Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2020-0095

Vol. **14**, No. 3, pp. 714-737
August 2021

Variant of Greedy Randomized Gauss-Seidel Method for Ridge Regression

Li-Xiao Duan and Guo-Feng Zhang*

School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, P.R. China

Received 1 July 2020; Accepted (in revised version) 29 November 2020

Abstract. The variants of randomized Kaczmarz and randomized Gauss-Seidel algorithms are two effective stochastic iterative methods for solving ridge regression problems. For solving ordinary least squares regression problems, the greedy randomized Gauss-Seidel (GRGS) algorithm always performs better than the randomized Gauss-Seidel algorithm (RGS) when the system is overdetermined. In this paper, inspired by the greedy modification technique of the GRGS algorithm, we extend the variant of the randomized Gauss-Seidel algorithm, obtaining a variant of greedy randomized Gauss-Seidel (VGRGS) algorithm for solving ridge regression problems. In addition, we propose a relaxed VGRGS algorithm and the corresponding convergence theorem is established. Numerical experiments show that our algorithms outperform the VRK-type and the VRGS algorithms when $m > n$.

AMS subject classifications: 65F10, 62J07, 68W20

Key words: Randomized algorithms, ridge regression, greedy randomized Gauss-Seidel, randomized Kaczmarz, iterative method.

1. Introduction

The Kaczmarz class algorithms [10, 14, 22, 23, 27, 30] and the Gauss-Seidel class algorithms [10, 15, 18] are usually effective methods to solve systems of linear equations $X\beta = y$, also sometimes called ordinary least squares (OLS) regression, where X is a complex $m \times n$ matrix, $y \in \mathbb{C}^m$ is a given m-dimensional complex vector. The Kaczmarz class and the Gauss-Seidel class algorithms select only one row or column of the matrix X as their workspace in each iteration. When $X\beta = y$ is ill-posed, Tikhonov-regularization (ridge regression) [28, 29] is the most commonly used method of regularization. The ridge regression possesses very wide range of applications, such as

*Corresponding author. *Email addresses:* duanlx17@lzu.edu.cn (L.-X. Duan), gf_zhang@lzu.edu.cn (G.-F. Zhang)

machine learning [16, 26], electricity networks [19, 20], statistical analysis [1, 21] and so on [9, 11], which corresponds to solving the convex optimization problem

$$\min_{\beta} \|y - X\beta\|_2^2 + \lambda\|\beta\|_2^2, \tag{1.1}$$

where λ is a given positive parameter. The solution of (1.1) is also equivalent to solving two dual linear systems. This means that we can directly apply the randomized Kaczmarz and randomized Gauss-Seidel algorithms to these two linear systems. Although their performance is poor, they are still instructive due to their simple iterative format. In [13, 33], Ivanov and Zhdanov proposed the augment projection method or IZ method for solving (1.1), which is a typical representative for the Kaczmarz-based algorithm.

Recently, Hefny *et al.* [12] proposed the variants of randomized Kaczmarz and randomized Gauss-Seidel algorithms to solve the ridge regression problems. In the absence of ambiguity, we call the VRK algorithm and the VRGS algorithm, respectively. They proved that the VRK algorithm is preferred in the underdetermined case ($m < n$) while the VRGS algorithm takes the lead in the overdetermined case ($m > n$).

For the OLS regression problems, Bai and Wu [3] proposed the greedy randomized Kaczmarz (GRK) algorithm by introducing an effective probability criterion for selecting the working rows from the coefficient matrix. Their theoretical results and numerical experiments show that the GRK algorithm is superior to the RK algorithm in CPU time and iteration steps. Based on the greedy strategy in [3], Liu and Gu [17] constructed a variant of greedy randomized Kaczmarz (VGRK) algorithm to solve the ridge regression problem. They also considered the VGRK method with relaxation parameter (VGRKRP). It can be showed from the convergence theorem and experimental results that the VRGK and the VGRKRP algorithms always perform better than the IZ0 and the VRGS algorithms for the underdetermined case.

In this paper, we propose a variant of greedy randomized Gauss-Seidel (VGRGS) method and its relaxed version. It can be viewed as an extension or variant of the greedy randomized Gauss-Seidel method introduced by Bai and Wu in [2] for solving the OLS problems to the ridge regression problems. As analyzed in [2], the greedy strategy of selecting working columns in the GRGS method can provide those more efficient column indicators. It was a natural idea to apply this approach to the VRGS algorithm to modify the way in which the column index was selected in proportion to the spectral norm of the column. We proved that the VGRGS algorithm and its relaxed version i.e. VGRGSRP algorithm converge faster than the VRK and VGRK algorithms in both theory and experiments for the overdetermined case ($m > n$).

The organization of this paper is as follows. In Section 2, we recap the five main existing algorithms mentioned in the introduction i.e. VRK, VRGS, IZ, VGRK and VGRKRP algorithms. In Section 3, we proposed the new algorithms i.e. VGRGS and VGRGSRP algorithms and the convergence theorem are given. Numerical results are reported in Section 4. The paper is ended with a brief conclusion in Section 5.