An orthogonal randomized Bregman projection method for linearly constrained optimization problems

葉雨欣 (Ye Yu-Xin)[1,2], 王澤 (Wang Ze)[1,2], 繆樹鑫(Miao Shu-Xin)[2,3], 殷俊鋒 (Yin Jun-Feng)[1], 張紹良 (Zhang Shao-Liang)[2]
[1] 同済大学 (Tongji University), [2] 名古屋大学 (Nagoya University), [3] 西北師範大学 (Northwest Normal University)
e-mail : yeyuxin@tongji.edu.cn

1 Introduction

Consider solving the linearly constrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \quad \text{s.t. } Ax = b, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and the objective function f is strongly convex but possibly nonsmooth. Such problem often arises in many areas of scientific computing, such as compressed sensing, image deblurring, and machine learning.

2 The orthogonal randomized Bregman projection (ORBP) method

The Bregman distance $D_f^{x^*}(x, y)$ between x and y with respect to a strongly convex function f and a subgradient $x^* \in \partial f(x)$ is defined as

$$D_f^{x^*}(x, y) := f(y) - f(x) - \langle x^*, y - x\rangle\,,$$

where $\partial f(x) = \{x^* \in \mathbb{R}^n \mid f(y) \geq f(x) + \langle x^*, y - x\rangle, \text{for all } y \in \mathbb{R}^n\}$ denotes the subdifferential of f at x. The Bregman projection of x onto C with respect to f and $x^* \in \partial f(x)$ is given by the unique minimizer

$$\Pi_C^{x^*}(x) = \arg\min_{y\in C} D_f^{x^*}(x, y),$$

where $f^*(y) := \sup_{x\in\mathbb{R}^n}\{\langle y, x\rangle - f(x)\}$ represents the conjugate function of f at y.

By integrating the Bregman projection-type method [1, 2] with the orthogonalization strategy of the two-subspace Kaczmarz method [3], the orthogonal randomized Bregman projection (ORBP) method is proposed. In this method, Bregman projections are successively performed onto the randomly selected and orthogonalized hyperplanes.

The convergence result of the ORBP method is presented as follows.

Theorem 1. *Assume $b \in \mathrm{Range}(A)$, f is an α-strongly convex function such that its subdifferential mapping grows at most linearly and is calm at the unique solution $\hat{x}$. Also, suppose $\{\partial f(\hat{x}), \mathrm{Range}(A^T)\}$ is linearly regular. If the initialization $x_0 \in \mathbb{R}^n$ and $x_0^* \in \partial f(x_0) \cap \mathrm{Range}(A^T)$. Then, the iteration sequence $\{x_k\}_{k=0}^{\infty}$ of the ORBP method converges to the unique solution $\hat{x}$ of problem (1) with a linear rate in expectation, and satisfies*

$$\mathbb{E}[D_f^{x_k^*}(x_k, \hat{x})] \leq \left[\left(1 - \frac{1}{2(1-\delta^2)\tau_{\max}\nu(\hat{x})}\right)\left(1 - \frac{1}{2\|A\|_F^2\nu(\hat{x})}\right)\right]^k D_f^{x_0^*}(x_0, \hat{x})\,,$$

where $\mathbb{E}[\cdot]$ *denotes the expectation with respect to the algorithmic randomness,* $\tau_{\max} = \|A\|_F^2 - \min\limits_{1\leq i\leq m} \|a_i\|_2^2$*, and* $\delta = \min_{p\neq q} \frac{|a_p^T a_q|}{\|a_p\|_2\|a_q\|_2}$.

