Applied Mathematics E-Notes, 13(2013), 270-276 © ISSN 1607-2510
Available free at mirror sites of http://www.math.nthu.edu.tw/∼amen/

Kaczmarz Algorithm For Tikhonov Regularization Problem*

Andrey Aleksandrovich Ivanov †, Aleksandr Ivanovich Zhdanov ‡

Received 25 February 2013

Abstract

In this article we present a modified version of Kaczmarz method for solving ill-posed systems of linear algebraic equations. This algorithm is based on transforming regularized normal equations to the equivalent augmented regularized normal system of equations. The proposed algorithm can effectively solve ill-posed problems of large dimensionality.

1 Statement of The Problem

Let us consider a standard A. N. Tikhonov's regularization problem

$$\min_{u\in\mathbb{R}^n}\left\{\|Au-f\|^2+\alpha\|u\|^2\right\}, \tag{1}$$

where $A\in\mathbb{R}^{m\times n}$, $f\in\mathbb{R}^m$, $\alpha>0$ is a regularization parameter, $\|\cdot\|=\|\cdot\|_2$ is Euclidean norm.

If matrix A has large dimensionality and is probably highly sparse, iterative methods often represent practically the only method for solving problem (1). However, most recent iterative algorithms for solving problem (1) are based on solving Euler equations (regularized normal equations)

$$(A^{\mathrm{T}}A+\alpha I_n)u=A^{\mathrm{T}}f, \tag{2}$$

where T denotes transposition, I_n is the identity matrix of order n.

Taking into account that condition number of problem (2) is approximately equal to the square of the condition number of the initial problem (1), it is practically impossible to solve problem (1) using the popular iterative methods.

The classical iterative methods easily lead to slow convergence when A has a large condition number in (1). Many studies were devoted to this problem. For example, an accelerated predictor–corrector iterated Tikhonov regularization was proposed [1] by combining the classical iterated Tikhonov regularization with modified Euler method.

*Mathematics Subject Classifications: 65F10, 47A52, 65F22, 65F20.

†Department of Applied Mathematics, S. P. Korolev Samara State Aerospace University, Samara 443086, Russia,

‡Department of Applied Mathematics and Informatics, Samara State Technical University, Samara 443120, Russia.

This paper is based on the famous Kaczmarz's method [2] and its block-based version suggested in [3]. We transform the system (1.2) to augmented regularized system of linear equations [4] which is always consistent and determined (for the $\alpha > 0$). We apply the Kaczmarz algorithm to this augmented system and propose original modification.

It is suggested here that the regularization parameter α is known. One of the easiest algorythms for selecting a regularization parameter was proposed by V. A. Morozov and S. F. Gilyazov [5]. According to this approach the regularization parameter α is selected $\alpha = h$, where h is a quantity that characterizes the error of specifying the elements of matrix A. This algorithm for selecting a regularization parameter means that the noise level in the initial task data is known. However, if estimation of the noise level is inaccurate, that would lead to many errors in regularizable solution. Here a set of stable rules for choosing a regularization parameter is offered [6]. They are proved to be stable from the point of view of perturbance level assessment.

It should be mentioned that there are regularization parameter selection rules which demand no prior information, for example [7].

In conclusion we present the computing experiment results. They demonstrate the efficiency of the suggested algorithm for solving the problem of regularization, and in particular, the efficiency when we apply the randomized Kaczmarz method [8].

2 Method of Augmented Regularized Normal Equations

As shown in [4], the regularized normal system of equation (2) can be written as

$$\begin{pmatrix} \omega I_m & A \\ A^{\mathrm{T}} & -\omega I_n \end{pmatrix} \begin{pmatrix} y \\ u \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad \tilde{A}_\omega z = \tilde{f}, \tag{3}$$

where $\omega = \sqrt{\alpha}$, and $I_n \in \mathbb{R}^{n\times n}$ and $I_m \in \mathbb{R}^{m\times m}$ are the identity matrices.

Matrix $\tilde{A}_\omega$ of the system (3) is nonsingular for all $\alpha > 0$ [4] and its only solution is a vector $z_* = (y_*^T, u_*^T)^T$, where $u_* = (A^TA + \alpha I_n)^{-1}A^Tf$, $y_* = \omega^{-1}r_*$, $r_* = f - Au_*$.

