SIAM J. MATRIX ANAL. APPL.
Vol. 31, No. 5, pp. 2400–2430

GMRES METHODS FOR LEAST SQUARES PROBLEMS*

KEN HAYAMI†, JUN-FENG YIN‡, AND TOKUSHI ITO§

Abstract. The standard iterative method for solving large sparse least squares problems $\min \|\boldsymbol{b}-A\boldsymbol{x}\|_2$, $A \in \mathbf{R}^{m\times n}$, is the CGLS method, or its stabilized version, LSQR, which is mathematically equivalent to applying the conjugate gradient method to the normal equation $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$. We consider alternative methods using a matrix $B \in \mathbf{R}^{n\times m}$ and applying the generalized minimal residual (GMRES) method to $\min \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$ or $\min \|B\boldsymbol{b} - BA\boldsymbol{x}\|_2$. We give a sufficient condition concerning B for the GMRES methods to give a least squares solution without breakdown for arbitrary $\boldsymbol{b}$, for overdetermined, underdetermined, and possibly rank-deficient problems. We then give a convergence analysis of the GMRES methods as well as the CGLS method. Then, we propose using the robust incomplete factorization (RIF) for B. Finally, we show by numerical experiments on overdetermined and underdetermined problems that, for ill-conditioned problems, the GMRES methods with RIF give least squares solutions faster than the CGLS and LSQR methods with RIF, and are similar in performance to the reorthogonalized CGLS with RIF.

Key words. least squares problems, iterative method, Krylov subspace method, GMRES method, CGLS method, LSQR method, robust incomplete factorization

AMS subject classifications. Primary, 65F10, 65F20, 65F50; Secondary, 15A06, 15A09

DOI. 10.1137/070696313

1. Introduction.

Consider the least squares problem

$$\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2, \tag{1.1}$$

where $A \in \mathbf{R}^{m\times n}$, and $m \geq n$ or $m < n$. We also allow the rank-deficient case when the equality in $\operatorname{rank} A \leq \min(m, n)$ does not hold.

The least squares problem (1.1) is equivalent to the normal equation

$$A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}, \tag{1.2}$$

for which a solution always exists. For $m < n$,

$$AA^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{b}, \quad \boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{y} \tag{1.3}$$

gives the minimum norm solution of (1.1), if one exists.

The standard direct method for solving the least squares problem (1.1) (where $m \geq n$ and A has full rank) is to use the QR decomposition: $A = QR$, where $Q \in \mathbf{R}^{m\times n}$ is an orthogonal matrix and $R \in \mathbf{R}^{n\times n}$ is an upper triangular matrix, which

*Received by the editors July 5, 2007; accepted for publication (in revised form) by M. Benzi May 11, 2010; published electronically August 12, 2010. This research was supported by the Grants-in-Aid for Scientific Research (C) of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

http://www.siam.org/journals/simax/31-5/69631.html

†National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan, and Department of Informatics, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies, Sokendai, Hayama, Japan (hayami@nii.ac.jp).

‡National Institute of Informatics, Tokyo, Japan. Current address: Department of Mathematics, Tongji University, 1239 Siping Road, Shanghai, 200092, People's Republic of China (yinjf@tongji.edu.cn).

§Business Design Laboratory Co. Ltd., 3-38-11, Yoyogi, Shibuya-ku, Tokyo 151-0053, Japan (ito047@gmail.com).

can be obtained using the Householder or Givens transformations or by the modified Gram–Schmidt (MGS) method. Then, (1.2) is transformed as $R^{\mathrm{T}}R\boldsymbol{x} = R^{\mathrm{T}}Q^{\mathrm{T}}\boldsymbol{b}$. If $\operatorname{rank} A = n$, R is nonsingular, so that $R\boldsymbol{x} = Q^{\mathrm{T}}\boldsymbol{b}$, and back substitution gives the least squares solution $\boldsymbol{x}$ of (1.1). When A is large and sparse, techniques are used to save memory and computation time [4].

However, for very large and sparse problems, iterative methods become necessary, among which the (preconditioned) conjugate gradient least squares (CGLS) method [4] or its stabilized version LSQR [23] are most commonly used. This is based on the observation that $A^{\mathrm{T}}A$ is symmetric, and also positive definite if $\operatorname{rank} A = n$, so that it is natural to apply a form of the conjugate gradient (CG) method to (1.2) [17, 4]. As shown in the numerical experiments, the convergence of CGLS deteriorates for ill-conditioned problems, so that reorthogonalization becomes necessary. Below, we give the reorthogonalized CGLS(CGNR)(l) method. See [20] for work on Lanczos (Golub–Kahan) bidiagonalization with partial reorthogonalization.

Method 1.1 (CGLS(CGNR)(l) method).

1. $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\boldsymbol{p}_0 = \boldsymbol{s}_0 = A^T\boldsymbol{r}_0$, $\gamma_0 = \|\boldsymbol{s}_0\|_2{}^2$
2. *for* $k = 0, 1, \ldots$ *until convergence, do:*
3. $\quad \boldsymbol{q}_k = A\boldsymbol{p}_k$
4. $\quad \alpha_k = \gamma_k / \|\boldsymbol{q}_k\|_2{}^2$
5. $\quad \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k$
6. $\quad \boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \boldsymbol{q}_k$
7. $\quad \boldsymbol{s}_{k+1} = A^T\boldsymbol{r}_{k+1}$
8. $\quad$ *for* $i = k, \ldots, \max(0, k+1-l)$
9. $\quad\quad \boldsymbol{s}_{k+1} = \boldsymbol{s}_{k+1} - \{(\boldsymbol{s}_{k+1}, \boldsymbol{s}_i)/\gamma_i\}\boldsymbol{s}_i$
10. $\quad$ *endfor*
11. $\quad \gamma_{k+1} = \|\boldsymbol{s}_{k+1}\|_2{}^2$
12. $\quad \beta_k = \gamma_{k+1}/\gamma_k$
13. $\quad \boldsymbol{p}_{k+1} = \boldsymbol{s}_{k+1} + \beta_k \boldsymbol{p}_k$
14. *endfor*

Note that the standard CGLS is given by deleting lines 8 to 10, and the fully reorthogonalized CGLS is given by setting $l = \infty$. In [5], different versions of the CGLS and the LSQR method are compared.

For underdetermined problems ($m < n$), the reorthogonalized CGNE(l) method is given as follows.

Method 1.2 (CGNE(l) method).

$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\boldsymbol{p}_0 = A^T\boldsymbol{r}_0$, $\gamma_0 = \|\boldsymbol{r}_0\|_2{}^2$
for $k = 0, 1, \ldots$ *until convergence, do:*
$\quad \alpha_k = \gamma_k / \|\boldsymbol{p}_k\|_2{}^2$
$\quad \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k$
$\quad \boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k A\boldsymbol{p}_k$
$\quad$ *for* $i = k, \ldots, \max(0, k+1-l)$
$\quad\quad \boldsymbol{r}_{k+1} = \boldsymbol{r}_{k+1} - \{(\boldsymbol{r}_{k+1}, \boldsymbol{r}_i)/\gamma_i\}\boldsymbol{r}_i$
$\quad$ *endfor*
$\quad \gamma_{k+1} = \|\boldsymbol{r}_{k+1}\|_2{}^2$
$\quad \beta_k = \gamma_{k+1}/\gamma_k$
$\quad \boldsymbol{p}_{k+1} = A^T\boldsymbol{r}_{k+1} + \beta_k \boldsymbol{p}_k$
endfor

Here, let $A^\dagger$ be the Moore–Penrose inverse of $A \in \mathbf{R}^{m\times n}$, $\operatorname{rank} A = r$, and let σ_1 and σ_r be the largest and smallest (nonzero) singular value of A, respectively.

