{\varepsilon}_k = \rho_z \max_{j \in [n]} \left\{ \frac{|A_{(j)}^T z^{(k)}|^2}{\|A_{(j)}\|_2^2} \right\} \geq \rho_z \frac{\|A^T z^{(k)}\|_2^2}{\|A\|_F^2 - \|A_{:,\mathcal{J}_{k-1}}\|_F^2} \geq \rho_z \frac{\sigma_r^2(A)\|\tilde{e}^k\|_2^2}{\|A\|_F^2 - \|A_{:,\mathcal{J}_{k-1}}\|_F^2}. \tag{2.4}$$

Substituting (2.3) and (2.4) into (2.2) yields

$$\|\tilde{e}^{(k+1)}\|_2^2 \leq \|\tilde{e}^{(k)}\|_2^2 - \frac{\tilde{\varepsilon}_k \|A_{:,\mathcal{J}_k}\|_F^2}{\sigma_1^2(A_{:,\mathcal{J}_k})} \leq \left(1 - \rho_z \frac{\|A_{:,\mathcal{J}_k}\|_F^2}{\sigma_1^2(A_{:,\mathcal{J}_k})} \frac{\sigma_{\min}^2(A)}{\|A\|_F^2 - \|A_{:,\mathcal{J}_{k-1}}\|_F^2} \right) \|\tilde{e}^{(k)}\|_2^2.$$

From the fact $\frac{\|A_{:,\mathcal{J}k}\|_F^2}{\sigma_1^2(A_{:,\mathcal{J}k})} \geq 1$ and the definition $\tilde{\phi}_{\min} := \min_{j \in [n]} \|A_{(j)}\|_2^2$, the recursive expression (2.1) is derived. □

Remark 1 In the extended randomized multiple rows method, the expected decrease in mean squared error at the $(k+1)$-th iteration is

$$\mathbb{E}\|\tilde{P}_k \tilde{e}^{(k)}\|_2^2 = \frac{\|A^T z^{(k)}\|_2^4}{\|AA^T z^{(k)}\|_2^2},$$

which is obtained by taking the expectation over $\mathcal{J}_k$ for the first equality of (2.3). Here $\mathbb{E}$ denotes the expected value conditional on the first k iterations. For the greedy block extended Kaczmarz method, the error reduction is $\frac{(\tilde{\eta}_k^T A^T \tilde{e}^{(k)})^2}{\|A\tilde{\eta}_k\|_2^2}$. It is obvious that

$$\frac{(\tilde{\eta}_k^T A^T \tilde{e}^{(k)})^2}{\|A\tilde{\eta}_k\|_2^2} = \frac{\sum\limits_{j\in\mathcal{J}_k} (A^T z^{(k)})_j^4}{\sum\limits_{j\in\mathcal{J}_k} \|A_{(j)}\|_2^2 (A^T z^{(k)})_j^2} \geq \frac{\sum\limits_{j\in[n]} (A^T z^{(k)})_j^4}{\sum\limits_{j\in[n]} \|A_{(j)}\|_2^2 (A^T z^{(k)})_j^2} = \frac{\|A^T z^{(k)}\|_2^4}{\|AA^T z^{(k)}\|_2^2},$$

which indicates that the convergence rate of $\{z^{(k)}\}_{k=0}^{\infty}$ in the greedy block extended Kaczmarz method is larger than that of the extended randomized multiple rows method.

The convergence analysis of $\{x^{(k)}\}_{k=0}^{\infty}$ in the greedy block extended Kaczmarz method relies on the utilization of the following lemma.

Lemma 2 *([5]). Let c_1, c_2 be real numbers such that $c_1 \in [0, 1)$, $\quad c_2 \geq -1$, $\quad c_2 - c_1 = c_1 c_2$, then*

$$(r_1 + r_2)^2 \geq c_1 r_1^2 - c_2 r_2^2, \quad \forall\, r_1, r_2 \in \mathbb{R}.$$

By Theorem 1 and Lemma 2, the convergence property for the greedy block extended Kaczmarz method is constructed as follows.

