Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 319–326

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Single Acceleration Methods of the Kaczmarz Algorithm Regularized Modifications

Yu. V. Sidorov *

Department of Higher Mathematics and Applied Computer Science, Samara State Technical University, Molodogvardeyskaya street, building №244,Samara,443100, Russia

Abstract

In this paper suggested methods for acceleration Kaczmarz algorithm regularized modifications to solve the standard regularization problem of A.N. Tikhonov. As shown in numerical experiments, for certain classes of problems, such methods allow reducing both the number of iterations and the time for finding solutions. For the two-dimensional problem of seismic tomography proposed greedy forms of Kaczmarz algorithm regularized modifications can reduce the number of iterations up to 28 times.

Keywords: *iterative methods, projection algorithms, Tikhonov's regularization, Kaczmarz algorithm, row-oriented form of the regularized Kaczmarz's algorithm, column - oriented form of the regularized Kaczmarz's algorithm;*

1. Introduction

Many problems in science and technology are often reduced to solving arbitrary systems of linear algebraic equations (SLAE) of the form

$$Ax = f,\ A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^{n},\ f \in \mathbb{R}^{m}. \quad (1)$$

* Corresponding author. Tel.: .:+ 7-927-262-0110.
*E-mail address:*linuxboy2007@gmail.com

10.1016/j.procs.2019.06.046

By the solution of system (1), in the following, we mean the pseudo - solution $x_* = A^+ f$, where A^+ - Moore – Penrose pseudoinverse matrix[1].

In the case of ill-conditioned of the matrix A and (or) in the presence of noise in the vector of the right side f , the computation of stable solutions of SLAE (1) requires the use of regularization methods. Similar problems arise, for example, when SLAE (1) is obtained as a result of discretization of the Fredholm integral equation of the first kind with a smooth kernel, in regression analysis in the presence of multicollinearity and in the theory of machine learning - “retraining” of regression models. In practice, the most common method for finding stable solutions of SLAE (1) is the regularization method of A. N. Tikhonov[2].

The standard regularization problem of A.N. Tikhonov can be written as follows:

$$\min_{x\in\mathbb{R}^n}\left\{\|Ax-f\|^2+\alpha\|x\|^2\right\}. \tag{2}$$

Here α - is a regularization parameter, $\alpha > 0$, $\|\bullet\| = \|\bullet\|_2$ denotes the Euclidean vector norm.

When the dimension of the matrix A is large and (or) matrix is sparse, the problem (2) is solved using iterative methods based on the Euler equations (regularized normal equations)

$$\left(A^T A+\alpha I_n\right)x = A^T f. \tag{3}$$

The main drawback of methods based on normal equations is that the spectral condition number of the problem (3) is equal to the square of the spectral condition number of the SLAE matrix (1).

An augmented regularized normal system of equations (ARNSE) is proposed[3] for elimination this drawback and solving the problem (2)

$$\begin{pmatrix}\omega I_m & A\\ A^T & -\omega I_n\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}f\\ 0\end{pmatrix}\Leftrightarrow B_\omega\theta=q, \tag{4}$$

where $B_\omega=\begin{pmatrix}b_1^T\\ \vdots\\ b_{m+n}^T\end{pmatrix}\in\mathbb{R}^{(m+n)\times(m+n)}$; $\theta=\begin{pmatrix}x\\ y\end{pmatrix}\in\mathbb{R}^{(m+n)}$; $q=\begin{pmatrix}f\\ 0\end{pmatrix}=\left(q_1, \;\ldots,\; q_{m+n}\right)^T\in\mathbb{R}^{(m+n)}$,

$q_i = f_i$, $i=1,\ldots,m$, $q_i=0$, $i=m+1,\ldots,m+n$; $\omega=\sqrt{\alpha}$; I_m, I_n - are identities matrices of appropriate sizes.

One of the main advantages of ARNSE is the significantly smaller spectral condition number of the matrix B_ω compared with the spectral condition number of matrix A of system (1)[3].

When $\alpha > 0$, the matrix B_ω is a nondegenerate square matrix of full rank, and a system (4) is consistent and has one unique solution $\theta_* = \left(x_*^T \quad y_*^T\right)$, where x_* - the solution to the problem (2), $y_* = \omega^{-1}\left(f - Ax_*\right)$[3]. The solution of the problem (4) can be calculated both direct and iterative methods.