Then, the convergence speed of CGLS is known to depend on the condition number $\kappa(A) := \|A\|_2 \|A^\dagger\|_2 = \sigma_1/\sigma_r$ of A [4]. Hence, the convergence of CGLS may be slow for ill-conditioned problems, so that preconditioning becomes necessary.

A simple preconditioning is accomplished by diagonal scaling using the diagonal elements of $A^{\mathrm{T}}A$. More sophisticated preconditionings are, for example, the incomplete Cholesky decomposition [21], incomplete QR decompositions using incomplete modified Gram–Schmidt methods [19, 26, 32] or incomplete Givens orthogonalizations [1, 24], and the robust incomplete factorization (RIF) [3].

For instance, the incomplete QR decomposition $A \sim QR$ can be used to precondition the normal equation (1.2) as

$$\tilde{A}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}}, \tag{1.4}$$

where $\tilde{A} = R^{-\mathrm{T}}A^{\mathrm{T}}AR^{-1}$, $\tilde{\boldsymbol{x}} = R\boldsymbol{x}$, $\tilde{\boldsymbol{b}} = R^{-\mathrm{T}}A^{\mathrm{T}}\boldsymbol{b}$, after which CG is applied to (1.4).

Even with preconditioning, the convergence behavior may deteriorate for highly ill-conditioned problems because rounding errors cause loss of orthogonality, as observed in our numerical experiments.

On the other hand, Zhang and Oyanagi [33] proposed applying the Orthomin(k) method directly to the least squares problem (1.1) instead of treating the normal equation (1.2). This was done by introducing a mapping matrix $B \in \mathbf{R}^{n\times m}$ to transform the problem to a system with a square matrix $AB \in \mathbf{R}^{m\times m}$, and then applying the Krylov subspace method Orthomin(k) to this nonsymmetric system.

In [18, 14, 15], we further extended their method by applying the GMRES method instead of the Orthomin(k) method, and also introduced an alternative method of treating the system with a matrix $BA \in \mathbf{R}^{n\times n}$, and gave a sufficient condition concerning B for the methods to give the least squares solution without breakdown for overdetermined full-rank systems.

Similar methods were also proposed by Vuik, Sevink, and Herman [31]. Related ideas can also be found in Dahlquist, Sjöberg, and Svensson [11], and Tanabe [30] for linear stationary iterations.

In this paper, we further extend the analysis to rank-deficient as well as underdetermined systems, and derive a sufficient condition concerning B for the general case. We also give some convergence analysis for the GMRES methods as well as the CGLS method. Then, we propose using the robust incomplete factorization of Benzi and Tůma [3] for B. Finally, we give numerical experiment results for overdetermined and underdetermined systems, showing that, for ill-conditioned problems, the GMRES-based methods using RIF for B are faster than the CGLS, CGNE, and LSQR methods with RIF, and are similar in performance to the reorthogonalized CGLS with RIF.

The rest of the paper is organized as follows. In section 2, we briefly review the GMRES method. In section 3, we present the AB- and BA-GMRES methods for over- and underdetermined least squares problems, and give a sufficient condition for B. In section 4, we discuss the properties of the eigenvalues of AB and BA, and give some convergence analysis of the GMRES methods as well as the preconditioned CGLS method. In section 5, we propose using RIF for B. In section 6, numerical experiment results are presented. Section 7 concludes the paper.

2. The GMRES method. The generalized minimal residual (GMRES) method [28] is an efficient and robust Krylov subspace iterative method for solving systems of linear equations $A\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbf{R}^{n\times n}$ is nonsingular and nonsymmetric. Usually, the method is implemented with restarts in order to reduce storage and computation

time, as in the GMRES(k) method below. ($k = \infty$ corresponds to the original full GMRES method.)

Method 2.1 (the GMRES(k) method).

Choose $\boldsymbol{x}_0$.
$*$ $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$
$\boldsymbol{v}_1 = \boldsymbol{r}_0/\|\boldsymbol{r}_0\|_2$
for $i = 1, 2, \ldots, k$
$\boldsymbol{w}_i = A\boldsymbol{v}_i$
for $j = 1, 2, \ldots, i$
$h_{j,i} = (\boldsymbol{w}_i, \boldsymbol{v}_j)$
$\boldsymbol{w}_i = \boldsymbol{w}_i - h_{j,i}\boldsymbol{v}_j$
endfor
$h_{i+1,i} = \|\boldsymbol{w}_i\|_2$
$\boldsymbol{v}_{i+1} = \boldsymbol{w}_i/h_{i+1,i}$
Find $\boldsymbol{y}_i \in \mathbf{R}^i$ *to minimize* $\|\boldsymbol{r}_i\|_2 = \left\| \|\boldsymbol{r}_0\|_2 \boldsymbol{e}_1 - \bar{H}_i \boldsymbol{y} \right\|_2$.
if $\|\boldsymbol{r}_i\|_2 < \epsilon$ *then*
$\boldsymbol{x}_i = \boldsymbol{x}_0 + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i$
stop
endif
endfor
$\boldsymbol{x}_0 = \boldsymbol{x}_k$
go to $*$

Here, $\bar{H}_i = (h_{pq}) \in \mathbf{R}^{(i+1)\times i}$ and $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbf{R}^{i+1}$. The method is designed to minimize the L$_2$-norm of the residual $\|\boldsymbol{r}_k\|_2$ for all $\boldsymbol{x}_k = \boldsymbol{x}_0 + \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$, where $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle = \langle \boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0 \rangle$ denotes the vector space spanned by the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$, and $(\boldsymbol{v}_i, \boldsymbol{v}_j) = \delta_{ij}$. The method is said to break down when $h_{i+1,i} = 0$.

The above is the modified Gram–Schmidt (MGS) version of the GMRES method, which is known to be backward stable [22].

In the following, we assume exact arithmetic.

When A is nonsingular, GMRES gives the exact solution for all $\boldsymbol{b} \in \mathbf{R}^n$ and $\boldsymbol{x}_0 \in \mathbf{R}^n$ within n steps [28]. When A is singular, the following holds [6] (see also [16]), where $\mathcal{R}(A)$ denotes the range space of A.

THEOREM 2.1. *Let* $A \in \mathbf{R}^{n\times n}$. *GMRES gives a solution to* $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ *without breakdown for arbitrary* $\boldsymbol{b} \in \mathbf{R}^n$ *and* $\boldsymbol{x}_0 \in \mathbf{R}^n$ *if and only if* $\mathcal{R}(A) = \mathcal{R}(A^{\mathrm{T}})$.

3. Solution of least squares problems using the GMRES method. Consider applying GMRES directly to the least squares problem (1.1). This would require multiplying $A \in \mathbf{R}^{m\times n}$ and the residual vector $\boldsymbol{r}_0 \in \mathbf{R}^m$, which is not feasible if $m \neq n$. In the following, we give two methods for overcoming this difficulty using a matrix $B \in \mathbf{R}^{n\times m}$.

3.1. The AB-GMRES method. The first method is to use the Krylov subspace $\mathcal{K}_i(AB, \boldsymbol{r}_0) = \langle\, \boldsymbol{r}_0, AB\boldsymbol{r}_0, \ldots, (AB)^{i-1}\boldsymbol{r}_0 \,\rangle$ in $\mathbf{R}^m$ generated by $AB \in \mathbf{R}^{m\times m}$, as in [33], and to solve the least squares problem $\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$ using GMRES.