Theorem 2 *Assume* $\operatorname{rank}(A) = r$. *The sequence $\{x^{(k)}\}_{k=0}^{\infty}$ with the initial guess $x^{(0)} = 0$ generated by the GBEK method converges to the least squares solution $x_{LS} = A^{\dagger} b$. Moreover, the solution error satisfies*

$$\|x^{(k+1)} - x_{LS}\|_2^2 \leq \max\{\alpha_x, \alpha_z\}^{k+1} \left(1 + (k+1)\beta\sigma_1^2(A)\right) \|x^{(0)} - x_{LS}\|_2^2,$$

where

$$\alpha_x := 1 - \rho_x c_1^2 \frac{\sigma_r^2(A)}{\|A\|_F^2 - \phi_{\min}}, \alpha_z := 1 - \rho_z \frac{\sigma_r^2(A)}{\|A\|_F^2 - \tilde{\phi}_{\min}}, \beta := \frac{\rho_x c_1 c_2}{\|A\|_F^2 - \phi_{\min}} + \frac{c_2 + 1}{\phi_{\min}},$$

with constants c_1, c_2 from Lemma 2, $\phi_{\min} = \min\limits_{i\in[m]} \|A^{(i)}\|_2^2$ and $\tilde{\phi}_{\min} = \min\limits_{j\in[n]} \|A_{(j)}\|_2^2$.

Proof Subtracting x_{LS} from both sides of step 9 in the Algorithm 1 leads to

$$x^{(k+1)} - x_{LS} = x^{(k)} - x_{LS} + \frac{\eta_k^T (b - z^{(k+1)} - Ax^{(k)})}{\|A^T \eta_k\|_2^2} A^T \eta_k.$$

For simplicity, let $e^{(k)} = x^{(k)} - x_{LS}$ and $P_k = \frac{A^T \eta_k \eta_k^T A}{\|A^T \eta_k\|_2^2}$. Then we have

$$e^{(k+1)} = e^{(k)} - P_k e^{(k)} - \frac{A^T \eta_k \eta_k^T \tilde{e}^{(k+1)}}{\|A^T \eta_k\|_2^2}.$$

Observe that $e^{(k)} - P_k e^{(k)}$ is perpendicular to $\frac{A^T \eta_k \eta_k^T \tilde{e}^{(k+1)}}{\|A^T \eta_k\|_2^2}$ and P_k is an orthogonal projection. By taking the norm of both sides of the above equality and using the Pythagorean Theorem, it yields

$$\|e^{(k+1)}\|_2^2 = \|e^{(k)}\|_2^2 - \|P_k e^{(k)}\|_2^2 + \frac{(\eta_k^T \tilde{e}^{(k+1)})^2}{\|A^T \eta_k\|_2^2}. \tag{2.5}$$

Since $e^{(0)} = x^{(0)} - x_{LS} = A^{\dagger} b \in R(A^T)$, $\frac{\eta_k^T (b - z^{(k+1)} - A x^{(k)})}{\|A^T \eta_k\|_2^2} A^T \eta_k \in R(A^T)$, it holds $e^{(k+1)} \in R(A^T)$ by induction. It follows that

$$\begin{aligned}
\|P_k e^{(k)}\|_2^2 - \frac{(\eta_k^T \tilde{e}^{(k+1)})^2}{\|A^T \eta_k\|_2^2} &= \frac{(\eta_k^T A e^{(k)})^2 - (\eta_k^T \tilde{e}^{(k+1)})^2}{\|A^T \eta_k\|_2^2} \\
&= \frac{\eta_k^T (A e^{(k)} + \tilde{e}^{(k+1)}) \eta_k^T (A e^{(k)} - \tilde{e}^{(k+1)})}{\|A^T \eta_k\|_2^2} \\
&\geq \frac{\eta_k^T (-A e^{(k)} + \tilde{e}^{(k+1)}) \|\eta_k\|_2^2}{\sigma_1^2 (A_{\mathcal{I}_k,:}) \|\eta_k\|_2^2} \\
&= \frac{\sum_{i \in \mathcal{I}_k} \left(-(A^{(i)})^T e^{(k)} - \tilde{e}_i^{(k+1)}\right) \left(-(A^{(i)})^T e^{(k)} + \tilde{e}_i^{(k+1)}\right)}{\sigma_1^2 (A_{\mathcal{I}_k,:})} \\
&= \frac{\|A_{\mathcal{I}_k,:} e^{(k)}\|_2^2 - \|\tilde{e}_{\mathcal{I}_k}^{(k+1)}\|_2^2}{\sigma_1^2 (A_{\mathcal{I}_k,:})}.
\end{aligned}$$

