（別紙様式３）

博　士　論　文　要　旨

平成 25 年　5 月　31 日

出願者

複合科学研究科　情報学専攻

ふりがな　もりくに　けいいち
氏　名　保 國　惠 一

（学籍番号）20081754

論文題目　Inner-iteration Preconditioning for Least Squares Problems

要旨　General-purpose preconditioners for the conjugate gradient (CG) and generalized minimal residual (GMRES) type methods are proposed for solving the linear least squares problem

$$\min_{\boldsymbol{x} \in \mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$$

and the general least squares problem

$$\min_{\boldsymbol{x} \in S} \|\boldsymbol{x}\|_2, \ S = \{\boldsymbol{x} \in \mathbf{R}^n : \|\boldsymbol{b} - A\boldsymbol{x}\|_2 = \min_{\boldsymbol{\xi} \in \mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{\xi}\|_2\},$$

where $A \in \mathbf{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbf{R}^m$. Numerical experiments show their effectiveness. We develop their convergence theory and evaluate one of the proposed methods for practical problems.

Background.

Solving linear least squares problems is a fundamental requirement in a wide range of areas across science, engineering, industry, and statistics, in particular, signal processing, control, tomography, geodetics, curve fitting, optimization etc. Hence, a significant point is to design robust, efficient, and reliable methods for computing least squares solutions. Conventional solution methods for the problems are direct methods such as the Cholesky and QR factorizations, and the singular value decomposition. Such methods are effective for solving relatively small or dense problems. However, there are ever-increasing demands for solving large-scale and complex problems. Such requirements are far beyond existing computers and computational methods. Thus, it is becoming difficult to apply those methods to recent large and sparse problems in terms of time and space complexity. With improvement of computational methods, this work attempts to take measures to deal with this coming serious situation.

Iterative methods are also well-established solution methods for the problems. The Jacobi and successive overrelaxation type methods (JOR and SOR) applied to the normal equations inspired by the Kaczmarz method and the Cimmino method, and Krylov subspace iterative methods such as the CGLS method proposed by Hestenes and Stiefel [12] and the LSQR methods developed by Paige and Saunders [16] have been studied and used for solving large and sparse linear least squares problems arising from many application fields. If the problems are well-conditioned, then these methods converge fast. Otherwise, the convergence becomes slow. Then, preconditioning is necessary to accelerate the convergence. It is known that appropriate preconditioners dramatically improve the convergence of Krylov subspace methods and achieve less storage requirement for Krylov subspace methods.

Motivation.

The majority of the studies on preconditioning Krylov subspace methods for least squares problems is devoted to incomplete matrix factorizations such as the robust incomplete factorization (RIF) developed by Benzi and Tůma [2]. However, such techniques require preconditioning time and memory for computing and storing incomplete matrix factors of the matrix.

Different types of preconditioners called inner iterations have been extensively studied and developed in the context of solutions to square systems of linear equations, e.g., DeLong and Ortega [8, 9] and Saad [17]. However, these types of preconditioners have not been studied so much and the sufficient conditions for their convergence are not well known in the context of solutions to least squares problems (cf Aoto, Ishiwata and Abe [1]).

On the other hand, an application of the generalized minimal residual (GMRES) method developed by Saad and Schultz to linear least squares problems was proposed by Hayami, Yin and Ito [11], where the right- and left-preconditioned GMRES methods (AB- and BA-GMRES) were combined with RIF. The convergence conditions of GMRES for singular systems are well understood due to the work by Brown and Walker [3] and Hayami and Sugihara [10]. These studies motivated us to investigate AB- and BA-GMRES more in depth in the thesis.

Previously AB- and BA-GMRES preconditioned by RIF was comparable with, but not definitely superior to, the reorthogonalized CGLS (or the CG normal error (CGNE) method) preconditioned by RIF in terms of time complexity. Moreover, many previous preconditioners based on incomplete matrix factorization such as RIF will break down for rank-deficient matrices. The Greville preconditioner proposed by Cui, Hayami and Yin is an exception, but was not definitely superior to previous methods [5]. Few authors addressed the problem of preconditioning Krylov subspace methods for solving rank-deficient least squares problems with sufficient theoretical justification. To the author's knowledge, little has been done in preconditioning in the rank-deficient case.