First, we give a theoretical justification for doing so. In the following, let $\mathcal{N}(A)$ denote the null space of A, and let $V^{\perp}$ denote the orthogonal complement of subspace V.

THEOREM 3.1.

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

holds for all $\boldsymbol{b} \in \mathbf{R}^m$ *if and only if* $\mathcal{R}(A) = \mathcal{R}(AB)$.

Proof. The sufficiency is obvious. The necessity is shown as follows:

$$\begin{aligned}
&\mathcal{R}(A) \neq \mathcal{R}(AB) \Longrightarrow \mathcal{R}(A) \supset \mathcal{R}(AB) \Longrightarrow {}^{\exists}\tilde{\boldsymbol{b}} \in \mathcal{R}(A) \backslash \mathcal{R}(AB) \\
&\Longleftrightarrow {}^{\exists}\tilde{\boldsymbol{x}} \in \mathbf{R}^n; \quad \tilde{\boldsymbol{b}} = A\tilde{\boldsymbol{x}}, \quad \tilde{\boldsymbol{b}} \neq AB\boldsymbol{z} \quad {}^{\forall}\boldsymbol{z} \in \mathbf{R}^m \\
&\Longleftrightarrow 0 = \min_{\boldsymbol{x} \in \mathbf{R}^n} \|\tilde{\boldsymbol{b}} - A\boldsymbol{x}\|_2 < \min_{\boldsymbol{z} \in \mathbf{R}^m} \|\tilde{\boldsymbol{b}} - AB\boldsymbol{z}\|_2. \qquad \square
\end{aligned}$$

LEMMA 3.2. $\mathcal{R}(AA^{\mathrm{T}}) = \mathcal{R}(A)$.

Proof.

$$\begin{aligned}
\mathcal{R}(AA^{\mathrm{T}}) &= \{A\boldsymbol{x} \,|\, \boldsymbol{x} \in \mathcal{R}(A^{\mathrm{T}})\} = \left\{A\boldsymbol{x} \,|\, \boldsymbol{x} \in \mathcal{N}(A)^{\perp}\right\} \\
&= \left\{A\boldsymbol{x} \,|\, \boldsymbol{x} \in \mathcal{N}(A)^{\perp} \cup \mathcal{N}(A)\right\} = \mathcal{R}(A). \qquad \square
\end{aligned}$$

This gives the following.

LEMMA 3.3. $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B) \Longrightarrow \mathcal{R}(A) = \mathcal{R}(AB)$.

For instance, if $\operatorname{rank} A = \operatorname{rank} B = n$, then $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B) = \mathbf{R}^n$ holds, and hence $\mathcal{R}(A) = \mathcal{R}(AB)$ holds.

Note also the following.

LEMMA 3.4. *Let* $B = CA^{\mathrm{T}}$, *where* C *is symmetric positive definite. Then,* $\mathcal{R}(A) = \mathcal{R}(AB)$ *holds.*

Proof. $AB = ACA^{\mathrm{T}} = \tilde{A}\tilde{A}^{\mathrm{T}}$, where $\tilde{A} = AC^{\frac{1}{2}}$. Hence, $\mathcal{R}(A) = \mathcal{R}(\tilde{A}) = \mathcal{R}(\tilde{A}\tilde{A}^{\mathrm{T}}) = \mathcal{R}(AB)$. $\square$

LEMMA 3.5. *Let* $B = A^{\mathrm{T}}C$, *where* C *is nonsingular. Then,* $\mathcal{R}(AB) = \mathcal{R}(A)$ *holds.*

Proof. $\mathcal{R}(AB) = \mathcal{R}(AA^{\mathrm{T}}C) = \mathcal{R}(AA^{\mathrm{T}}) = \mathcal{R}(A)$. $\square$

Thus, assume $\mathcal{R}(A) = \mathcal{R}(AB)$ holds. Then, for arbitrary $\boldsymbol{x}_0 \in \mathbf{R}^n$, there exists a $\boldsymbol{z}_0 \in \mathbf{R}^m$ such that $A\boldsymbol{x}_0 = AB\boldsymbol{z}_0$, and $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0 = \boldsymbol{r}_0 - AB\boldsymbol{z}_0$. Hence, consider solving the least squares problem

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

using GMRES(k), letting the initial approximate solution be $\boldsymbol{z} = \boldsymbol{z}_0$ such that $AB\boldsymbol{z}_0 = A\boldsymbol{x}_0$. Then, we have the following algorithm.

Method 3.1 (the AB-GMRES(k) method).

Choose $\boldsymbol{x}_0 \quad (A\boldsymbol{x}_0 = AB\boldsymbol{z}_0)$.

$*$ $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0 \quad (= \boldsymbol{b} - AB\boldsymbol{z}_0)$

$\boldsymbol{v}_1 = \boldsymbol{r}_0 / \|\boldsymbol{r}_0\|_2$

for $i = 1, 2, \ldots, k$

  $\boldsymbol{w}_i = AB\boldsymbol{v}_i$

  for $j = 1, 2, \ldots, i$

    $h_{j,i} = (\boldsymbol{w}_i, \boldsymbol{v}_j)$

    $\boldsymbol{w}_i = \boldsymbol{w}_i - h_{j,i}\boldsymbol{v}_j$

  endfor

  $h_{i+1,i} = \|\boldsymbol{w}_i\|_2$

  $\boldsymbol{v}_{i+1} = \boldsymbol{w}_i / h_{i+1,i}$

  Find $\boldsymbol{y}_i \in \mathbf{R}^i$ *to minimize* $\|\boldsymbol{r}_i\|_2 = \left\| \|\boldsymbol{r}_0\|_2 \boldsymbol{e}_1 - \bar{H}_i \boldsymbol{y} \right\|_2$.

  $\boldsymbol{x}_i = \boldsymbol{x}_0 + B[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i \quad (\Longleftrightarrow \boldsymbol{z}_i = \boldsymbol{z}_0 + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i)$

  $\boldsymbol{r}_i = \boldsymbol{b} - A\boldsymbol{x}_i$

  if $\|A^{\mathrm{T}}\boldsymbol{r}_i\|_2 < \epsilon$ *stop*

endfor
$\boldsymbol{x}_0 = \boldsymbol{x}_k \quad (\Longleftrightarrow \boldsymbol{z}_0 = \boldsymbol{z}_k)$
go to $*$

Note here that the convergence is assessed by explicitly computing $\|A^{\mathrm{T}}\boldsymbol{r}_i\|_2$ since $\|\boldsymbol{r}_i\|_2$ does not necessarily converge to 0 for the general inconsistent case ($\boldsymbol{b} \notin \mathcal{R}(A)$). The following holds.

THEOREM 3.6. *If* $\mathcal{R}(A^T) = \mathcal{R}(B)$, *then* $\mathcal{R}(AB) = \mathcal{R}(B^TA^T) \iff \mathcal{R}(A) = \mathcal{R}(B^T)$.

Proof. If $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$, Lemma 3.2 gives $\mathcal{R}(AB) = \mathcal{R}(AA^{\mathrm{T}}) = \mathcal{R}(A)$ and $\mathcal{R}(B^{\mathrm{T}}A^{\mathrm{T}}) = \mathcal{R}(B^{\mathrm{T}}B) = \mathcal{R}(B^{\mathrm{T}})$. ∎

Let AB-GMRES be the AB-GMRES(k) method with $k = \infty$ (no restarts). Then, Theorems 2.1 and 3.6 give the following.