For the term $\|A_{\mathcal{I}_k,:} e^{(k)}\|_2^2$,

$$\begin{aligned}
\|A_{\mathcal{I}_k,:} e^{(k)}\|_2^2 &= \sum_{i \in \mathcal{I}_k} |(A^{(i)})^T e^{(k)}|^2 = \sum_{i \in \mathcal{I}_k} |b_{R(A)_i} - (A^{(i)})^T x^{(k)}|^2 \\
&\geq \sum_{i \in \mathcal{I}_k} \left(c_1 \left(b_i - z_i^{(k+1)} - (A^{(i)})^T x^{(k)} \right)^2 - c_2 \left(z_i^{(k+1)} - b_{R(A)_i^{\perp}} \right)^2 \right) \\
&\geq \sum_{i \in \mathcal{I}_k} \left(c_1 \varepsilon_k \|A^{(i)}\|_2^2 - c_2 |\tilde{e}_i^{(k+1)}|^2 \right) \\
&= c_1 \varepsilon_k \|A_{\mathcal{I}_k,:}\|_F^2 - c_2 \|\tilde{e}_{\mathcal{I}_k}^{(k+1)}\|_2^2,
\end{aligned} \tag{2.6}$$

where the first inequality holds because of Lemma 2. Therefore equation (2.5) becomes

$$\begin{aligned}
\|e^{(k+1)}\|_2^2 &\leq \|e^{(k)}\|_2^2 - \frac{c_1 \varepsilon_k \|A_{\mathcal{I}_k,:}\|_F^2}{\sigma_1^2 (A_{\mathcal{I}_k,:})} + \frac{c_2 + 1}{\sigma_1^2 (A_{\mathcal{I}_k,:})} \|\tilde{e}_{\mathcal{I}_k}^{(k+1)}\|_2^2 \\
&\leq \|e^{(k)}\|_2^2 - c_1 \varepsilon_k + \frac{c_2 + 1}{\min_{i \in [m]} \|A^{(i)}\|_2^2} \|\tilde{e}_{\mathcal{I}_k}^{(k+1)}\|_2^2,
\end{aligned} \tag{2.7}$$

where the second inequality holds because $\frac{\|A_{\mathcal{I}_k,:}\|_F^2}{\sigma_1^2(A_{\mathcal{I}_k,:})} \geq 1$ and $\sigma_1^2(A_{\mathcal{I}_k,:}) \geq \min_{i\in[m]} \|A^{(i)}\|_2^2$. For the lower bound of ε_k, it holds that

$$\begin{aligned}
\varepsilon_k &= \rho_x \max_{i\in[m]} \left\{ \frac{|b(i) - z^{(k+1)}(i) - (A^{(i)})^T x^{(k)}|^2}{\|A^{(i)}\|_2^2} \right\} \geq \rho_x \frac{\|b - z^{(k+1)} - Ax^{(k)}\|_2^2}{\|A\|_F^2 - \min_{i\in[m]} \|A^{(i)}\|_2^2} \\
&\geq \rho_x \frac{c_1\|b_{R(A)} - Ax^{(k)}\|_2^2 - c_2\|z^{(k+1)} - b_{R(A)^\perp}\|_2^2}{\|A\|_F^2 - \min_{i\in[m]} \|A^{(i)}\|_2^2} \\
&\geq \rho_x \frac{c_1\sigma_r^2(A)\|e^{(k)}\|_2^2 - c_2\|\tilde{e}^{(k+1)}\|_2^2}{\|A\|_F^2 - \min_{i\in[m]} \|A^{(i)}\|_2^2}.
\end{aligned}$$

Let $\phi_{\min} = \min_{i\in[m]} \|A^{(i)}\|_2^2$. With the lower bound of ε_k, the inequality (2.7) is reformulated as

$$\begin{aligned}
\|e^{(k+1)}\|_2^2 \leq &\left(1 - \rho_x c_1^2 \frac{\sigma_r^2(A)}{\|A\|_F^2 - \phi_{\min}}\right) \|e_k\|_2^2 \\
&+ \left(\frac{\rho_x c_1 c_2}{\|A\|_F^2 - \phi_{\min}} + \frac{c_2 + 1}{\phi_{\min}}\right) \|\tilde{e}^{(k+1)}\|_2^2.
\end{aligned} \tag{2.8}$$