3 Numerical Experiments

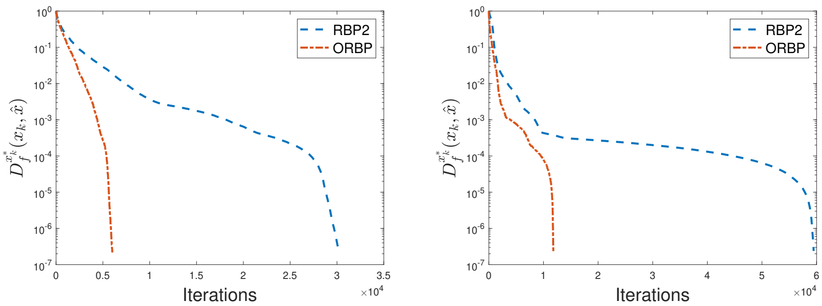
In the experiments, the ORBP method is compared with the two-step randomized Bregman projection (RBP2) method [4]. we consider the initialization $x_0 = x_0^* = 0$, and $f = \frac{1}{2}\|x\|_2^2+\lambda\|x\|_1$ with $\lambda = 2$. The stopping criterion is the relative solution error

$$\mathrm{ERR} = \frac{\|x_k - \hat{x}\|_2}{\|\hat{x}\|_2} < 10^{-3}$$

or the number of iterations exceed 200,000.

Example 1. *The test matrices are random Gaussian matrices with terms in [0.8,1]. The sparse solution* $\hat{x}$ *is randomly generated with the density 0.01 and the right-hand side* $b = A\hat{x}$*.*

The curves of the Bregman distance $D_f^{x_k^*}(x_k, \hat{x})$ versus the iteration steps of different coefficient matrices for ORBP and RBP2 methods are plotted in the following figures.



(a) Size: 4000×6000 (b) Size: 6000×4000

Figure1. Convergenve curves.

From Figure 1, it is seen that the ORBP method exhibits a faster convergence rate compared to the RBP2 method.

Reference

[1] D. A. Lorenz, F. Schöpfer, and S. Wenger. The linearized Bregman method via split feasibility problems: analysis and generalizations. SIAM Journal on Imaging Sciences, 7.2(2014), 1237– 1262.

[2] Z.-Y. Yuan, L. Zhang, H.-X. Wang, and H. Zhang. Adaptively sketched Bregman projection methods for linear systems. Inverse Problems, 38.6(2022), 065005.

[3] W.-T. Wu. On two-subspace randomized extended Kaczmarz method for solving large linear least-squares problems. Numerical Algorithms, 89.1(2022), 1–31.

[4] F. Schöpfer and D. A. Lorenz. Linear convergence of the randomized sparse Kaczmarz method. Mathematical Programming, 173(2019), 509–536.

Design and analysis of a predefined-time zeroing neural network model for solving the Stein tensor equation

Shu-Xin Miao[1, 2], Tomohiro Sogabe[2], Shao-Liang Zhang[2]
[1] Department of Mathematics, Northwest Normal University, Lanzhou 730070, P.R. China
[2] Department of Applied Physics, Nagoya University, Chikusa-ku, Nagoya 464–8603, Japan
e-mail : miaosx@nwnu.edu.cn

1 Introduction

The Stein tensor equation $\mathcal{X} - \mathcal{C} *_N \mathcal{X} *_N \mathcal{D} = \mathcal{F}$ plays a vital role in stability analysis and model reduction of dynamical systems [1], where $*_N$ denotes the Einstein product between tensors. In practical applications, the tensors involved are time-dependent. In this paper, we consider zeroing neural network (ZNN) model [2] for solving time-varying Stein tensor equation

$$\mathcal{X}(t) - \mathcal{C}(t) *_N \mathcal{X}(t) *_N \mathcal{D}(t) = \mathcal{F}(t), \quad 0 \le t < +\infty. \tag{1}$$

2 ZNN model for matrix form of (1)

It follows from [3] that (1) can be transformed into the matrix equation

$$X(t) - C(t)X(t)D(t) = F(t), \ \ 0 \le t < \infty. \tag{2}$$

Therefore, the solution to (1) can be obtained by solving equation (2).