THEOREM 3.7. *If* $\mathcal{R}(A^T) = \mathcal{R}(B)$ *holds, then AB-GMRES determines a least squares solution of* $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ *for all* $\boldsymbol{b} \in \mathbf{R}^m$ *and for all* $\boldsymbol{x}_0 \in \mathbf{R}^n$ *without breakdown if and only if* $\mathcal{R}(A) = \mathcal{R}(B^T)$.

As a corollary, we have the following sufficient condition.

COROLLARY 3.8. *If* $\mathcal{R}(A^T) = \mathcal{R}(B)$ *and* $\mathcal{R}(A) = \mathcal{R}(B^T)$, *then AB-GMRES determines a least squares solution of* $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ *for all* $\boldsymbol{b} \in \mathbf{R}^m$ *and for all* $\boldsymbol{x}_0 \in \mathbf{R}^n$ *without breakdown.*

Note that the condition $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ is satisfied if B can be expressed as

$$B = CA^{\mathrm{T}}, \tag{3.1}$$

where C is a nonsingular matrix.

Note also the following.

LEMMA 3.9. *If* $B = CA^T$, *where* $C = C^T$, *then* $AB = B^TA^T$.

Proof. $B^{\mathrm{T}}A^{\mathrm{T}} = AC^{\mathrm{T}}A^{\mathrm{T}} = ACA^{\mathrm{T}} = AB$. ∎

Hence, from Lemma 3.4 we have the following.

LEMMA 3.10. *If* $B = CA^T$, *where* C *is symmetric positive definite, then* $\mathcal{R}(A) = \mathcal{R}(AB)$ *and* $AB = B^TA^T$ *hold.*

Hence, we are fully justified in using AB-GMRES with $B = CA^{\mathrm{T}}$, where C is symmetric positive definite, even when $\operatorname{rank} A < \min(m, n)$.

We note here that Calvetti, Lewis, and Reichel [7] proposed a related method for solving overdetermined ($m \geq n$) least squares problems using GMRES. Their method is to append $(m - n)$ zero column vectors to the right-hand side of the matrix A, to obtain a square singular matrix $\tilde{A} = [A, 0] \in \mathbf{R}^{m\times m}$, and apply GMRES to

$$\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - \tilde{A}\boldsymbol{z}\|_2 \left(= \min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2\right).$$

This corresponds to a special case of AB-GMRES with $B = [\,\mathrm{I}_n, 0\,] \in \mathbf{R}^{n\times m}$, where $\mathrm{I}_n \in \mathbf{R}^{n\times n}$ is an identity matrix, i.e., $\tilde{A} = AB$. In this case, $\mathcal{R}(AB) = \mathcal{R}(\tilde{A}) = \mathcal{R}(A)$ holds, but $\mathcal{R}(\tilde{A}) = \mathcal{R}(\tilde{A}^{\mathrm{T}})$ does not necessarily hold. Hence, their method may break down before giving a least squares solution. In fact, in [25], Reichel and Ye propose a breakdown-free GMRES method to circumvent this difficulty.

3.2. The BA-GMRES method. The other alternative is to use the same matrix $B \in \mathbf{R}^{n\times m}$ to map the initial residual vector $\boldsymbol{r}_0 \in \mathbf{R}^m$ to $\tilde{\boldsymbol{r}}_0 = B\boldsymbol{r}_0 \in \mathbf{R}^n$, and then to construct the Krylov subspace $\mathcal{K}_i(BA, \tilde{\boldsymbol{r}}_0) = \langle \tilde{\boldsymbol{r}}_0, BA\tilde{\boldsymbol{r}}_0, \ldots, (BA)^{i-1}\tilde{\boldsymbol{r}}_0\rangle$ in $\mathbf{R}^n$ and to solve the least squares problem $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|B\boldsymbol{b} - BA\boldsymbol{x}\|_2$ using GMRES.

First, we give a theoretical justification for doing so.

THEOREM 3.11. $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ *and* $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|B\boldsymbol{b}-BA\boldsymbol{x}\|_2$ *are equivalent for all* $\boldsymbol{b}\in\mathbf{R}^m$ *if and only if* $\mathcal{R}(A)=\mathcal{R}(B^TBA)$.

Proof. Note the following:

$$\|\boldsymbol{b}-A\boldsymbol{x}^*\|_2=\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2\Longleftrightarrow A^{\mathrm{T}}(\boldsymbol{b}-A\boldsymbol{x}^*)=\boldsymbol{0},$$

$$\begin{aligned}\|B\boldsymbol{b}-BA\boldsymbol{x}^*\|_2=\min_{\boldsymbol{x}\in\mathbf{R}^n}\|B\boldsymbol{b}-BA\boldsymbol{x}\|_2 &\iff (BA)^{\mathrm{T}}B(\boldsymbol{b}-A\boldsymbol{x}^*)=\boldsymbol{0}\\ &\iff A^{\mathrm{T}}B^{\mathrm{T}}B(\boldsymbol{b}-A\boldsymbol{x}^*)=\boldsymbol{0}\,.\end{aligned}$$

Then, note that $A^{\mathrm{T}}(\boldsymbol{b}-A\boldsymbol{x}^*)=\boldsymbol{0}$ is equivalent to $A^{\mathrm{T}}B^{\mathrm{T}}B(\boldsymbol{b}-A\boldsymbol{x}^*)=\boldsymbol{0}$ for all $\boldsymbol{b}\in\mathbf{R}^m$, if and only if $\mathcal{N}(A^{\mathrm{T}})=\mathcal{N}(A^{\mathrm{T}}B^{\mathrm{T}}B)$, which is equivalent to $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}}BA)$. ∎

LEMMA 3.12. $\mathcal{R}(A)=\mathcal{R}(B^T)\Longrightarrow\mathcal{R}(BA)=\mathcal{R}(B)$.

Proof. $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})\Longrightarrow\mathcal{R}(BA)=\mathcal{R}(BB^{\mathrm{T}})=\mathcal{R}(B)$ from Lemma 3.2. ∎

LEMMA 3.13. $\mathcal{R}(BA)=\mathcal{R}(B)\Longrightarrow\mathcal{R}(B^TBA)=\mathcal{R}(B^T)$.

Proof. $\mathcal{R}(BA)=\mathcal{R}(B)\Longrightarrow\mathcal{R}(B^{\mathrm{T}}BA)=\mathcal{R}(B^{\mathrm{T}}B)=\mathcal{R}(B^{\mathrm{T}})$. ∎

LEMMA 3.14. $\mathcal{R}(A)=\mathcal{R}(B^T)\Longrightarrow\mathcal{R}(A)=\mathcal{R}(B^TBA)$.

Proof. From Lemmas 3.12 and 3.13, $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})\Longrightarrow\mathcal{R}(BA)=\mathcal{R}(B)\Longrightarrow\mathcal{R}(B^{\mathrm{T}}BA)=\mathcal{R}(B^{\mathrm{T}})=\mathcal{R}(A)$. ∎

The following theorem also holds.[1]

THEOREM 3.15. *For all* $\boldsymbol{b}\in\mathbf{R}^m$*, the equation* $BA\boldsymbol{x}=B\boldsymbol{b}$ *has a solution, and the solution attains* $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ *if and only if* $\mathcal{R}(A)=\mathcal{R}(B^T)$.