In 4,5,6, column and row oriented forms of the regularized Kaczmarz algorithm are proposed to solve problem (2), based on applying Kaczmarz’s algorithm7 to the system of augmented regularized normal equations (4). The considered forms of the algorithm are effective in solving problems of regularization of large and extra-large dimensions. There is a parallel implementation for multicore (multiprocessor) systems for the column - oriented form of a regularized algorithm8.

It is known that the rules for choosing the projection sequence for Kaczmarz -type algorithms have a very strong effect on the rate of convergence.

The ways to reduce the number of iterations of regularized forms of the Kaczmarz's algorithm while calculating the solution of the problem (2), using special rules for constructing a projection sequence are discussed in the considered article.

2. Regularized Forms of the Kaczmarz's algorithm

We write the system (4) in the following form:

$$\omega y + Ax = f, \tag{5}$$

$$A^T y - \omega x = 0. \tag{6}$$

If we use the equation (5) as the condition of matching and apply the Kaczmarz's algorithm to equation (6), we obtain the column-oriented form of the algorithm 5:

$$\begin{aligned} y_{k+1} &= y_k - \rho_k a_j, \\ x_{k+1} &= x_k + \rho_k e_j. \end{aligned}, \ \rho_k = \frac{a_j^T y_k - \alpha e_j^T x_k}{\left\| a_j \right\|^2 + \alpha}. \tag{7}$$

Here $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \in \mathbb{R}^{m\times n}$, $E_n = (e_1, \ldots, e_n)$, $k = 0,1,\ldots$, $j = j(k) = (k \bmod n) + 1$, $\{j(k)\}_{k=0}^{\infty}$ is periodic sequence of the form $1,2,\ldots,n,1,2,\ldots,n,\ldots$. The initial conditions x_0 and y_0 satisfy $y_0 = \omega^{-1}(f - Ax_0)$.

The column-oriented regularized form (7) corresponds to the application of the Gauss – Seidel's algorithm to the equation (5) with the initial conditions $y_0 = \omega^{-1}(f - Ax_0)$.

If, however, we use equation (6) as the matching conditions, and apply Kaczmarz's algorithm to equation (5), we obtain the row-oriented form of the algorithm6:

$$\begin{aligned} y_{k+1} &= y_k + \omega \rho_k e_j, \\ x_{k+1} &= x_k + \rho_k a_j. \end{aligned}, \ \rho_k = \frac{f_j - \omega e_j^T y_k - a_j^T x_k}{\left\| a_j \right\|^2 + \omega^2}. \tag{8}$$

Here $A = (a_1, \ldots, a_m)^T \in \mathbb{R}^{m\times n}$, $f = (f_1, \ldots, f_m) \in \mathbb{R}^m$, $E_m = (e_1, \ldots, e_m)$, $k = 0,1,\ldots$, $j = j(k) = (k \bmod m) + 1$, $\{j(k)\}_{k=0}^{\infty}$ is periodic sequence of the form $1,2,\ldots,m,1,2,\ldots,m,\ldots$. The initial conditions x_0 and y_0 satisfy $x_0 = \omega^{-1} A^T y_0$.

The sequence $j(k)$ is chosen according to the cyclic rule in the considered regularized forms of the algorithm (7) and (8), what greatly simplifies the implementation and computational costs of the algorithms. However, such rule of this choice has one drawback - the slow rate of convergence.

Various approaches are used to select a projection sequence for increasing the rate of convergence of Kaczmarz - type algorithms, one of them is the random (randomized) selection rule for the sequence $j(k)$ [9,10]. For example, in[11], the sequence is proposed to choose in proportion to the square of the norm of the matrix rows. However, if the row (or rows) of the matrix have a rate much larger than the others, the random selection rule for the sequence $j(k)$ will not lead to a decrease in the number of iterations. The column-oriented regularized form of the Kaczmarz's algorithm (7), with the random selection rule $j(k)$, considered in this work, is successfully used in solving certain classes of applied problems[5,12].

Greedy rules for the selection of the sequence $j(k)$ were proposed for the Kaczmarz -type algorithms in[13] and[14] - the maximum distance rule and the maximum residual rule, respectively. It is shown in[15] that, the computational costs for the greedy and random[11] selection rules are comparable for a certain class of problems, but the algorithm with the greedy selection rules has a higher convergence rate.

Then, the variants of the regularized forms of the Kaczmarz's algorithm, in which the sequences $j(k)$ are based on the greedy selection rules, will be considered.