According to Theorem 1,

$$\|\tilde{e}^{(k+1)}\|_2^2 \leq \left(1 - \rho_z \frac{\sigma_r^2(A)}{\|A\|_F^2 - \tilde{\phi}_{\min}}\right) \|\tilde{e}^{(k)}\|_2^2,$$

where $\tilde{\phi}_{\min} = \min_{j\in[n]} \|A_{(j)}\|_2^2$. For simplicity, let

$$\alpha_x := 1 - \rho_x c_1^2 \frac{\sigma_r^2(A)}{\|A\|_F^2 - \phi_{\min}}, \alpha_z := 1 - \rho_z \frac{\sigma_r^2(A)}{\|A\|_F^2 - \tilde{\phi}_{\min}}, \beta := \frac{\rho_x c_1 c_2}{\|A\|_F^2 - \phi_{\min}} + \frac{c_2 + 1}{\phi_{\min}},$$

then the inequality (2.8) is rewritten as

$$\begin{aligned}
\|e^{(k+1)}\|_2^2 &\leq \alpha_x \|e^{(k)}\|_2^2 + \beta \|\tilde{e}^{(k+1)}\|_2^2 \\
&\leq \alpha_x^{k+1} \|e^{(0)}\|_2^2 + \beta \sum_{l=0}^{k} \alpha_x^l \|\tilde{e}^{(k+1-l)}\|_2^2 \\
&\leq \alpha_x^{k+1} \|e^{(0)}\|_2^2 + \beta \sum_{l=0}^{k} \alpha_x^l \alpha_z^{k+1-l} \|\tilde{e}^{(0)}\|_2^2 \\
&\leq \max\{\alpha_x, \alpha_z\}^{k+1} \|e^{(0)}\|_2^2 + (k+1)\beta \max\{\alpha_x, \alpha_z\}^{k+1} \|b_{R(A)}\|_2^2
\end{aligned} \tag{2.9}$$

$$\leq \max\{\alpha_x, \alpha_z\}^{k+1} \left(1 + (k+1)\beta\sigma_1^2(A)\right) \|e^{(0)}\|_2^2, \tag{2.10}$$

where the third inequality holds because of $\tilde{e}^{(0)} = z^{(0)} - b_{R(A)^\perp} = b_{R(A)}$, and the last inequality holds because of $x^{(0)} = 0$ and $\|b_{R(A)}\|_2^2 \leq \sigma_1^2(A)\|x_{LS}\|_2^2$. This completes the proof. □

3 Numerical Experiments

In this section, the numerical examples are presented to show the efficiency of the greedy block extended Kaczmarz (GBEK) method compared with the randomized double block Kaczmarz (RDBK) method, the randomized extended average block Kaczmarz (REABK) method and the extended randomized multiple rows (ERMR) method.

The inconsistent system is $Ax + \varepsilon = b$, where ε is a noise vector whose entries are drawn from a normal distribution and satisfies $\|\varepsilon\|_2 = 0.01 \times \|Ax\|_2$. The number of iteration steps (denoted as "IT") and the computational time in seconds (denoted as "CPU") are used for evaluation. The row blocks $\{\mathcal{I}_i\}_{i=1}^s$ and the column blocks $\left\{\mathcal{J}_j\right\}_{j=1}^t$ of the RDBK, REABK and ERMR methods are partitioned as follows:

$$\begin{aligned}
\mathcal{I}_i &= \{(i-1)\tau_r + 1, (i-1)\tau_r + 2, \ldots, i\tau_r\}, \quad i = 1, 2, \ldots, s-1, \\
\mathcal{I}_s &= \{(s-1)\tau_r + 1, (s-1)\tau_r + 2, \ldots, m\}, \quad |\mathcal{I}_s| \leq \tau_r,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{J}_j &= \{(j-1)\tau_c + 1, (j-1)\tau_c + 2, \ldots, j\tau_c\}, \quad j = 1, 2, \ldots, t-1, \\
\mathcal{J}_t &= \{(t-1)\tau_c + 1, (t-1)\tau_c + 2, \ldots, n\}, \quad |\mathcal{J}_t| \leq \tau_c,
\end{aligned}$$

where τ_r and τ_c are block sizes for the row and column partitions respectively. To ensure a fair comparison, it is necessary to use the same block size for all four methods. This is achieved by initially applying the GBEK method to get the average sizes of the row and column blocks, then utilizing these sizes to partition the rows and columns for the RDBK, REABK and ERMR methods.