Let $E(t) = X(t) - C(t)X(t)D(t) - F(t) = (e_{ij}(t))$ be the error function of (2), to derive $E(t)$ converges to zero, we adopt the design formula [4] $\dot{E}(t) = \frac{\mathrm{d}E(t)}{\mathrm{d}t} = -\frac{1}{T_c}G(E(t))$, where $T_c > 0$ is a predefined and tunable parameter. Then we have the predefined-time ZNN (PTZNN) model

$$\begin{aligned}\dot{X}(t) - C(t)\dot{X}(t)D(t) = \dot{C}(t)X(t)D(t) + C(t)X(t)\dot{D}(t) + \dot{F}(t)\\ - \frac{1}{T_c}G\left(X(t) - C(t)X(t)D(t) - F(t)\right)\end{aligned} \tag{3}$$

and its noise-perturbed PTZNN (NPTZNN) model

$$\begin{aligned}\dot{X}(t) - C(t)\dot{X}(t)D(t) = \dot{C}(t)X(t)D(t) + C(t)X(t)\dot{D}(t) + \dot{F}(t)\\ - \frac{1}{T_c}G\left(X(t) - C(t)X(t)D(t) - F(t)\right) + N(t)\end{aligned} \tag{4}$$

for solving (2). Here $N(t) = (n_{ij}(t))$ represents the noise, $G = (g(u))$ with $g(u) = \vartheta(\iota_1 \mid u \mid^p + \iota_2 \mid u \mid^q)^k \mathrm{sign}(u) + \iota_3 u + \iota_4 \mathrm{sign}(u)$, $\iota_1,\ \iota_2 > 0$, $\iota_3,\ \iota_4 \ge 0$, $0 < p < 1$ and $q > 1$, $k > 0$, $0 < kp < 1$, $kq > 1$ and $\vartheta = \frac{\Gamma\left(\frac{1-kp}{q-p}\right)\Gamma\left(\frac{kq-1}{q-p}\right)}{\iota_1^k \Gamma(k)(q-p)} \left(\frac{\iota_1}{\iota_2}\right)^{\frac{1-kp}{q-p}}$, $\Gamma(\cdot)$ denotes the Gamma function.

3 Theoretical results

Following theorem gives the convergence of the PTZNN model (3) for solving (2).

Theorem 1. *Assume that the PTZNN model* (3) *is applied to solve the time-varying equation* (2). *Then the state matrix* $X(t)$ *of the PTZNN model, starting from arbitrary original state matrix* $X(0)$, *will converge to the analytical solution* $X^*(t)$ *within the predefined time* T_c.

For the convergence of the NPTZNN model (4) with DBV noise $\mid n_{ij}(t) \mid \leq \varepsilon \mid e_{ij}(t) \mid$ and DBN noise $\mid n_{ij}(t) \mid \leq \varepsilon$, we have the following result, where $\varepsilon \in (0, +\infty)$.

Theorem 2. *Assume that the NPTZNN model with DBV (or DBN) noise is applied to solve the time-varying equation* (2)*. If $\varepsilon \leq \frac{\iota_3}{T_c}$ (or $\varepsilon \leq \frac{\iota_4}{T_c}$), then the state matrix $X(t)$ of the NPTZNN model, starting from arbitrary initial state matrix $X(0)$, will converge to the analytical solution $X^*(t)$ within the predefined time T_c.*

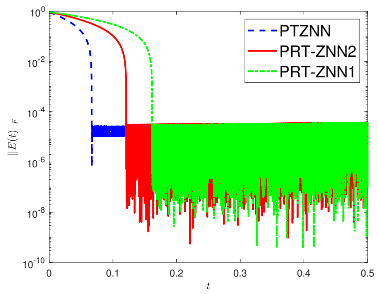
4 Numrical results

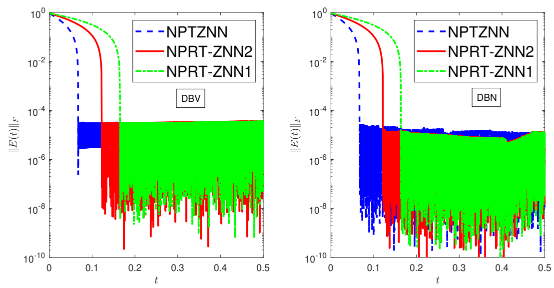
Example 3. *Let* $\mathcal{C}(t) = (\mathsf{c}_{i_1 i_2 j_1 j_2}(t))$, $\mathcal{D}(t) = (\mathsf{d}_{i_1 i_2 j_1 j_2}(t)) \in \mathcal{R}^{2\times2\times2\times2}$, *where*

$$\mathsf{c}_{i_1 i_2 j_1 j_2}(t) = \begin{cases} \frac{1}{2} + \frac{1}{2}\cos(t), & i_1 = j_1,\ i_2 = j_2, \\ 0, & \text{otherwise}, \end{cases} \quad \mathsf{d}_{i_1 i_2 j_1 j_2}(t) = \begin{cases} \frac{1}{3} + \frac{1}{3}\cos(t), & i_1 = j_1,\ i_2 = j_2, \\ 0, & \text{otherwise}. \end{cases}$$