Least squares problems have an infinite number of solutions if the problems are not full column rank. On the other hand, general least squares problems have a unique solution called the pseudo-inverse solution, whose Euclidean norm is minimum. CGLS without preconditioning with an appropriate initial approximate solution determines the pseudo-inverse solution. However, little has been done in preconditioning general least squares problems even though the pseudo-inverse solution is useful in many applications such as inverse problems and control.

Objectives.

Based on the above mentioned points, the main objectives for proposing the new preconditioners is to reduce time and space complexity significantly, broaden the scope of problems that can be solved to the rank-deficient case, and remove the above-mentioned drawbacks. Based on the understanding of the solution methods for least square problems, we describe a comprehensive treatment of preconditioners for least squares problems and general least squares problems.

Organization.

The thesis is organized as follows. In Chapter 1, we describe the background, motivation, and objectives of the thesis. In Chapter 2, we prepare relevant basics and notations, and explain about the least squares problems which to be solved. In Chapter 3, we describe existing methods for solving least squares problems, including direct approaches, stationary iterations, Krylov subspace methods, and preconditioners.

In Chapter 4, we present the main results. We design new general-purpose preconditioners based on linear stationary iterative methods for Krylov subspace methods such as CGLS, CGNE, and AB- and BA-GMRES for solving linear least squares problems including the rank-deficient case. The proposed preconditioners are given by several steps of linear stationary iterative methods, which are regarded as inner iterations, and serve as preconditioners for the Krylov subspace methods, which are regarded as outer iterations.

We first consider using general linear stationary iterative methods applied to the normal equations as the inner-iteration preconditioners. With the help of the convergence theory for GMRES- and CG-type methods for least squares problems including the rank-deficient case and linear stationary iterative methods for singular systems of linear equations, we develop a general convergence theory for AB- and BA-GMRES, CGNE, and CGLS preconditioned by inner iterations. That is, we show that a sufficient condition for the proposed methods to determine a least squares or the pseudo-inverse solution within r iterations without breakdown for arbitrary initial approximate solution is that the inner-iteration matrix H is *semi-convergent*, i.e., $\lim_{i \to \infty} H^i$ exists, where $r = \mathrm{rank} A$. This theory holds irrespective of whether A is over- or under-determined and whether A is of full-rank or rank-deficient. In addition, we correct the previous convergence theory for AB- and BA-GMRES given by Hayami, Yin and Ito [11] .

We next consider using specific linear stationary iterative methods as the inner-iteration preconditioners. It was shown by Dax [7] that JOR and (S)SOR applied to the normal equations give a semi-convergent iteration matrix with a value of the relaxation parameter in an appropriate range. There exist efficient implementations of these specific iterative methods called the Cimmino-NE and NR and NE- and NR-(S)SOR methods in terms of time and space complexity and we use them.

Compared to previous preconditioners based on incomplete matrix factorizations, the advantage of these methods is that they enable one to avoid computing and storing the normal equation matrix and factors of the preconditioning matrix explicitly.

We analyze the spectrum of the preconditioned coefficient matrix, and characterize it by the spectral radius of the inner-iteration matrix $\rho(H)$ and the number of inner iterations ℓ and then show that the nonzero eigenvalues lie inside the circle of radius $\rho(H)^\ell$ with center at unity and they approach unity as the number of inner iterations ℓ increases. This theoretical result is tested by numerical examples. In addition, we show that the zero eigenvalues do not affect the residual norm convergence of the proposed GMRES-type methods.

We left the analysis of the convergence rate for the proposed methods for the future. It is know that the residual norm convergence of GMRES for general matrices is not necessarily determined only by the spectrum of the matrices. On the other hand, experiments often show a correlation between the spectrum and the convergence property. Hence, we hope to study more in depth the convergence analysis in connection with the spectrum of the preconditioned matrix given by inner iterations.