The spectral condition number of the regularized normal systems 2 is

$$\kappa_2(A^TA + \alpha I_n) = \frac{\sigma_{\max}^2 + \alpha}{\sigma_{\min}^2 + \alpha},$$

where $\sigma_{\min}$ and $\sigma_{\max}$ are the minimal and maximal singular numbers of matrix A respectively. The spectral condition number of the augmented regularized normal system (3) is considerably smaller and is equal to

$$\kappa_2(\tilde{A}_\omega) = \sqrt{\kappa_2(A^TA + \alpha I_n)}.$$

3 Projection Algorithm

Let us write the augmented system (3) as a system of two equations:

$$(\omega I_m | A)z = f, \tag{4}$$

$$(A^T | -\omega I_n)z = 0. \tag{5}$$

To solve the problem (4)–(5), we use the block Kaczmarz algorithm [3] (projection algorithm):

$$z_{i,1} = z_{i,0} + (\omega I_m | A)^+ [(\omega I_m | A) z_{i,0} - f], \tag{6}$$

$$z_{i,t} = z_{i,t-1} - \beta_{i,t-1} \begin{pmatrix} a_{t-1} \\ -\omega e_{t-1} \end{pmatrix}, \tag{7}$$

where

$$z_{i+1,0} = z_{i,n+1}, \quad i = 1, 2, \ldots, \beta_{i,t-1} = \frac{(a_{t-1}^T, -\omega e_{t-1}^T) z_{i,t-1}}{\|(a_{t-1}^T, -\omega e_{t-1}^T)^T\|^2}, \quad t = 2, 3, \ldots, n+1,$$

i is the number of external iterations, and t is the number of internal iterations (the total number of internal iterations for algorithm (6)-(7) is equal to $n+1$), $(e_1, \ldots, e_n) = I_n$, $A = (a_1, \ldots, a_n)$, $(\cdot)^+$ is a pseudoinverse matrix.

Matrix $\tilde{A}_\omega$ is nonsingular for all cases of $\alpha > 0$ [4], consequently [2], [3],

$$z_{i,n+1} \overset{i \to \infty}{\longrightarrow} z_*$$

for any initial value of vector $z_{1,0}$.

If we denote $u = (u^{(1)}, \ldots, u^{(n)})^T$, then, taking into account that vector $z = (y^T, u^T)^T$, the recurrent equations (7) can be written as two recurrent equations:

$$y_{i,t} = y_{i,t-1} - \beta_{i,t-1} a_{t-1}, \quad \beta_{i,t-1} = \frac{a_{t-1}^T y_{i,t-1} - \omega u_{i,t-1}^{(t-1)}}{\|a_{t-1}\|^2 + \omega^2}, \tag{8}$$

$$u_{i,t}^{(t-1)} = u_{i,t-1}^{(t-1)} + \omega \beta_{i,t-1}, \quad t = 2, 3, \ldots, n+1. \tag{9}$$

Since $y = \omega^{-1} r$, where $r = f - Au$, recurrent equations (8)–(9) can be transformed as follows:

$$r_{i,t} = r_{i,t-1} - \rho_{i,t-1} a_{t-1}, \quad \rho_{i,t-1} = \frac{a_{t-1}^T r_{i,t-1} - \alpha u_{i,t-1}^{(t-1)}}{\|a_{t-1}\|^2 + \alpha}, \tag{10}$$

$$u_{i,t}^{(t-1)} = u_{i,t-1}^{(t-1)} + \rho_{i,t-1}, \quad t = 2, 3, \ldots, n+1. \tag{11}$$

We note that classifying iteration steps $u_{i,t}$ and $r_{i,t}$ as internal and external ones in equations (10)–(11) is optional and the computation process can be presented as the calculation of "microiterations" according to a certain single parameter $k = 1, 2, \ldots$. Thus, in this case, it is necessary for convergence that all the columns of matrix A participate in the sequence of vectors $a_1, a_2, \ldots$.