The right hand side tensor $\mathcal{F}(t)$ is chosen such that the analytical solution is $\mathcal{X}^*(t) = (\mathsf{x}_{i_1 i_2 j_1 j_2}(t)) \in \mathcal{R}^{2\times2\times2\times2}$ with $\mathsf{x}_{i_1 i_2 j_1 j_2}(t) = \sin(t)$ for all $i_1,\ i_2,\ j_1,\ j_2 = 1, 2$.

Let $\iota_1 = \iota_3 = \iota_4 = 1$, $\iota_2 = 2$ $p = 0.5$, $q = 2$, $k = 1.5$, $\gamma = 1$ and $T_c = 0.5s$, by comparing PTZNN with PRT-ZNN1 and PRT-ZNN2 [5], and NPTZNN with NPRT-ZNN1 and NPRT-ZNN2 [5], the comparison results for $\|E(t)\|_F$ w.r.t. t are presented in the following figure.





参考文献

[1] B. Datta, Numerical Methods for Linear Control Systems, Academic Press, 2003.

[2] Y. Zhang, S. Ge, Design and analysis of a general recurrent neural network model for time-varying matrix inversion, IEEE Transactions on Neural Networks, 16 (6) (2005) 1477–1490.

[3] Q.-W. Wang, X. Xu, Iterative algorithms for solving some tensor equations, Linear and Multilinear Algebra 67 (7) (2019) 1325–1349.

[4] Y. Yue, L. Mi, C. Chen, Y. Yang, Predefined-time stability-based zeroing neural networks and their application in solving the Lyapunov equation, Neural Processing Letters 56 (2024) art.14, 15 pages.

[5] J. Dai, L. Jia, L. Xiao, Design and analysis of two prescribed-time and robust ZNN models with applications to time-variant Stein matrix equation, IEEE Transactions on Neural Networks and Learning Systems 32 (4) (2021) 1668–1677.

Extended sparse Kaczmarz method with surrogate hyperplane for sparse solutions to inconsistent linear systems

Ze Wang[1,2], Yu-Xin Ye[1,2],Jun-Feng Yin[2],Shao-Liang Zhang[1]
[1]Nagoya University, Japan [2]Tongji University, China
e-mail : math_wangze@tongji.edu.cn

1 Introduction

Consider the following combined optimization problem:

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{s.t.} \quad Ax = \hat{y}, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ is the noisy observation, and $\hat{y}$ denotes the projection of b onto the range of A, defined as $\hat{y} = \underset{y\in\mathcal{R}(A)}{\text{argmin}}\frac{1}{2}\|b-y\|_2^2$. The sparsity-promoting objective function is $f(x) = \lambda\|x\|_1 + \frac{1}{2}\|x\|_2^2$, which corresponds to the regularized basis pursuit problem. This optimization problem arises widely from many scientific computing and engineering applications fields, such as machine learning, compressed sensing, and image processing.

The sparse Kaczmarz method [1] solves the regularized Basis Pursuit problem, while the extended Kaczmarz method [2] addresses least squares problems. To approximate sparse solutions of inconsistent linear systems, a surrogate hyperplane sparse extended Kaczmarz framework is constructed for the combined optimization problem (1). Convergence theories are established, and numerical experiments confirm the method's effectiveness.