The SOR-type inner-iteration preconditioner uses two parameters, the relaxation parameter and the number of inner iterations. CPU time for the preconditioned iterative method varies with the values of these parameters. Hence, we need to determine the values of these parameters before the outer iterations start. In Chapter 5, we propose a procedure to automatically tune the value of the parameters in terms of time complexity, and showed that it is effective.

In Chapter 6, numerical experiments on large and sparse overdetermined least squares problems with artificial ill-conditioned and practical matrices from [6] illustrate that the proposed methods are efficient and robust, and serve as powerful preconditioners especially for ill-conditioned and rank-deficient problems, outperforming previous methods such as the CGLS and CGNE methods preconditioned by the diagonal scaling and RIF.

More work has to be done regarding the stopping criteria. In the experiments, we judged the convergence of the methods explicitly by using the residual norm of the normal equations $\|A^\mathsf{T}\boldsymbol{r}_k\|_2$. However, it would be ideal to judge the convergence of the methods in terms of estimates of the residual norm $\|A^\mathsf{T}\boldsymbol{r}_k\|_2$, error norm $\|\boldsymbol{x}_k - \boldsymbol{x}_*\|_2$, and backward error of $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$, where $\boldsymbol{x}_*$ is a least squares solution.

The inner-iteration preconditioning in Chapter 4, the automatic parameter tuning procedure in Chapter 5, and numerical experiment results in Chapter 5 were presented by the author and Hayami [13] (see also [?]). The convergence theory in Chapter 4 was given by the author and Hayami [14].

Numerical experiments on the proposed methods comparing with previous methods on test least squares problems led us to focus on the behavior of AB-GMRES for inconsistent problems. In Chapter 7, we indicate that AB-GMRES numerically fails to converge for inconsistent problems even though it is theoretically convergent. Based on this observation, in order to overcome the defect, we propose using BA-GMRES for solving inconsistent problems instead. We have not completed the analysis of the numerical behavior of the GMRES-type methods applied to inconsistent problems. We hope to study more in depth these subjects in conjunction with numerical experiments on the proposed methods in the future.

In Chapter 8, we consider preconditioning and solving general least squares problems. We present preconditioned CG and GMRES-type methods for general least squares problems, give their convergence theories, and evaluate the methods through numerical experiments. However, more numerical experiments on test problems are required to fully evaluate the proposed methods comparing them with CGLS.

In Chapter 9, in order to evaluate the proposed methods in a practical application, we apply AB-GMRES preconditioned by SOR-type inner iterations to image reconstruction problems arising from the use of electron microscopes in biology, where large least square problems arise. Here, we combine the method with the Tikhonov regularization in order to smooth the images, and with restarts in order to satisfy the non-negativity constraint, and call this the restarted regularized AB-GMRES (RRAB-GMRES) method. Numerical experiments on relatively small and large problems show that the method performs competitively with previous methods such as the Algebraic Reconstruction Technique (ART) and CG-type methods, although not decisively.

A part of the results in Chapter 9 was presented by the author, Hosoda and Hayami [15].

We did not address the problem of how to choose the regularization parameter and operator and set the stopping criterion. Further investigation in terms of numerical experiments on more test problems is required. Also, in order to improve the accuracy, we need to utilize the characteristic of the problem. Introducing appropriate constraints to the approximate solution would give a better solution.

In Chapter 10, we conclude the thesis and summarize the results obtained for the proposed methods for least squares problems and general least squares problems, respectively. BA-GMRES with NR-SOR is recommended for solving overdetermined and inconsistent underdetermined problems. AB-GMRES with NE-SOR is recommended for solving consistent underdetermined problems. Our theory filled a gap by investigating preconditioners for rank-deficient cases and applying AB- and BA-GMRES to the case of general least squares problems.

Contributions.

This work contributes to the development of new methods for solving least squares problems.