Moreover, it will be shown that it is *not necessary to use the first recurrent equation* (6) in the iteration process if some additional prerequisite of matching the initial conditions $u_0 = u_{1,0}$ and $r_0 = r_{1,0}$ is fulfilled. Equation (4) is used *only for matching the initial conditions* u_0 *and* r_0.

Let $k = 1, 2, \ldots$ and $j(k) = (k-1) \bmod(n) + 1$. Consequently, $\{j(k)\}_{k=1}^{\infty}$ is a periodic sequence of the kind $1, 2, \ldots, n, 1, 2, \ldots, n, \ldots$. Then recurrent equations (10)–(11) can be written as follows

$$r_k = r_{k-1} - \rho_{k-1} a_{j(k)}, \tag{12}$$

$$u_k = u_{k-1} + \rho_{k-1} e_{j(k)}, \tag{13}$$

where

$$\rho_{k-1} = \frac{a_{j(k)}^T r_{k-1} - \alpha e_{j(k)}^T u_{k-1}}{\|a_{j(k)}\|^2 + \alpha}, \quad k = 1, 2, \ldots .$$

Index k in equations (12)–(13) and indices i, t and n in equations (10)–(11) are related by $k = (i-1)n + t - 1$, and $t - 1 = j(k)$. It is obvious, then, that $r_0 = r_{1,0}$ and $u_0 = u_{1,0}$.

Let us introduce vector $\theta_k = (r_k^T, u_k^T)^T$. Then recurrent equations (12)–(13) can be written as one recurrent equation

$$\theta_k = \theta_{k-1} + \rho_{k-1} \begin{pmatrix} a_{j(k))} \\ -e_{j(k))} \end{pmatrix}, \tag{14}$$

where θ_0 is the vector of initial values and $\rho_{k-1} = \frac{(a_{j(k)}^T, -\alpha e_{j(k)}^T)\theta_{k-1}}{\|a_{j(k)}\|^2 + \alpha}, \quad k = 1, 2, \ldots .$

THEOREM 1. Let vector $\theta_0 = (r_0^T, u_0^T)^T$ satisfy the condition of matching

$$r_0 = f - Au_0. \tag{15}$$

in recurrent equation (14). Then, for an arbitrary initial vector u_0, $\theta_k \longrightarrow \theta_*$ as $k \longrightarrow \infty$, where $\theta_* = (r_*^T, u_*^T)^T$.

PROOF. Let us construct the proof by mathematical induction. From the condition of matching (15) of initial values it follows that the condition of matching is fulfilled

$$r_k = f - Au_k \tag{16}$$

for any $k \geq 0$ where r_k and u_k are calculated from recurrent equations (12)–(13).

For $k = 1$ we obtain from (12), (13)

$$f - Au_1 = f - A(u_0 + \rho_0 e_1) = (f - Au_0) - \rho_0 a_1 = r_0 - \rho_0 a_1 = r_1.$$

Thus, the condition of matching (16) is fulfilled for $k = 1$. Let us assume that (16) is fulfilled for some arbitrary $k = \nu > 1$.

Let us show that if (16) for $k = \nu$ holds, then also (16) for $k = \nu + 1$ holds. From recurrent equation (12), (13) we immediately obtain

$$f - Au_{j(\nu+1)} = f - A(u_\nu + \rho_\nu e_{j(\nu+1)}) = (f - Au_\nu) - \rho_\nu a_{j(\nu+1)} = r_\nu - \rho_\nu a_{j(\nu+1)} = r_{j(\nu+1)}.$$

Since both the basis and the inductive step have been performed, by mathematical induction, the statement (16) holds for any $k = 1, 2, \ldots$.

If in recurrent equations (6)–(7) $y_{1,0}$ and $u_{1,0}$ satisfy the condition

$$\omega y_{1,0} = f - Au_{1,0},$$

$z_{1,1} = z_{1,0}$ and consequently vectors $r_{i,t}$ and $u_{i,t}$ from (10), (11) for all $k = (i-1)n + t - 1$ completely coincide with vector θ_k from equation (14). From the validity of condition (16) for all $k = 1, 2, \ldots$ we immediately find that $z_{i,1} = z_{i,0}$ and recurrent equation (6) is *not necessary* if condition (15) is fulfilled.