2 Sparse extended Kaczmarz method with surrogate hyperplane

Let a_i^T denote the ith row of A, and b_i the ith element of b. To enhance the computational efficiency of the extended Kaczmarz method, the residual $u_k = -A^T z_k$ is used to define the surrogate hyperplane $u_k^T A^T z = 0$ in the first stage, yielding z_{k+1} as an approximation of $b_{\mathcal{R}(A)^\perp}$. In the second stage, x_k is projected onto $\eta_k^T Ax = \eta_k^T(b - z_{k+1})$. By combining original hyperplanes via weight parameters, the surrogate hyperplane extended Kaczmarz method is established.

For undetermined weight vector η_k, the following theorem provides a general upper bound for the surrogate hyperplane sparse extended Kaczmarz method.

Theorem 2.1. *Assume that both $b_{\mathcal{R}(A)} \neq 0$ and $b_{\mathcal{R}(A)^\perp} \neq 0$. Let the initial vectors be $x_0 = x_0^* = 0$. Then, there exists a constant $c > 0$ such that the iterative sequence $\{x_k\}_{k=0}^\infty$ converges to the unique solution $\hat{x}$ of (1), it satisfies that*

$$\|x_k - \hat{x}\|_2^2 \leq \frac{c}{k}.$$

In particular, three specific strategies for generating surrogate hyperplanes are proposed.

- Randomized sparse extended Kaczmarz method (RSEK): where $\eta_k = e_i$ and i is selected based on the probability $p_i = \|a_i\|_2^2/\|A\|_F^2$,
- Residual-based surrogate hyperplane sparse extended Kaczmarz method (SEK-RSH): where $\eta_k = b - Ax_k$ during each iteration,

- Partial residual-based surrogate hyperplane sparse extended Kaczmarz method (SEK-PRSH): where $\eta_k = \sum_{i\in\tau_k}(b_i - a_i^T x_k - z_{i_{k+1}}^*)e_i$, with the index set determined by

$$\tau_k = \left\{ i \,\middle|\, |b_i - a_i^T x_k - z_{i_{k+1}}^*|^2 \geq \epsilon_k \|b - Ax_k - z_{k+1}^*\|_2^2 \|a_i\|_2^2 \right\},$$

and

$$\epsilon_k = \frac{1}{2\left\|b - Ax_k - z_{k+1}^*\right\|_2^2} \max_{1\leq i\leq m} \left\{ \frac{\left|b_i - a_i^T x_k - z_{i_{k+1}}^*\right|^2}{\|a_i\|_2^2} \right\} + \frac{1}{2\|A\|_F^2}.$$

The convergence and contraction factors of the proposed methods are established, with corresponding theorems and analyses given in our paper.

3 Numerical experiments

The surrogate hyperplane sparse extended Kaczmarz methods are compared with the residual-based surrogate hyperplane sparse Kaczmarz method (SK-RSH) [3] and the residual-based surrogate hyperplane extended Kaczmarz method (EK-RSH) [4], respectively.

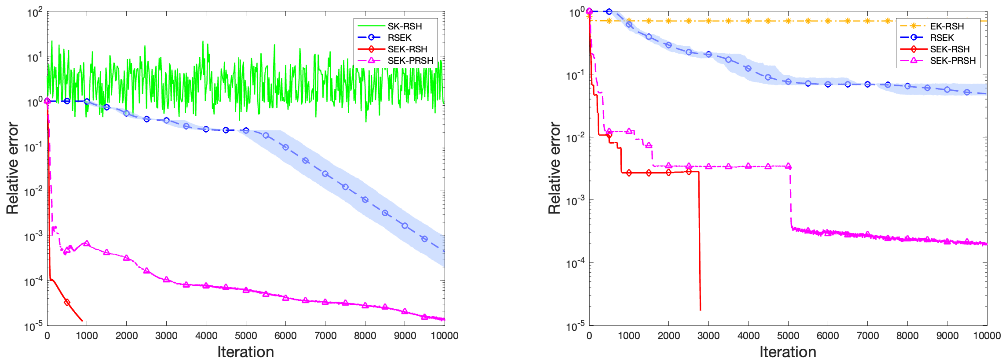


Figure1. Convergence curves of relative error versus iteration

From Figure 1, it is observed that the SEK-type methods converge to a sparse least squares solution, with the SEK-RSH method requiring the fewest iteration steps.