3. Accelerated Regularized Forms of Kaczmarz's Algorithm

The formula for generating the sequence $j(k)$ based on the maximum distance rule (MDR) for algorithm (7) has the following form:

$$j(k) = \arg\max_{i\in[1,n]} \left| a_i^T y_k - \alpha e_i^T x_k \right|, k = 0,1,\ldots. \tag{9}$$

In the case of the maximum residual rule (MRR)

$$j(k) = \arg\max_{i\in[1,n]} \left| \frac{a_i^T y_k - \alpha e_i^T x_k}{\|a_i\| + \alpha} \right|, k = 0,1,\ldots. \tag{10}$$

For the algorithm (8) MDR and MRR respectively:

$$j(k) = \arg\max_{i\in[1,m]} \left| f_i - \omega e_i^T y_k - a_i^T x_k \right|, k = 0,1,\ldots. \tag{11}$$

$$j(k) = \arg\max_{i\in[1,m]} \left| \frac{f_i - \omega e_i^T y_k - a_i^T x_k}{\|a_i\| + \omega} \right|, k = 0,1,\ldots. \tag{12}$$

The pseudocodes of regularized modifications of the Kaczmarz's algorithm (7) and (8), with the greedy selection rules (9) - (12) of the projection sequence j (k), are presented below.

Input: $A \in \mathbb{R}^{m\times n}$, $f \in \mathbb{R}^m$

Set x_0 and α, calculate $y_0 = \omega^{-1}\left(f - Ax_0\right)$, $k = 0$;

while do

choose $j = j(k)$ using the formula (9) or (10);

$$\rho_k = \frac{a_j^T y_k - \alpha e_j^T x_k}{\left\|a_j\right\|^2 + \alpha};$$

$$y_{k+1} = y_k - \rho_k a_j;$$

$$x_{k+1} = x_k + \rho_k e_j;$$

$$k = k+1;$$

Algorithm 1 Column - oriented regularized form of Kaczmarz's algorithm (CRKMR, CRKMD).

Input: $A \in \mathbb{R}^{m\times n}$, $f \in \mathbb{R}^m$

Set y_0 and $w = \sqrt{\alpha}$, calculate $x_0 = \omega^{-1} A^T y_0$, $k = 0$;

while do

choose $j = j(k)$ using the formula (11) or (12);

$$\rho_k = \frac{f_j - \omega e_j^T y_k - a_j x_k}{\left\|a_j\right\|^2 + \omega^2};$$

$$y_{k+1} = y_k + \omega \rho_k e_j,$$

$$x_{k+1} = x_k + \rho_k a_j.$$

$$k = k+1;$$

Algorithm 2 Row - oriented regularized form of Kaczmarz's algorithm (RRKMR, RRKMD).

4. Numerical Experiments

We consider an example from the package AIR Tools II[16] for comparison the rate of convergence of the considered regularized forms of the Kaczmarz's algorithm. All calculations were performed in MATLAB R2018a 64-bit (Intel Core i7 3770 @ 3.4GHz processor and 32 Gb of RAM).

We enter the following notation:

x_* - exact pseudo-solution;

$\mathrm{Err} = \left\|x_* - x_k\right\| / \left\|x_*\right\|$ - relative error;

$\delta = \left\|\varepsilon\right\| / \left\|f\right\|$ - noise level;

k - - total number of iterations;

T - - algorithm execution time;

RRKS - is a row-oriented regularized form of the Kaczmarz's algorithm, with a sequential rule for choosing a projection sequence;

RRKRND – is a row-oriented regularized form of the Kaczmarz's algorithm, with a randomized selection rule for the projection sequence;

CRKS – is a column-oriented regularized form of the Kaczmarz's algorithm, with a sequential rule for choosing a projection sequence;

CRKRND – is a column-oriented regularized form of the Kaczmarz's algorithm, with a randomized selection rule for the projection sequence.

We consider the two-dimensional problem of seismic tomography as an example. This problem is reduced to finding the solution of the SLAE

$$Ax = \tilde{f},\ A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^{n},\ \tilde{f} = \left(f + \varepsilon\right) \in \mathbb{R}^{m}, \tag{13}$$

where f - the exact right-hand side, ε - Gaussian noise $\left(N\left(0,\sigma^2\right)\right)$. The regularization parameter α was calculated by the method proposed in[5]. The matrix, the exact solution, and the right-hand side were generated using the seismictomo function from16. The matrix is sparse and rectangular $(m>n)$ in the considered problem. The stopping criterion for iterative algorithms is $\left\|\theta_{k+1} - \theta_k\right\| \le 1e-06$.