The validity of the theorem conclusively follows from the complete equivalence of recurrent equations (10), (11) and (14) if condition (15) is fulfilled.

4 Numerical Experiment

In this section, we consider Phillips's "famous" test problem [10]. Consider the Fredholm integral equation of first kind on the square $[-6,6]\times[-6,6]$ with kernel function

$$K(s,t)=\phi(s-t),\ u(t)=\phi(t),$$

and right-hand side

$$f(s)=(6-|s|)\left(1+\frac{1}{2}\cos\left(\frac{s\pi}{3}\right)\right)+\frac{9}{2\pi}\sin\left(\frac{|s|\,\pi}{3}\right). \tag{17}$$

The pertubed problem $A\tilde{u}=\tilde{f}$, $A\in R^{n\times n}$, $\tilde{u}\in R^n$, $\tilde{f}\in R^n$ is constructed for

$$\tilde{f}=\frac{1}{K}\sum_{k=1}^{K}\tilde{f}_k,$$

where $\tilde{f}_k=(f+\varepsilon_k)$, ε_k - is Gaussian noise with $\frac{\|f-f_k\|}{\|f\|}=0.1$, $\left\|f-\tilde{f}\right\|\approx 0.014$ and $K=50$ [9].

The algorithm regularization parameter α is chosen applying the following rule [5]

$$\alpha=\frac{\delta\sigma_{max}^2(A)}{\left\|\tilde{f}\right\|_2+\delta},\ \left\|f-\tilde{f}\right\|\le\delta, \tag{18}$$

where $\sigma_{\max}(A)$ - is maximum singular value of matrix A. The parameter δ is approximately assessed as

$$\delta^2=\frac{1}{K}\sum_{k=1}^{K}\left\|\tilde{f}-\tilde{f}_k\right\|^2. \tag{19}$$

This way of choosing regularization parameter was also successfully used by authors in this article [9].

For formulating the problem two algorithms were used: classical Kaczmarz algorithm [10] and regularized Kaczmarz algorithm. Randomized modification with using the result from [8] was also represented for regularized algorithm. For each algorithm no more than 6000 iterations were made.

All source codes for MATLAB were published by authors in [11].

There are solutions of problems in Figure 1 received by using under examination algorithm and the solution of exact problem. In particular absolute and relative errors while solving Kaczmarz algorithm reach 1.49 and 0.49 correspondingly. While solving this problem by proposed algorithm, absolute and relative errors reach 0.16 and 0.056 correspondingly.

Randomization algorithm slightly influences the accuracy of regularized solution but the rate of convergence is much better. There are diagrams in Figure 2 which show the degree of convergence of each under examination algorithm. Classical Kaczmarz algorithm was not convergent to the solution even for 6000 iterations, regularized Kaczmarz algorithm was convergent to the solution for 1485 iteration while randomization regularized algorithm made it possible to get the solution only for 47 iteration.

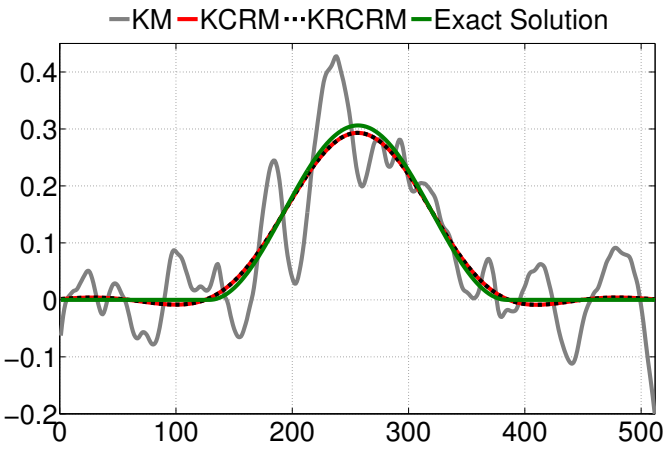

Figure 1: The solutions of testing perturbed problem received by using classical Kaczmarz method (KM), regularized method (proposed in this article, KCRM - Kaczmarz Column Regularized Method) and randomized regularized version (KRCRM); and also the solution of the exact problem (green line).