References

[1] D. A. Lorenz, F. Schöpfer, and S. Wenger. The linearized Bregman method via split feasibility problems: analysis and generalizations. SIAM Journal on Imaging Sciences, 7(2):1237–1262, 2014.

[2] A. Zouzias and N. M. Freris. Randomized extended Kaczmarz for solving least squares. SIAM Journal on Matrix Analysis and Applications, 34(2):773–793, 2013.

[3] Z. Wang, J.-F. Yin, and J.-C. Zhao. The sparse Kaczmarz method with surrogate hyperplane for the regularized basis pursuit problem. Journal of Computational and Applied Mathematics, 454:116182, 2025.

[4] K. Zhang, X.-X. Chen, and X.-L. Jiang. A residual-based surrogate hyperplane extended Kaczmarz algorithm for large least squares problems. Calcolo, 61(3), 2024.

任意インデックスの正方行列の特異系とランク落ち長方行列の最小二乗問題に対する NR-SSOR 右前処理 RRGMRES 法

NR-SSOR right preconditioned RRGMRES for square singular systems with arbitrary index and rank-deficient rectangular least squares problems

杉原 光太 (Sugihara Kota)[1], 速水 謙 (Hayami Ken)[2]
[1] 所属無し (No Affiliation), [2] 国立情報学研究所、総合研究大学院大学 (Professor Emeritus, National Institute of Informatics, and The Graduate University for Advanced Studies (SOKENDAI))
e-mail：kouta.sugihara@gmail.com

1 Abstract

GMRES is known to determine a least squares (LS) solution of $A\boldsymbol{x} = \boldsymbol{b}$ where $A \in \mathbf{R}^{n\times n}$ without breakdown for arbitrary $\boldsymbol{b} \in \mathbf{R}^n$, and initial iterate $\boldsymbol{x}_0 \in \mathbf{R}^n$ if and only if A is range-symmetric, i.e. $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(A)$ holds, where A may be singular and $\boldsymbol{b}$ may not be in the range space $\mathcal{R}(A)$ of A [1, 2]. On the other hand, when $\mathcal{R}(A^{\mathrm{T}}) \neq \mathcal{R}(A)$, including the case when the index of A is greater than or equal to 1, there exist $\boldsymbol{x}_0$ and $\boldsymbol{b}$ such that GMRES breaks down without giving a LS solution [2].

In this paper, to solve this difficulty, we propose applying the Range Restricted GMRES (RRGMRES)[3] to $ACA^{\mathrm{T}}\boldsymbol{z} = \boldsymbol{b}$, where $C \in \mathbf{R}^{n\times n}$ is symmetric positive definite (SPD). Here, we use RRGMRES instead of GMRES since RRGMRES is more stable compared to GMRES for inconsistent range-symmetric systems, where $A\boldsymbol{x} = \boldsymbol{b}$ is called inconsistent when $\boldsymbol{b} \notin \mathcal{R}(A)$. The proposed RRGMRES determines a least squares solution $\boldsymbol{x} = CA^{\mathrm{T}}\boldsymbol{z}$ of $A\boldsymbol{x} = \boldsymbol{b}$ without breakdown for arbitrary (singular) matrix $A \in \mathbf{R}^{n\times n}$ and $\boldsymbol{b}, \boldsymbol{x}_0 \in \mathbf{R}^n$ since ACA^{T} is SPD even if $\mathcal{R}(A^{\mathrm{T}}) \neq \mathcal{R}(A)$. In particular, we prove that $C \in \mathbf{R}^{n\times n}$ for NR-SSOR right preconditioner is SPD and propose the NR-SSOR right preconditioned RRGMRES, which also works efficiently for rectangular least squares problems $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ for $A \in \mathbf{R}^{m\times n}$ and arbitrary $\boldsymbol{b} \in \mathbf{R}^m$.

Numerical experiments demonstrate the validity of the proposed method.