Recently, linear stationary iterative methods alone are regarded as classical approaches and not useful in general. However, this work showed that some of them combined with Krylov subspace methods as preconditioners serve as powerful preconditioners and play a significant role in accelerating the convergence of the Krylov subspace methods, when used as inner iterations for least squares problems. Investigations of stationary iterative methods would still contribute to the development of preconditioners.

From the theoretical point of view, we showed that *semi-convergence* of the inner-iteration matrix is intrinsically significant in the determination of convergence of CG- and GMRES-type methods preconditioned by inner iterations. Semi-convergence is a useful property for the convergence analysis of linear stationary iterative methods for solving singular systems of linear equations. Here, we utilize this property for the convergence analysis of the Krylov subspace methods preconditioned by inner iterations for least squares problems in the full rank and rank-deficient cases.

Our methods are promising candidates which would meet the increasing demands for solving large-scale least squares problems in the future.

Future work.

We are interested in the application of the proposed methods to least squares problems arising from the interior point method for linear programming problems [4] and tomographic imaging reconstruction problems arising from the use of electronic astronomical telescopes in which least squares problems with sparse, ill-conditioned and sometimes rank deficient coefficient matrices arise.

References

[1] D. Aoto, E. Ishiwata, and K. Abe, *A variable preconditioned GCR(m) method using the GSOR method for singular and rectangular linear systems*, J. Comput. Appl. Math., 234 (2010), pp. 703–712.

[2] M. Benzi and M. Tůma, *A robust preconditioner with low memory requirements for large sparse least squares problems*, SIAM J. Sci. Comp., 25 (2003), pp. 499–512.

[3] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.

[4] X. Cui, *Approximate Generalized Inverse Preconditioning Methods for Least Squares Problems*, PhD thesis, The Graduate University for Advanced Studies, Japan, 2009.

[5] X. Cui, K. Hayami, and J.-F. Yin, *Greville's method for preconditioning least squares problems*, Adv. Comput. Math., 35 (2011), pp. 243–269.

[6] T. Davis, *The University of Florida Sparse Matrix Collection*, `http://www.cise.ufl.edu/research/sparse/matrices/`.

[7] A. Dax, *The convergece of linear stationary iterative processes for solving singular unstructured systems of linear equations*, SIAM Review, 32 (1990), pp. 611–635.

[8] M. A. DeLong and J. M. Ortega, *SOR as a preconditioner*, Appl. Numer. Math., 18 (1995), pp. 431–440.

[9] ——, *SOR as a preconditioner II*, Appl. Numer. Math., 26 (1998), pp. 465–481.

[10] K. Hayami and M. Sugihara, *A geometric view of Krylov subspace methods on singular systems*, Numer. Linear Algebra Appl., 18 (2011), pp. 449–469.

[11] K. Hayami, J.-F. Yin, and T. Ito, *GMRES methods for least squares problems*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2400–2430.

[12] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436.

[13] K. Morikuni and K. Hayami, *Inner-iteration Krylov subspace methods for least squares problems*, NII Technical Report NII-2011-001E, `http://www.nii.ac.jp/TechReports/ 11-001E.html`, (2011), pp. 1–27.

[14] ——, *Convergence of inner-iteration GMRES methods for least squares problems*, NII Technical Report NII-2012-005E, `http://www.nii.ac.jp/TechReports/ 12-005E.html`, (2012), pp. 1–9.

[15] K. Morikuni, Y. Hosoda, and K. Hayami, *Inverse problems arising from image reconstruction in electron microscopes*, Meeting of Electronics, Information and System Society, The Institute of Electrical Engineers of Japan, (2012), pp. 931–933 (Book of Abstracts, in Japanese).

[16] C. C. Paige and M. Saunders, *LSQR: An algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software, 8 (1982), pp. 43–71.

[17] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.

（備考）1　和文で作成する場合は 2,000 字～ 3,000 字、英文で作成する場合は 700 語～2,000 語程度で作成すること。

2　用紙の大きさは、日本工業規格Ａ４縦型とすること。