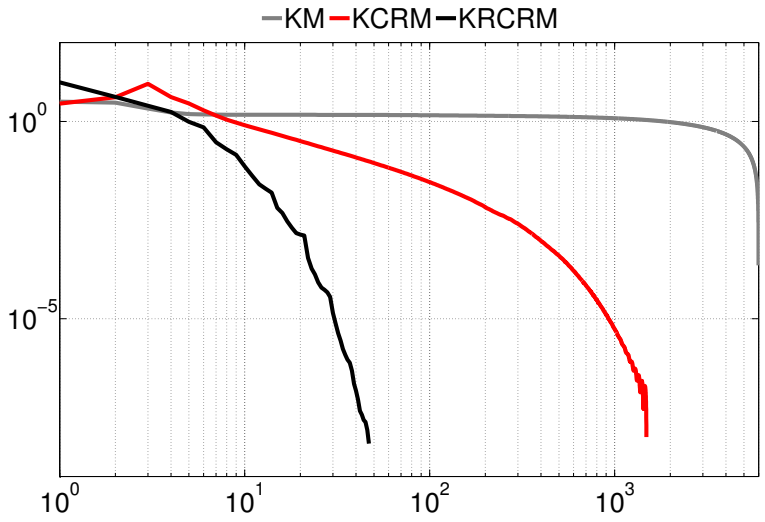

Figure 2: The dependence of $e_k = \|u_{\text{inf}} - u_k\|$ on iteration number for classical Kaczmarz algorithm, regularized modification (proposed in this article) and randomized regularized version.

5 Conclusion

The obtained results make it possible to suppose that the usage of regularized algorithms, especially in randomization modification, allows for less iteration get more accurate solution of perturbed problem. It means that this method is highly effective. It is also noticed that only two matrix operations are used: dot product and saxpy operation in one iteration. This method allows to solve the incorrect and ill-conditioned problems of large dimensionality efficiently.

References

[1] W. Liu and C. Wu, A predictor–corrector iterated Tikhonov regularization for linear ill-posed inverse problems, Appl. Math. Comput., 221(2013), 802–818.

[2] S. Kaczmarz, Angenäherte Auflösung von Systemen linearer Gleichungen, Bull. Internat. Acad. Polon. Sci. Lett. A, 35(1937), 335–357.

[3] G. P. Vasil'chenko and A. A. Svetlakov, A projection algorithm for solving systems of linear algebraic equations of high dimensionality, Comput. Math. Math. Phys., 20(1980), 1–8.

[4] A. I. Zhdanov, The method of augmented regularized normal equations, Comput. Math. Math. Phys., 52(2012), 194–197.

[5] V. A. Morozov, Methods of Solving Incorrectly Posed Problems, Springer Verlag, New York, 1984.

[6] U. Hämarik, R. Palm and T. Raus, A family of rules for parameter choice in Tikhonov regularization of ill–posed problems with inexact noise level, J. Comput. Appl. Math., 236(2012), 2146–2157.

[7] A. I. Zhdanov, Optimal regularization of solutions of approximate stochastic systems of linear algebraic equations, Comput. Math. Math. Phys., 01(1990), 224–229.

[8] T. Strohmer and R. Vershynin, A randomized Kaczmarz algorithm for linear systems with exponential convergence, J. Fourier Anal. Appl., 15(2009), 262–278.

[9] A. K. Kaveev, G. I. Kropotov, E. V. Tsygankova, I. A. Tzibizov, S. D. Ganichev, S. N. Danilov, P. Olbrich, C. Zoth, E. G. Kaveeva, A. I. Zhdanov, A. A. Ivanov, R. Z. Deyanov and B. Redlich, Terahertz polarization conversion with quartz waveplate sets, Applied Optics, 52(4)(2013), B60–B69.

[10] P. C. Hansen, Regularization Tools Version 4.0 for Matlab 7.3, Numer. Algorithms, 46(2007), 189–194.

[11] A. A. Ivanov, Regularization Kaczmarz Tools Version 1.0 for Matlab, Matlabcentral Fileexchange,
URL: *http://www.mathworks.com/matlabcentral/fileexchange/43791.*