2 Index of matrices, GP matrices and its applications

index(A) of $A \in \mathbf{R}^{n\times n}$ denotes the smallest nonnegative integer i such that rank(A^i) = rank(A^{i+1}). If $\mathcal{R}(A) \cap \mathcal{N}(A) = \mathbf{0}$, A is called a GP (group) matrix. If A is singular, A is GP if and only if index$(A) = 1$. GP matrices arise, for example, in the analysis of ergodic homogeneous finite Markov chains.

3 Right preconditioned GMRES and RRGMRES for arbitrary singular systems

RRGMRES determines the minimum-norm least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ without breakdown for arbitrary $\boldsymbol{b} \in \mathbf{R}^n$, and initial iterate $\boldsymbol{x}_0 \in \mathcal{R}(A)$ where A may be singular if $\mathcal{R}(A) =$

$\mathcal{R}(A^{\mathrm{T}})$ [3]. Furthermore, using $B = CA^{\mathrm{T}}$ as a right preconditioner where C is SPD, the right preconditioned GMRES and RRGMRES determine a least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ without breakdown for arbitrary $\boldsymbol{b} \in \mathbf{R}^n$, and initial iterate $\boldsymbol{x}_0 \in \mathbf{R}^n$ [4, 5].

4 Convergence theory of NR-SSOR right preconditioned RRGMRES

NR-SSOR is mathematically equivalent to SSOR applied to $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$.

Then, consider the first kind normal equations $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$ where $A \in \mathbf{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbf{R}^m$. Let $A^{\mathrm{T}}A = M - N \in \mathbf{R}^{n\times n}$ where M is nonsingular, and $H = M^{-1}N = I - M^{-1}A^{\mathrm{T}}A$.

Now consider the stationary iterative method applied to $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$ with $\boldsymbol{x}^{(0)} = \mathbf{0}$. Then, $\boldsymbol{x}^{(\ell)} = H\boldsymbol{x}^{(\ell-1)} + M^{-1}A^{\mathrm{T}}\boldsymbol{b} = \sum_{i=0}^{\ell-1} H^i M^{-1}A^{\mathrm{T}}\boldsymbol{b}$. Therefore, let $B^{(\ell)} = C^{(\ell)}A^{\mathrm{T}}$ be the preconditioner where $C^{(\ell)} = \sum_{i=0}^{\ell-1} H^i M^{-1}$.

For NR-SSOR, $M = \omega^{-1}(2-\omega)^{-1}(D+\omega L)D^{-1}(D+\omega L^{\mathrm{T}})$, where $A^{\mathrm{T}}A = L + D + L^{\mathrm{T}}$, L is a strictly lower triangular matrix, D is a diagonal matrix, and ω is the relaxation parameter. Thus, M is symmetric. Then, the iteration matrix $H = M^{-1}N$ for NR-SSOR with $0 < \omega < 2$ is semiconvergent. M is nonsingular if A has no zero columns. Furthermore, M is positive definite if $0 < \omega < 2$.

Using NR-SSOR right preconditioning, the preconditioner is $B = CA^{\mathrm{T}}$ where C is symmetric positive definite according to the following theorem.

Theorem 1. *([5], Theorem 2) Assume that A has no zero columns and $0 < \omega < 2$ holds. Then,* $C^{(\ell)} = \displaystyle\sum_{i=0}^{\ell-1} H^i M^{-1}$ *is positive definite.*

Therefore, the right preconditioned GMRES and RRGMRES using the NR-SSOR right preconditioner $B = CA^{\mathrm{T}}$ determine a least squares solution of arbitrary singular systems and rank-deficient rectangular least squares problems without breakdown for arbitrary $\boldsymbol{b} \in \mathbf{R}^m$, and initial iterate $\boldsymbol{x}_0 \in \mathbf{R}^n$. We will report numerical results for GP, Index 2 inconsistent problems and rectangular least squares problems in our presentation.

参考文献

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

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

[3] D. Calvetti, B. Lewis, and L. Reichel, *GMRES-type methods for inconsistent systems*, Linear Algebra Appl., 316 (2000), pp. 157–169.

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

[5] K. Sugihara and K. Hayami, *NR-SSOR right preconditioned RRGMRES for arbitrary singular systems and least squares problems*, `https://arxiv.org/abs/2504.09891` (2025).