All the methods are started from the initial vectors $x^{(0)} = 0$ and $z^{(0)} = b$ and stopped if the relative solution error (RSE) satisfies

$$\text{RSE} = \frac{\|x^{(k)} - x_{LS}\|_2^2}{\|x_{LS}\|_2^2} \leq 10^{-6},$$

or the number of iteration steps exceeds 50000. To compare the difference in computational time between the proposed method and other methods, define the following speed-ups:

$$\text{speed-up}_1 = \frac{\text{CPU of RDBK}}{\text{CPU of GBEK}}, \quad \text{speed-up}_2 = \frac{\text{CPU of REABK}}{\text{CPU of GBEK}},$$

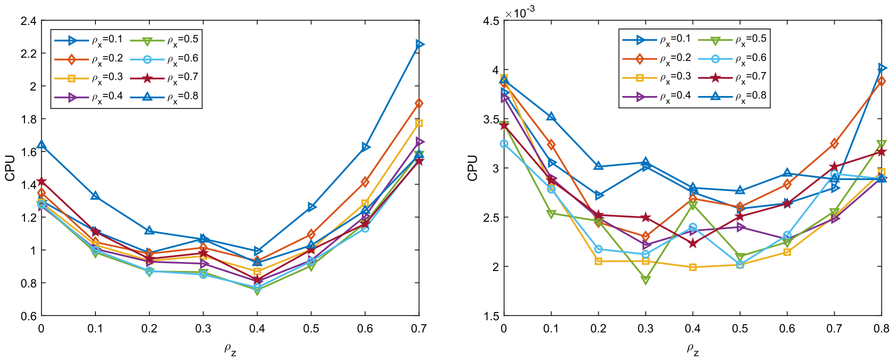

Fig. 1 Curves of the computing time versus ρ_z with fixed ρ_x (left: $A \in \mathbb{R}^{1000\times 100}$, $\mathrm{cond}(A) = 50$, right: $A = \mathrm{rel5}$, $\mathrm{cond}(A) = \mathrm{Inf}$)

$$\text{speed-up}_3 = \frac{\text{CPU of ERMR}}{\text{CPU of GBEK}}.$$

Example 1 Apply the GBEK method with different parameters ρ_x and ρ_z to solve the problem, where A is either a random Gaussian matrix or a sparse matrix from [8].

The influence of parameters ρ_x and ρ_z on the efficiency of the GBEK method is firstly explored in Example 1.

In Fig. 1, the curves of the computational time versus ρ_z with fixed ρ_x of the GBEK method for two different matrices are presented respectively. For $A \in \mathbb{R}^{1000\times 100}$ and $\mathrm{cond}(A) = 50$, it is observed that the computing time first decreases and then increases when the value of ρ_x is fixed and value of ρ_x is increasing. Similar phenomenon exists in $A = \mathrm{rel5}$. For the rest of the numerical examples, we set the parameters $\rho_x = \rho_z = 0.5$ in the GBEK method.

Example 2 The coefficient matrix A is an overdetermined random Gaussian matrix.

In Table 1, the number of iteration steps, computational time and speed-ups for the RDBK, REABK, ERMR and GBEK methods for solving Example 2 are presented respectively.

From Table 1, it is obvious that the GBEK method significantly reduces the number of iteration steps and computing time, compared with the RDBK, REABK, ERMR methods. The GBEK method demonstrates a noticeable performance against the ERMR method, with a maximum value of speed-up$_3$ reaching 12.9021. Due to the fact that the GBEK method and the ERMR method employ the same iterative format, the possible reason of the superior performance of the GBEK method is the use of greedy block criterion.

Example 3 The coefficient matrix A is an underdetermined random Gaussian matrix.