Proof. ($\Longrightarrow$) First, for all $\boldsymbol{b}\in\mathbf{R}^m$, $BA\boldsymbol{x}=B\boldsymbol{b}$ has a solution if and only if

$$\mathcal{R}(BA)=\mathcal{R}(B). \tag{3.2}$$

Next, let $BA\boldsymbol{x}=B\boldsymbol{b}$ have a solution, and let the solution attain $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$. That is, for all $\boldsymbol{b}\in\mathbf{R}^m$, if $B(\boldsymbol{b}-A\boldsymbol{x})=\boldsymbol{0}$, then $A^{\mathrm{T}}(\boldsymbol{b}-A\boldsymbol{x})=\boldsymbol{0}$, which holds if and only if $\mathcal{N}(B)\subseteq\mathcal{N}(A^{\mathrm{T}})$. This is equivalent to

$$\mathcal{R}(A)\subseteq\mathcal{R}(B^{\mathrm{T}}). \tag{3.3}$$

Noting that $\operatorname{rank}BA\le\min\{\operatorname{rank}A,\operatorname{rank}B\}$, that is,

$$\dim\mathcal{R}(BA)\le\min\{\dim\mathcal{R}(B),\dim\mathcal{R}(A)\},$$

we have from (3.2) that $\dim\mathcal{R}(B)\le\dim\mathcal{R}(A)$. This is equivalent to $\dim\mathcal{R}(B^{\mathrm{T}})\le\dim\mathcal{R}(A)$, which together with (3.3) gives $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$.

($\Longleftarrow$) $\quad\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})\Longrightarrow\mathcal{R}(BA)=\mathcal{R}(BB^{\mathrm{T}})=\mathcal{R}(B)$,

and also

$$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})\Longrightarrow\mathcal{R}(A)\subseteq\mathcal{R}(B^{\mathrm{T}}).$$ ∎

Note the following.

LEMMA 3.16. *If* $B=CA^T$*, where* C *is nonsingular, then* $\mathcal{R}(A)=\mathcal{R}(B^T)$ *holds.*

Proof. $\mathcal{R}(B^{\mathrm{T}})=\mathcal{R}(AC^{\mathrm{T}})=\mathcal{R}(A)$. ∎

Also, if $\operatorname{rank}A=\operatorname{rank}B=m$, then $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})=\mathbf{R}^m$ holds, and hence $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}}BA)$ holds.

Thus, assume $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}}BA)$ holds, and apply GMRES(k) to the least squares problem $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|B\boldsymbol{b}-BA\boldsymbol{x}\|_2$. This gives the following algorithm.

[1] This theorem is due to Professor Masaaki Sugihara.

Method 3.2 (the BA-GMRES(k) method).

Choose $\boldsymbol{x}_0$.
$*$ $\tilde{\boldsymbol{r}}_0 = B(\boldsymbol{b} - A\boldsymbol{x}_0)$
$\boldsymbol{v}_1 = \tilde{\boldsymbol{r}}_0 / \|\tilde{\boldsymbol{r}}_0\|_2$
for $i = 1, 2, \ldots, k$
$\boldsymbol{w}_i = BA\boldsymbol{v}_i$
for $j = 1, 2, \ldots, i$
$h_{j,i} = (\boldsymbol{w}_i, \boldsymbol{v}_j)$
$\boldsymbol{w}_i = \boldsymbol{w}_i - h_{j,i}\boldsymbol{v}_j$
endfor
$h_{i+1,i} = \|\boldsymbol{w}_i\|_2$
$\boldsymbol{v}_{i+1} = \boldsymbol{w}_i / h_{i+1,i}$
Find $\boldsymbol{y}_i \in \mathbf{R}^i$ *to minimize* $\|\tilde{\boldsymbol{r}}_i\|_2 = \left\| \, \|\tilde{\boldsymbol{r}}_0\|_2 \, \boldsymbol{e}_1 - \bar{H}_i \, \boldsymbol{y} \, \right\|_2$.
$\boldsymbol{x}_i = \boldsymbol{x}_0 + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i$
$\boldsymbol{r}_i = \boldsymbol{b} - A\boldsymbol{x}_i$
if $\|A^T\boldsymbol{r}_i\|_2 < \epsilon$ *stop*
endfor
$\boldsymbol{x}_0 = \boldsymbol{x}_k$
go to $*$
Here, $\tilde{\boldsymbol{r}}_i = B\boldsymbol{r}_i$.

Similarly to Theorem 3.6, the following holds.

THEOREM 3.17. *If* $\mathcal{R}(A) = \mathcal{R}(B^T)$, *then* $\mathcal{R}(BA) = \mathcal{R}(A^TB^T) \Longleftrightarrow \mathcal{R}(A^T) = \mathcal{R}(B)$ *holds.*

Proof. If $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$, Lemma 3.2 gives $\mathcal{R}(BA) = \mathcal{R}(BB^{\mathrm{T}}) = \mathcal{R}(B)$ and $\mathcal{R}(A^{\mathrm{T}}B^{\mathrm{T}}) = \mathcal{R}(A^{\mathrm{T}}A) = \mathcal{R}(A^{\mathrm{T}})$. ∎

Let BA-GMRES be BA-GMRES(k) with $k = \infty$ (no restarts). Then, Theorems 2.1 and 3.17 give the following.

THEOREM 3.18. *If* $\mathcal{R}(A) = \mathcal{R}(B^T)$ *holds, then BA-GMRES determines a least squares solution of* $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ *for all* $\boldsymbol{b} \in \mathbf{R}^m$ *and for all* $\boldsymbol{x}_0 \in \mathbf{R}^n$ *without breakdown if and only if* $\mathcal{R}(A^T) = \mathcal{R}(B)$.

As a corollary, we have the following sufficient condition.

COROLLARY 3.19. *If* $\mathcal{R}(A) = \mathcal{R}(B^T)$ *and* $\mathcal{R}(A^T) = \mathcal{R}(B)$, *then BA-GMRES determines a solution of* $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ *for all* $\boldsymbol{b} \in \mathbf{R}^m$ *and all* $\boldsymbol{x}_0 \in \mathbf{R}^n$ *without breakdown.*

We note here that Reichel and Ye [25] proposed a related method for solving underdetermined ($m \le n$) least squares problems using GMRES. Their method is to append $(n - m)$ zero row vectors to the bottom of A to obtain a square singular matrix

$$\tilde{A} = \begin{bmatrix} A \\ 0 \end{bmatrix} \in \mathbf{R}^{n\times n}, \quad \text{and also} \quad \tilde{\boldsymbol{b}} = \begin{bmatrix} \boldsymbol{b} \\ 0 \end{bmatrix} \in \mathbf{R}^n,$$

and apply GMRES to $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\tilde{\boldsymbol{b}} - \tilde{A}\boldsymbol{x}\|_2$. This corresponds to a special case of BA-GMRES with

$$B = \begin{bmatrix} \mathrm{I}_m \\ 0 \end{bmatrix} \in \mathbf{R}^{n\times m},$$

where $\mathrm{I}_m \in \mathbf{R}^{m\times m}$ is an identity matrix, i.e., $\tilde{A} = BA$ and $\tilde{\boldsymbol{b}} = B\boldsymbol{b}$. In this case, $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}}) = \mathbf{R}^m$ holds if and only if $\operatorname{rank} A = m$, but $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ does not necessarily hold. Hence, from Theorem 3.18, their method may break down before

giving a least squares solution. In fact, in [25], Reichel and Ye propose a breakdown-free GMRES to circumvent this difficulty.

