East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.160721.160122

Vol. **12**, No. 2, pp. 406-420
May 2022

A Fast Block Coordinate Descent Method for Solving Linear Least-Squares Problems

Jia-Qi Chen and Zheng-Da Huang*

School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, Zhejiang, P.R. China.

Received 16 July 2021; Accepted (in revised version) 16 January 2022.

Abstract. A fast block coordinate descent method for solving linear least-squares problems is proposed. The method is based on a greedy criterion of the column selection used at each iteration. It is proven that if the coefficient matrix of the corresponding system has full column rank, the method converges to the unique solution of the linear least-squares problem. Numerical experiments show the advantage of this approach over similar methods in terms of CPU time and computational cost, does not matter whether the coefficient matrix is of full column rank or not.

AMS subject classifications: 65F10, 65F20, 65K05, 90C25, 15A06

Key words: Linear least-squares problem, Kaczmarz method, coordinate descent method, greedy blocks, convergence.

1. Introduction

We consider iteration methods for approximation of the least-norm least-squares solution $x_\star = A^\dagger b$ of the system of linear equations

$$Ax = b, \tag{1.1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $A^\dagger$ refers to the Moore-Penrose pseudoinverse of the matrix A.

For consistent linear systems (1.1), the Kaczmarz method [14] is one of the typical deterministic row-action solvers, and the randomized Kaczmarz (RK) method [23] is an earlier randomized version of the Kaczmarz method. For the convergence analysis of these methods we refer to [1, 13, 23].

If (1.1) is an inconsistent system, the RK method may not converge to $x_\star$ [20]. To fix the problem, Zouzias and Freris [26] proposed a randomized extended Kaczmarz (REK) method. Using different strategies for the column choice in the REK method or a consistent

*Corresponding author. *Email addresses:* zdhuang@zju.edu.cn (Z.-D. Huang), 11835035@zju.edu.cn (J.-Q. Chen)

augmented linear system, Bai and Wu [6, 7] constructed a partially randomized extended Kaczmarz (PREK) method and greedy randomized augmented Kaczmarz (GRAK) method.

In order to accelerate the convergence of the REK method, Needell *et al*. [21] introduced a block variant of the REK method, named as the randomized double block Kaczmarz (RDBK) method. In this approach, the partitions of both row and column indices of the coefficient matrix A have to be predetermined. Du *et al*. [11] proposed a randomized extended average block Kaczmarz (REABK) method. This method can be viewed as another block variant of the REK method. Wu [25] constructed a two-subspace randomized extended Kaczmarz (TREK) method, which does not need to pre-determine any partition of the rows or the columns. It is proven that for properly chosen initial values, the sequences obtained by the three above mentioned methods converge to $x_\star$ in expectation, does not matter whether the linear system (1.1) is overdetermined ($m \geq n$), underdetermined ($m < n$), consistent, inconsistent, full-rank, or rank-deficient [8, 11, 21, 25].

The coordinate descent-like method, as one kind of well-known method in solving the optimization problems, enjoys a long history —- cf. [17–19, 22, 24] and references therein. For solving consistent and inconsistent linear systems (1.1), Leventhal and Lewis [15] proposed a randomized coordinate descent (RCD) method, which at each iteration chooses a column of A with the probability proportional to the square of its Euclidean norm. They also estimated the convergence rate in terms of a natural condition measure in linear algebra for coordinate descent methods. The RCD method can be regarded as a column-action method. Modifying the probability criterion in [3, 4], Bai and Wu [5] proposed a greedy randomized coordinate descent (GRCD) method with a faster convergence in expectation, both in theory and computations.

A block version of the RCD method for solving the linear system (1.1) is the randomized block coordinate descent (RBCD) method proposed by Needell *et al*. [21]. At each step, the RBCD method selects randomly a column index set from a predetermined partition of $\{1, 2, \ldots, n\}$. It should be mentioned that only a partition of column indices of the coefficient matrix A needs to be predetermined in the RBCD method. By using different strategies to select the block index set and adopting the same iterative format of the RBCD method, Liu *et al*. [16] proposed a maximal residual block Gauss-Seidel (MRBGS) method, which is a deterministic coordinate descent method and does not need to pre-determine any partition of the column indices of the matrix A.

For overdetermined linear systems with a full-column rank coefficient matrix, Dumitrescu [12] showed that, in average (over the random draw of columns and rows), the RCD method requires less iterations and less arithmetic operations to reach the stopping criterion than the REK method. In other words, in this case the coordinate descent-like methods have advantages. This motivates us to try to construct a more efficient block version of the coordinate descent method.

In this paper, we use the greedy rule in [5] for the column block selection and develop a fast block coordinate descent (FBCD) method based on an iterative format different from the one in the RBCD and MRBGS methods. We prove that if the coefficient matrix A has full column rank, this method converges to the unique least-norm least-squares solution of the linear system (1.1) and derive the corresponding error estimate. Testing the method