The numerical results of Example 3 are listed in Table 2. When the coefficient matrix A is underdetermined, the GBEK method outperforms the RDBK, REABK and ERMR methods in terms of both iteration counts and computing time. The GBEK

Table 1 Numerical results for Example 2

$m \times n$		5000×500	6000×500	7000×500	8000×500	9000×500	10000×500
RDBK	IT	249	174	505	249	409	142
	CPU	1.3913	1.0503	3.0794	1.8150	3.1954	1.2546
REABK	IT	1865	391	817	1349	963	338
	CPU	1.9426	0.5579	1.0848	2.0793	1.6602	0.7687
ERMR	IT	182	182	336	460	352	188
	CPU	1.5045	1.8212	4.0031	6.5170	5.5494	3.4387
GBEK	IT	43	40	38	35	33	32
	CPU	0.3568	0.3945	0.5000	0.5051	0.5205	0.5778
speed-up$_1$		3.8990	2.6624	6.8451	3.5932	6.1390	2.1713
speed-up$_2$		5.4438	1.4142	2.4113	4.1166	3.1895	1.3304
speed-up$_3$		4.2163	4.6166	8.8982	12.9021	10.6614	5.9510

Table 2 Numerical results for Example 3

$m \times n$		500×5000	500×6000	500×7000	500×8000	500×9000	500×10000
RDBK	IT	210	209	173	169	127	245
	CPU	0.8866	1.4028	1.3351	1.4177	1.3707	2.0258
REABK	IT	254	503	374	253	263	397
	CPU	0.7520	1.4202	1.2908	0.9645	1.2887	2.3532
ERMR	IT	216	196	169	203	151	164
	CPU	2.2721	2.5736	2.5964	3.6264	3.0398	3.8439
GBEK	IT	45	41	39	36	34	35
	CPU	0.4883	0.5501	0.6027	0.3501	0.6934	0.8565
speed-up$_1$		1.8157	2.5498	2.2153	2.1683	1.9768	2.3653
speed-up$_2$		1.5399	2.5815	2.1418	1.4750	1.8586	2.7476
speed-up$_3$		4.6528	4.6781	4.3083	5.5463	4.3840	4.4881

Table 3 Numerical results for Example 4

name		rel5	abtaha1	relat6	df2177	lp_qap8
$m \times n$		340 × 35	14596 × 209	2340 × 157	630 × 10358	912 × 1632
density		5.51%	1.68%	2.21%	0.34%	0.49%
cond(A)		Inf	12.2283	Inf	2.0066	2.32×10^{17}
rank(A)		24	209	137	630	742
RDBK	IT	170	3921	419	158	374
	CPU	0.1200	16.5563	0.5023	0.4192	0.3356
REABK	IT	792	12435	6976	405	2665
	CPU	0.0121	2.7548	0.2851	0.2597	0.1423
ERMR	IT	325	5779	1146	333	1162
	CPU	0.0061	2.6146	0.0734	0.0985	0.0772
GBEK	IT	92	793	319	45	912
	CPU	0.0022	0.4685	0.0280	0.0196	0.0103
speed-up$_1$		55.3191	35.3376	17.9583	21.4212	32.4548
speed-up$_2$		5.5783	5.8798	10.1924	13.2687	13.7592
speed-up$_3$		2.8219	5.5806	2.6243	5.0345	7.4629

method exhibits the lowest iteration counts and the shortest computational time to achieve the desired accuracy.

Example 4 The matrix A is taken from the SuiteSparse Matrix Collection [8].

For solving Example 4, the numbers of iteration steps, the computational time and the speed-ups for the RDBK, REABK, ERMR and GBEK methods are provided in Table 3. All the coefficient matrices are sparse and rank-deficient, with different matrix sizes, densities, and condition numbers. Here the density of A is defined as

$$\text{density} = \frac{\text{number of nonzeros of } A}{mn},$$

which accurately describes the sparsity of A.

From Table 3, it is seen that the GBEK method outperforms the RDBK, REABK and ERMR methods in terms of the number of iteration steps and computational time. Furthermore, in Table 3, the maximum values of speed-up$_1$ and speed-up$_2$ are 55.3191 and 13.7952 respectively, which further confirms the superiority of the GBEK method for solving large sparse least squares problems.