3.3. Summary on condition for B. Summing up the above, we have three cases.

General A, including rank-deficient case. For $\operatorname{rank} A \leq \min(m,n)$, including the rank-deficient case, the following holds. If

$$\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}}), \quad \mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B), \tag{3.4}$$

then from Corollaries 3.8 and 3.19, AB-GMRES and BA-GMRES determine a least squares solution of $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ for all $\boldsymbol{b} \in \mathbf{R}^m$ and all $\boldsymbol{x}_0 \in \mathbf{R}^n$ without breakdown.

Condition (3.4) is satisfied if $B = \alpha A^{\mathrm{T}}$, where $0 \neq \alpha \in \mathbf{R}$. For a more sophisticated preconditioner that satisfies (3.4) for the rank-deficient case, see [8, 9, 10].

Full-rank overdetermined case. For $m \geq n = \operatorname{rank} A$, let

$$B = CA^{\mathrm{T}}, \tag{3.5}$$

where $C \in \mathbf{R}^{n\times n}$ is an arbitrary nonsingular matrix. Then, the following holds:

$$B = CA^{\mathrm{T}} \Longleftrightarrow B^{\mathrm{T}} = AC^{\mathrm{T}} \Longrightarrow \mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}}).$$

Hence, $n = \operatorname{rank} A^{\mathrm{T}} = \operatorname{rank} A = \operatorname{rank} B^{\mathrm{T}} = \operatorname{rank} B$ gives

$$\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B) = \mathbf{R}^n.$$

Since $AB \in \mathbf{R}^{m\times m}, BA \in \mathbf{R}^{n\times n}$, with $m \geq n$, the amount of computation per iteration is less for BA-GMRES than for AB-GMRES. This is because BA-GMRES works in a space of smaller dimension, so that the amount of computation for the modified Gram–Schmidt procedure is less. If $\operatorname{rank} A = n$, $A^{\mathrm{T}}A$ and $\operatorname{diag}(A^{\mathrm{T}}A)$ are nonsingular. Hence, a simple example for C is

$$C := \{\operatorname{diag}(A^{\mathrm{T}}A)\}^{-1}, \tag{3.6}$$

i.e., $B = \{\operatorname{diag}(A^{\mathrm{T}}A)\}^{-1}A^{\mathrm{T}}$, as in [33].

Full-rank underdetermined case. For $\operatorname{rank} A = m \leq n$, let

$$B = A^{\mathrm{T}}C, \tag{3.7}$$

where $C \in \mathbf{R}^{m\times m}$ is an arbitrary nonsingular matrix. Then, the following holds:

$$B = A^{\mathrm{T}}C \Longrightarrow \mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B).$$

Hence, from $m = \operatorname{rank} A = \operatorname{rank} A^{\mathrm{T}} = \operatorname{rank} B = \operatorname{rank} B^{\mathrm{T}}$, we also have

$$\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}}) = \mathbf{R}^m.$$

Note here that, since $AB \in \mathbf{R}^{m\times m}$, $BA \in \mathbf{R}^{n\times n}$, $m \leq n$, the amount of computation per iteration is less for AB-GMRES than for BA-GMRES. If $\operatorname{rank} A = m$, then AA^{T} and $\operatorname{diag}(AA^{\mathrm{T}})$ are nonsingular. Hence, a simple example for C is

$$C := \{\operatorname{diag}(AA^{\mathrm{T}})\}^{-1}, \tag{3.8}$$

i.e., $B = A^{\mathrm{T}}\{\operatorname{diag}(AA^{\mathrm{T}})\}^{-1}$.

Note also that when $\operatorname{rank} A = m, \mathcal{R}(A) = \mathbf{R}^m \ni \boldsymbol{b}$, so that

$$\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2 = \min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2 = \min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AA^{\mathrm{T}}C\boldsymbol{z}\|_2 = 0.$$

Hence, AB-GMRES with $B = A^{\mathrm{T}}C$ gives the minimum-norm least squares solution $\boldsymbol{x}^* = B\boldsymbol{z}^* = A^{\mathrm{T}}(C\boldsymbol{z}^*)$ of (1.1), since $AA^{\mathrm{T}}(C\boldsymbol{z}^*) = \boldsymbol{b}$.

4. Convergence analysis. Next, we analyze the convergence of AB-GMRES and BA-GMRES, as well as preconditioned CGLS (PCGLS) and preconditioned CGNE (PCGNE).[2]

4.1. Overdetermined case. First, we consider the overdetermined case $m \geq n$ with $B := CA^{\mathrm{T}}$ as in (3.5), with a restricted choice of C.

THEOREM 4.1. *Let $A \in \mathbf{R}^{m\times n}$, $m \geq n$, and $B := CA^{\mathrm{T}}$, where $C \in \mathbf{R}^{n\times n}$ is symmetric and positive definite. Let the singular values of $\tilde{A} := AC^{\frac{1}{2}}$ be $\sigma_i (1 \leq i \leq n)$. Then, ${\sigma_i}^2$ $(1 \leq i \leq n)$ are eigenvalues of AB and BA. If $m > n$, all the other eigenvalues of AB are 0.*

Proof. Let $\tilde{A} := AC^{\frac{1}{2}} = U\Sigma V^{\mathrm{T}}$ be the singular decomposition of $\tilde{A}$. Here, $U \in \mathbf{R}^{m\times m}, V \in \mathbf{R}^{n\times n}$ are orthogonal matrices, and

$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ & & \sigma_n \\ 0 & & \end{bmatrix} \in \mathbf{R}^{m\times n},$$

where $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ are the singular values of $\tilde{A}$. Then,

$$AB = ACA^{\mathrm{T}} = \tilde{A}\,\tilde{A}^{\mathrm{T}} = U\Sigma\Sigma^{\mathrm{T}}U^{\mathrm{T}}, \tag{4.1}$$

$$BA = CA^{\mathrm{T}}A = C^{\frac{1}{2}}\tilde{A}^{\mathrm{T}}\tilde{A}C^{-\frac{1}{2}} = C^{\frac{1}{2}}V\Sigma^{\mathrm{T}}\Sigma(C^{\frac{1}{2}}V)^{-1}. \tag{4.2}$$

∎

Note that AB is symmetric. It is also positive definite if $\operatorname{rank} A = n$.

If $\operatorname{rank} A = n$, then $C := \{\operatorname{diag}(A^{\mathrm{T}}A)\}^{-1}$ in the example in (3.6) is symmetric positive definite, and $C := (LDL^{\mathrm{T}})^{-1}$ for the RIF preconditioner introduced in section 5 is also symmetric positive definite.

Now note the following theorem. (See, for instance, [27].)

THEOREM 4.2. *Assume $A' \in \mathbf{R}^{n\times n}$ is diagonalizable, and let $A' = X\Lambda X^{-1}$, where $\Lambda = \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$. Then, the residual norm achieved by the kth step of GMRES applied to $A'\boldsymbol{x}' = \boldsymbol{b}'$ satisfies*

$$\|\boldsymbol{r}'_m\|_2 \leq \kappa_2(X)\epsilon^{(k)}\|\boldsymbol{r}'_0\|_2,$$

where $\kappa_2(X) = \|X\|_2\|X^{-1}\|_2$ and

$$\epsilon^{(k)} = \min_{p_k\in Q_k} \max_{i=1,\ldots,n} |p_k(\lambda_i)| \leq 2\left[\frac{\sqrt{\kappa(A')}-1}{\sqrt{\kappa(A')}+1}\right]^k,$$

where $Q_k := \{p_k \mid p_k(x) : \text{polynomial of } x \text{ with degree} \leq k,\ p_k(0) = 1\}$ and $\kappa(A') = \lambda_1/\lambda_n$.