The results of the computational experiments for different sizes of the matrix A are presented in Tables 1, 2 and 3 and Figures 1, 2, 3.

Table 1. Results for seismic tomography test problem: m = 392, n = 192, δ = 1e-04.

Algorithm's form	Number of iterations, k	Err	T, sec
RRKS	2801624	4.08e-02	21.09
RRKRND	999600	4.08e-02	9.87
RRKMR	331632	4.08e-02	5.18
RRKMD	331240	4.08e-02	6.13
CRKS	1345736	4.08e-02	10.93
CRKRND	970004	4.08e-02	12.79
CRKMR	248332	4.08e-02	3.90
CRKMD	249312	4.08e-02	5.81

Table 2. Results for seismic tomography test problem: m = 512, n = 256, δ = 1e-04.

Algorithm's form	Number of iterations, k	Err	T, sec
RRKS	4084736	7.65e-02	30.75
RRKRND	1165312	7.65e-02	11.55
RRKMR	390144	7.65e-02	7.16
RRKMD	391680	7.65e-02	8.63
CRKS	2328832	7.65e-02	18.84
CRKRND	1174528	7.65e-02	15.71
CRKMR	291584	7.65e-02	5.38
CRKMD	329984	7.65e-02	8.90

Table 3. Results for seismic tomography test problem: m = 1800, n = 900, δ = 1e-04.

Algorithm's form	Number of iterations, k	Err	T, sec
RRKS	39468600	1.01e-01	319.73
RRKRND	2453400	1.01e-01	28.95
RRKMR	819000	1.01e-01	27.20
RRKMD	828000	1.01e-01	37.31
CRKS	21584700	1.01e-01	202.95
CRKRND	2706300	1.01e-01	41.95
CRKMR	752400	1.01e-01	25.41
CRKMD	760500	1.01e-01	36.36

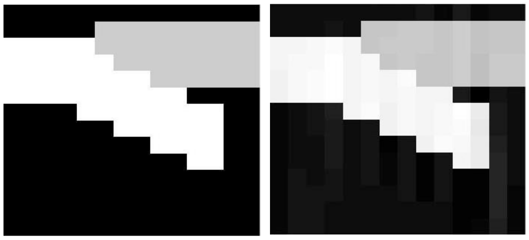
a  b

Fig. 1. First test problem m = 392, n = 192, δ = 1e-04. (a) exact solution; (b) computed solution.

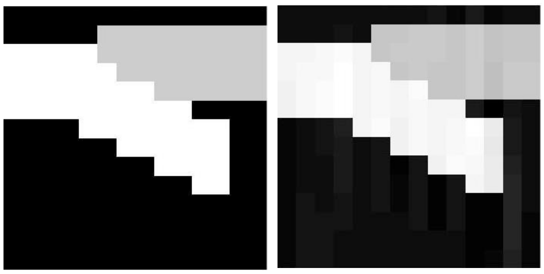
a  b

Fig. 2. Second test problem m = 512, n = 256, δ = 1e-04. (a) exact solution; (b) computed solution.

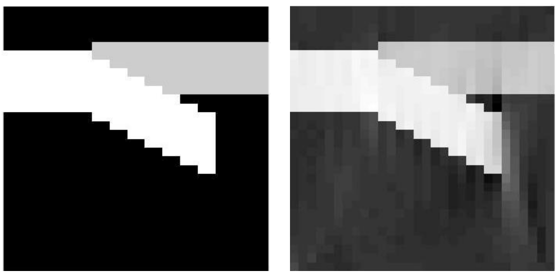
a  b

Fig. 3. Third test problem m = 1800, n =900, δ = 1e-04. (a) exact solution; (b) computed solution.

It can be seen from Tables 1, 2, 3 that the considered regularized forms of the Kaczmarz's algorithm with greedy rules for choosing a projection sequence can reduce the number of iterations from 5 to 28 times compare to sequential and randomized rules with an increase in the dimension of the problem. In spite of increasing the computational cost of the greedy selection rules for comparison with the rest of the considered rules, execution time is reduced to 3 times. In the example, the best variant is the regularized column-oriented form of the Kaczmarz's algorithm with the choice of the projection sequence according to the maximum distance rule (9).