The curves of the relative solution error versus the iteration counts for the RDBK, REABK, ERMR and GBEK methods for different matrices are showed in Fig. 2. It is clear that the relative solution error of the GBEK method decreases the fastest as the number of iteration steps increases for these three examples.

Example 5 Consider solving the X-ray computed tomography problem in AIR Tools II [14]. The size of the matrix A is set to be 15300×3600.

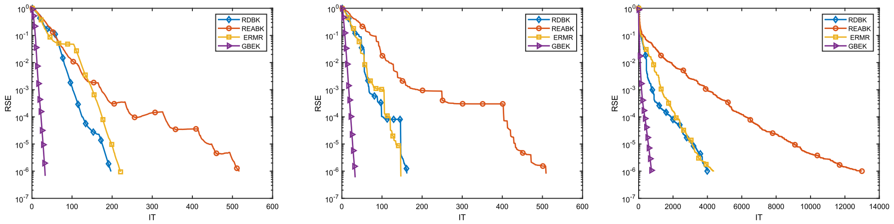

Fig. 2 Convergence curves of the RDBK, REABK, ERMR and GBEK methods for different matrices (left: $A \in \mathbb{R}^{10000\times500}$, middle: $A \in \mathbb{R}^{500\times10000}$, right: $A = \text{abtaha1}$)

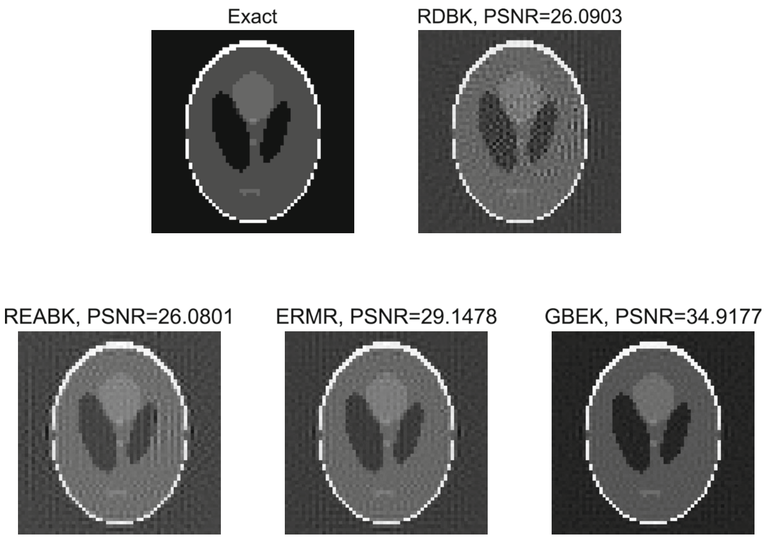

Fig. 3 Numerical results for Example 5

In Example 5, the effectiveness of the RDBK, REABK, ERMR and GBEK methods is evaluated by the Peak Signal-to-Noise Ratio (PSNR), which is a widely used metric in image processing to measure the similarity between two images. The higher PSNR value indicates the better image quality. All methods were run with the same number of iteration steps.

The original image and the approximate images recovered by the four methods are given in Fig. 3. It is obvious that the image reconstructed by the GBEK method is the best and attains the highest PSNR value of 34.9177.

4 Conclusions

A greedy block extended Kaczmarz method is proposed for solving least squares problems. Theoretical analysis is established and a linear convergence rate is derived. Numerical experiments show the proposed method exhibits a better performance than randomized block extended Kaczmarz methods in terms of both the number of iteration steps and computational time.

Declarations

Conflict of interest The authors declare that there is no Conflict of interest.