First, consider AB-GMRES. Let $\operatorname{rank} A = r$ and $U = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$. Then, $AB\boldsymbol{u}_i = {\sigma_i}^2\boldsymbol{u}_i$ $(i = 1, \ldots, r)$, $AB\boldsymbol{u}_i = \boldsymbol{0}$ $(i = r+1, \ldots, m)$, so that $\mathcal{R}(AB) = \mathcal{R}(A) = \operatorname{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$, and $\mathcal{N}(AB) = \mathcal{R}(AB)^{\perp} = \mathcal{R}(A)^{\perp} = \operatorname{span}\{\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m\}$. Let

$$\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0 = \sum_{i=1}^{m} \rho_i\boldsymbol{u}_i\ . \tag{4.3}$$

[2] We would like to thank Professor Michael Eiermann for discussions that led to the following analysis.

The kth residual vector $\boldsymbol{r}_k = \boldsymbol{b} - A\boldsymbol{x}_k$ satisfies

$$\|\boldsymbol{r}_k\|_2 = \min_{\boldsymbol{\zeta} \in \mathcal{K}_k(AB, \boldsymbol{r}_0)} \|\boldsymbol{b} - AB(\boldsymbol{z}_0 + \boldsymbol{\zeta})\|_2 = \min_{p_k \in Q_k} \|p_k(AB)\boldsymbol{r}_0\|_2 \, .$$

Note from (4.1) that

$$p_k(AB) = U p_k(\Sigma\Sigma^{\mathrm{T}}) U^{\mathrm{T}} = \sum_{i=1}^{r} p_k({\sigma_i}^2) \boldsymbol{u}_i {\boldsymbol{u}_i}^{\mathrm{T}} + p_k(0) \sum_{i=r+1}^{m} \boldsymbol{u}_i {\boldsymbol{u}_i}^{\mathrm{T}}.$$

Hence, (4.3) gives

$$p_k(AB)\boldsymbol{r}_0 = \sum_{i=1}^{r} \rho_i \, p_k({\sigma_i}^2) \boldsymbol{u}_i + p_k(0) \sum_{i=r+1}^{m} \rho_i \boldsymbol{u}_i.$$

Hence, the norm of the $\mathcal{R}(A)$-component of the residual is given by

$$\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2{}^2 = \min_{p_k \in Q_k} \sum_{i=1}^{r} {\rho_i}^2 \{p_k({\sigma_i}^2)\}^2,$$

where

$$\|\boldsymbol{r}_0|_{\mathcal{R}(A)}\|_2{}^2 = \sum_{i=1}^{r} {\rho_i}^2.$$

Thus,

$$\begin{aligned} \frac{\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2}{\|\boldsymbol{r}_0|_{\mathcal{R}(A)}\|_2} &\leq \min_{p_k \in Q_k} \max_{1 \leq i \leq r} |p_k({\sigma_i}^2)| \\ &\leq 2 \left[\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right]^k = 2 \left(\frac{\sigma_1 - \sigma_r}{\sigma_1 + \sigma_r} \right)^k, \end{aligned}$$

where $\kappa = (\sigma_1/\sigma_r)^2$.

Thus, AB-GMRES minimizes $\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2$, and we have the following under the same assumptions and notation as in Theorem 4.1.

THEOREM 4.3. *The residual $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$, achieved by the kth step of AB-GMRES, satisfies*

$$\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2 \leq 2 \left(\frac{\sigma_1 - \sigma_r}{\sigma_1 + \sigma_r} \right)^k \|\boldsymbol{r}_0|_{\mathcal{R}(A)}\|_2.$$

Next, consider BA-GMRES. Assume $\operatorname{rank} A = n$. In Theorem 4.2, let $A' = BA, \boldsymbol{b}' = B\boldsymbol{b}$, and let $\boldsymbol{x}' = \boldsymbol{x}$. Then, we have $\boldsymbol{r}' = B\boldsymbol{r} = CA^{\mathrm{T}}\boldsymbol{r}$, $\lambda_i = {\sigma_i}^2$ $(i = 1, \ldots, n)$, $X = C^{\frac{1}{2}}V, \kappa_2(X) = \sqrt{\kappa(C)}$, where $\kappa(C) = \frac{\lambda_{\max}(C)}{\lambda_{\min}(C)}$, $\kappa(A') = \kappa(BA) = (\frac{\sigma_1}{\sigma_n})^2$, $\frac{\sqrt{\kappa(A')}-1}{\sqrt{\kappa(A')}+1} = \frac{\sigma_1 - \sigma_n}{\sigma_1 + \sigma_n}$.

Thus, BA-GMRES minimizes $\|B\boldsymbol{r}_k\|_2$, and we have the following under the same assumptions and notation as in Theorem 4.1.

THEOREM 4.4. *The residual $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$, achieved by the kth step of BA-GMRES, satisfies*

$$\|B\boldsymbol{r}_k\|_2 = \|CA^{T}\boldsymbol{r}_k\| \leq 2\sqrt{\kappa(C)} \left(\frac{\sigma_1 - \sigma_n}{\sigma_1 + \sigma_n} \right)^k \|B\boldsymbol{r}_0\|_2.$$

Next, we turn to CGLS. Consider a symmetric positive definite preconditioner matrix $C = (\tilde{L}\tilde{L}^{\mathrm{T}})^{-1} \in \mathbf{R}^{n\times n}$. Here, C may be based on an incomplete factorization $LDL^{\mathrm{T}} \sim A^{\mathrm{T}}A$, $\tilde{L} = LD^{\frac{1}{2}}$ as in the RIF. For the diagonal scaling, we have $\mathrm{diag}(A^{\mathrm{T}}A) \sim A^{\mathrm{T}}A$, $\tilde{L} = (\mathrm{diag}(A^{\mathrm{T}}A))^{\frac{1}{2}}$.

Then, the natural way to precondition the CGLS (CGNR) method, or the mathematically equivalent LSQR method, is to apply CG to

$$A'\boldsymbol{x}' = \boldsymbol{b}', \tag{4.4}$$

where $A' = \tilde{L}^{-1}A^{\mathrm{T}}A\tilde{L}^{-\mathrm{T}}$ is symmetric positive definite, $\boldsymbol{x}' = \tilde{L}^{\mathrm{T}}\boldsymbol{x}$, and $\boldsymbol{b}' = \tilde{L}^{-1}A^{\mathrm{T}}\boldsymbol{b}$.

Note that (4.4) is equivalent to $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$, whose solution is $\boldsymbol{x}_* = A^{\dagger}\boldsymbol{b}$. Let $\boldsymbol{e} = \boldsymbol{x} - \boldsymbol{x}_*$, $\boldsymbol{x}'_* = \tilde{L}^{\mathrm{T}}\boldsymbol{x}_*$, and $\boldsymbol{e}' = \tilde{L}^{\mathrm{T}}\boldsymbol{e}$.

From here, assume $\mathrm{rank}\, A = n$. Then, $\boldsymbol{x}_* = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$.

CG applied to (4.4) minimizes $\|\boldsymbol{e}'_k\|_{A'}$, which is bounded as follows. (See, e.g., [27].)