5. Conclusion

The "greedy" methods for generating the projection sequence $j(k)$ to accelerate the convergence of regularized forms of the Kaczmarz's algorithm 4,5,6 are considered in this article.

As shown in numerical experiments, for certain classes of problems, such methods allow reducing both the number of iterations and the time for finding solutions. However, it should be noted that the "greedy" rules are more costly from a computational point of view compared to randomized ones. It is necessary to take into account not only the structure of the system matrix, but also the parallel computing systems used, when solving applied problems of large and extra-large dimensions, in order to choose the variant of accelerated forms of the considered algorithm.

It should be noted that the advantage of the row-oriented form of the Kaczmarz's regularized algorithm is sequential access to the incoming data, which can be effectively used for solving applied problems arising in computed tomography, image processing and signal processing.

6. References

1. Bjorck A. Linear Least Squares Problems / Numerical methods in matrix computations Texts in Applied Mathematics, 59. Berlin: Springer, 2015. pp. 211–430. doi: 10.1007978-3-319-05089-8_2.

2 Tikhonov A.N., Arsenin V.Ya. Methods for solving ill-posed problems. M .: Science, 1986. - 288 p.

3. Zhdanov A. The method of augmented regularized normal equations // Comput. Math. and Math. Phys., 2012. no. 52. pp. 194-197.

4. Zhdanov A. I., Ivanov A. A, "Projection Regularization Algorithm for Solving Linear Algebraic System of Large Dimension", Vestn. Samar. Gos. Tekhn. Univ., Ser. Fiz.-Mat. Nauki [J. Samara State Tech. Univ., Ser. Phys. Math. Sci.], 5(21)(2010), 309–312.

5. Ivanov A., Zhdanov A. Kaczmarz algorithm for Tikhonov regularization problem // Applied Mathematics E-Notes, 2013. no. 13. pp. 270-276.

6 Zhdanov A. I., Sidorov Yu. V. The row-oriented form of the regularized Kaczmarz's method, Vestn. Samar. Gos. Tekhn. Univ., Ser. Fiz.-Mat. Nauki [J. Samara State Tech. Univ., Ser. Phys. Math. Sci.], 2017, vol. 21, no. 3, pp. 546–555. doi: 10.14498/vsgtu1548 (In Russian).

7.Kaczmarz S. Angenäherte Auflösung von Systemen linearer Gleichungen // Bull. Acad. Polon. Sci. Lett. A, 1937. - Vol. 35. - P. 355–357.

8. Zhdanov A. I., Sidorov Yu. V., "Parallel implementation of a randomized regularized Kaczmarz's algorithm", CO, 39:4 (2015), 536–541 - 2452-2015-39-4-536-541.

9. Feichtinger H. G. Cenker C., Mayer M., Steier H., Strohmer T. New variants of the POCS method using affine subspaces of finite codimension with applications to irregular sampling. SPIE: VCIP, pages 299–310, 1992.

10 Herman G. T., Meyer L. B. Algebraic reconstruction techniques can be made computationally efficient. IEEE Trans. Medical Imaging, 12(3):600–609, 1993.

11. Strohmer T., Vershynin R. A randomized Kaczmarz algorithm with exponential convergence. J. Fourier Anal. Appl., 15:262–278, 2009.

12. Sinha, L.et al. Quantification of the binding potential of cell-surface receptors in fresh excised specimens via dual- probe modeling of SERS nanoparticles. Sci. Rep. 5, 8582; DOI:10.1038/srep08582 (2015).

13. Eldar Y. C., Needell D. Acceleration of randomized Kaczmarz methods via the Johnson – Lindenstrauss Lemma. Numer. Algor., 58:163–177, 2011.

14. Griebel M, Oswald P. Greedy and randomized versions of the multiplicative Schwartz method. Lin. Alg. Appl., 437:1596–1610, 2012.

15. Julie Nutini, Behrooz Sepehry, Issam Laradji, Mark Schmidt, Hoyt Koepke, and Alim Virani. 2016. Convergence rates for greedy Kaczmarz algorithms, and faster randomized Kaczmarz rules using the orthogonality graph. In Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence (UAI'16), Alexander Ihler and Dominik Janzing (Eds.). AUAI Press, Arlington, Virginia, United States, 547-556.

16. Hansen, P.C., Jørgensen, J.S. Numer Algor (2018) 79: 107. https://doi.org/10.1007/s11075-017-0430-x.