References

1. Arioli, M., Duff, I.S.: Preconditioning linear least-squares problems by identifying a basis matrix. SIAM J. Sci. Comput. **37**(5), S544–S561 (2015)
2. Bai, Z.-Z., Wen-Ting, W.: On greedy randomized Kaczmarz method for solving large sparse linear systems. SIAM J. Sci. Comput. **40**(1), A592–A606 (2018)
3. Bai, Z.-Z., Wen-Ting, W.: On partially randomized extended Kaczmarz method for solving large sparse overdetermined inconsistent linear systems. Linear Algebra Appl. **578**, 225–250 (2019)
4. Bai, Z.-Z., Wen-Ting, W.: On greedy randomized augmented Kaczmarz method for solving large sparse inconsistent linear systems. SIAM J. Sci. Comput. **43**(6), A3892–A3911 (2021)
5. Bao, W., Lv, Z., Zhang, F., Li, W.: A class of residual-based extended Kaczmarz methods for solving inconsistent linear systems. J. Comput. Appl. Math. **416**, 114529 (2022)
6. Björck, Åke: *Numerical methods for least squares problems*. SIAM, (1996)
7. Chen, J.-Q., Huang, Z.-D.: On a fast deterministic block Kaczmarz method for solving large-scale linear systems. Numer. Algorithms **89**(3), 1007–1029 (2022)
8. Davis, T.A., Hu, Y.: The University of Florida sparse matrix collection. ACM Trans. Math. Softw. (TOMS) **38**(1), 1–25 (2011)
9. Kui, D.: Tight upper bounds for the convergence of the randomized extended Kaczmarz and Gauss-Seidel algorithms. Numer. Linear Algebra Appl. **26**(3), e2233 (2019)
10. Kui, D., Si, W.-T., Sun, X.-H.: Randomized extended average block Kaczmarz for solving least squares. SIAM J. Sci. Comput. **42**(6), A3541–A3559 (2020)
11. Gower, R.M., Molitor, D., Moorman, J., Needell, D.: On adaptive sketch-and-project for solving linear systems. SIAM J. Matrix Anal. Appl. **42**(2), 954–989 (2021)
12. Haddock, J., Ma, A.: Greed works: An improved analysis of sampling Kaczmarz-Motzkin. SIAM J. Math. Data Sci. **3**(1), 342–368 (2021)
13. Haddock, J., Needell, D.: On Motzkin's method for inconsistent linear systems. BIT Numer. Math. **59**, 387–401 (2019)
14. Hansen, P.C., Jørgensen, J.S.: AIR Tools II: algebraic iterative reconstruction methods, improved implementation. Numer. Algorithms **79**(1), 107–137 (2018)
15. Herman, G T.: *Fundamentals of computerized tomography: image reconstruction from projections*. Springer Science & Business Media, (2009)
16. Hoyos-Idrobo, A., Weiss, P., Massire, A., Amadon, A., Boulant, N.: On variant strategies to solve the magnitude least squares optimization problem in parallel transmission pulse design and under strict SAR and power constraints. IEEE Trans. Med. Imaging **33**(3), 739–748 (2013)
17. Ma, A., Needell, D., Ramdas, A.: Convergence properties of the randomized extended Gauss-Seidel and Kaczmarz methods. SIAM J. Matrix Anal. Appl. **36**(4), 1590–1604 (2015)
18. Mccormick, S.F.: The method of Kaczmarz and row orthogonalization for solving linear equations and least squares problems in hilbert space. Indiana Univ. Math. J. **26**(6), 1137–1150 (1977)
19. Necoara, I.: Faster randomized block Kaczmarz algorithms. SIAM J. Matrix Anal. Appl. **40**(4), 1425–1452 (2019)
20. Needell, D., Zhao, R., Zouzias, A.: Randomized block Kaczmarz method with projection for solving least squares. Linear Algebra Appl. **484**, 322–343 (2015)
21. Niu, Y.-Q., Zheng, B.: A greedy block Kaczmarz algorithm for solving large-scale linear systems. Appl. Math. Lett. **104**, 106294 (2020)
22. Strohmer, T., Vershynin, R.: A randomized Kaczmarz algorithm with exponential convergence. J. Fourier Anal. Appl. **15**(2), 262 (2009)

23. Wu, N-C, Liu, C, Wang, Y, Zuo, Q: On the extended randomized multiple row method for solving linear least-squares problems. arXiv preprint arXiv:2210.03478, (2022)
24. Wen-Ting, W.: On two-subspace randomized extended Kaczmarz method for solving large linear least-squares problems. Numer. Algorithms **89**(1), 1–31 (2022)
25. Yang, X.: A geometric probability randomized Kaczmarz method for large scale linear systems. Appl. Numer. Math. **164**, 139–160 (2021)
26. Zouzias, A., Freris, N.M.: Randomized extended Kaczmarz for solving least squares. SIAM J. Matrix Anal. Appl. **34**(2), 773–793 (2013)