THEOREM 4.5. *Let $\boldsymbol{x}_k$ be the kth iterate of the CG method applied to $A'\boldsymbol{x}' = \boldsymbol{b}'$, where A' is symmetric positive definite. Then, CG minimizes $\|\boldsymbol{e}'_k\|_{A'}$, which is bounded as follows:*

$$\|\boldsymbol{e}'_k\|_{A'} \le 2\left\{\frac{\sqrt{\kappa(A')}-1}{\sqrt{\kappa(A')}+1}\right\}^k \|\boldsymbol{e}'_0\|_{A'},$$

where $\boldsymbol{e}'_k = \boldsymbol{x}'_* - \boldsymbol{x}'_k$, $\boldsymbol{x}'_* = {A'}^{-1}\boldsymbol{b}'$, and $\kappa(A') = \lambda_{\max}(A')/\lambda_{\min}(A')$.

Since $(\tilde{L}^{\mathrm{T}})^{-1}A'(\tilde{L}^{\mathrm{T}}) = CA^{\mathrm{T}}A = BA$, $\lambda_i(A') = \lambda_i(BA) = {\sigma_i}^2$ $(i = 1,\dots,n)$. Hence, $\kappa(A') = (\frac{\sigma_1}{\sigma_n})^2$ and $\frac{\sqrt{\kappa(A')}-1}{\sqrt{\kappa(A')}+1} = \frac{\sigma_1-\sigma_n}{\sigma_1+\sigma_n}$.

Note

$$\boldsymbol{r}|_{\mathcal{R}(A)} = P_{\mathcal{R}(A)}\boldsymbol{r} = AA^{\dagger}\boldsymbol{r} = AA^{\dagger}(\boldsymbol{b} - A\boldsymbol{x}) = A\boldsymbol{x}^* - A\boldsymbol{x} = -A\boldsymbol{e},$$

where P_X denotes the orthogonal projector onto subspace X, and

$$\|\boldsymbol{e}'\|_{A'}{}^2 = (\boldsymbol{e}', A'\boldsymbol{e}') = (\boldsymbol{e}, A^{\mathrm{T}}A\boldsymbol{e}) = (A\boldsymbol{e}, A\boldsymbol{e}) = \|\boldsymbol{r}|_{\mathcal{R}(A)}\|_2{}^2.$$

Hence, the preconditioned CGLS (PCGLS) minimizes $\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2$, and we have the following under the same assumptions and notation as in Theorem 4.1.

THEOREM 4.6. *The residual $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$ of the kth step of preconditioned CGLS (CG applied to* (4.4)*) satisfies*

$$\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2 = \|\boldsymbol{e}_k\|_{A^TA} \le 2\left(\frac{\sigma_1-\sigma_n}{\sigma_1+\sigma_n}\right)^k \|\boldsymbol{r}_0|_{\mathcal{R}(A)}\|_2.$$

Note that the kth approximate solution $\boldsymbol{x}_k$ of AB-GMRES, BA-GMRES, and PCGLS all belong to $\boldsymbol{x}_0 + \mathcal{K}_k(BA, B\boldsymbol{r}_0)$. AB-GMRES and PCGLS both minimize $\|\boldsymbol{r}_k|_{\mathcal{R}(A)}\|_2$, whereas BA-GMRES minimizes $\|B\boldsymbol{r}_k\|_2$. Hence, AB-GMRES and PCGLS are mathematically equivalent.

From the above, we expect that the convergence behavior of AB-GMRES and PCGLS are the same, and BA-GMRES will exhibit similar convergence behavior for the full-rank overdetermined case. This is confirmed by the experiments relative to reorthogonalized CGLS in section 6.

4.2. Underdetermined case. Similarly, consider the underdetermined case $m \leq n$ with $B = A^{\mathrm{T}}C$ as in (3.7), where $C \in \mathbf{R}^{m\times m}$ is restricted to be symmetric and positive definite. The following holds.

THEOREM 4.7. *Let $A \in \mathbf{R}^{m\times n}$, $m \leq n$, where $B := A^{\mathrm{T}}C$ and $C \in \mathbf{R}^{m\times m}$ is symmetric and positive definite. Let the singular values of $\tilde{A} := C^{\frac{1}{2}}A$ be σ_i $(1 \leq i \leq m)$. Then, ${\sigma_i}^2$ $(1 \leq i \leq m)$ are the eigenvalues of AB and BA. If $m < n$, all the other eigenvalues of BA are 0.*

Proof. Let $\tilde{A} := C^{\frac{1}{2}}A = U\Sigma V^{\mathrm{T}}$ be the singular value decomposition of $\tilde{A}$. Here, $U \in \mathbf{R}^{m\times m}, V \in \mathbf{R}^{n\times n}$ are orthogonal matrices, and

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_m & \end{bmatrix} \in \mathbf{R}^{m\times n},$$

where $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$ are the singular values of $\tilde{A}$. Then,

$$AB = AA^{\mathrm{T}}C = C^{-\frac{1}{2}}\tilde{A}\,\tilde{A}^{\mathrm{T}}C^{\frac{1}{2}} = C^{-\frac{1}{2}}U\Sigma\Sigma^{\mathrm{T}}(C^{-\frac{1}{2}}U)^{-1},$$

$$BA = A^{\mathrm{T}}CA = \tilde{A}^{\mathrm{T}}\tilde{A} = V\Sigma^{\mathrm{T}}\Sigma V^{\mathrm{T}}. \qquad \square$$

First, consider AB-GMRES. Assume $\operatorname{rank} A = m$. In Theorem 4.2, let $A' = AB$, $\boldsymbol{b}' = \boldsymbol{b}$, $\boldsymbol{x}' = \boldsymbol{z}$. Then we have $\boldsymbol{r}' = \boldsymbol{b}' - A'\boldsymbol{x}' = \boldsymbol{b} - AB\boldsymbol{z} = \boldsymbol{r}$, $\lambda_i = {\sigma_i}^2$ $(i = 1, \ldots, m)$, $X = C^{-\frac{1}{2}}U$, $\kappa_2(X) = \sqrt{\kappa(C)}$, where $\kappa(C) = \lambda_{\max}(C)/\lambda_{\min}(C)$, $\kappa(A') = \kappa(AB) = (\sigma_1/\sigma_m)^2$.

AB-GMRES minimizes $\|\boldsymbol{r}\|_2$, and we have the following theorem under the same assumptions and notation as in Theorem 4.7.

THEOREM 4.8. *The residual $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$ achieved by the kth step of AB-GMRES satisfies*

$$\|\boldsymbol{r}_k\|_2 \leq 2\sqrt{\kappa(C)}\left(\frac{\sigma_1 - \sigma_m}{\sigma_1 + \sigma_m}\right)^k \|\boldsymbol{r}_0\|_2.$$

Next, consider BA-GMRES. Assume $\operatorname{rank} A = r$. Note that $BA = A^{\mathrm{T}}CA$ is symmetric, $BA\boldsymbol{v}_i = {\sigma_i}^2\boldsymbol{v}_i$ $(i = 1, \ldots, r)$, and $BA\boldsymbol{v}_i = \boldsymbol{0}$ $(i = r+1, \ldots, n)$, so that $\mathcal{R}(BA) = \mathcal{R}(B) = \mathcal{R}(A^{\mathrm{T}}) = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ and $\mathcal{N}(BA) = \mathcal{R}(BA)^{\perp} = \mathcal{R}(B)^{\perp} = \operatorname{span}\{\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n\}$.

Let

$$B\boldsymbol{r}_0 = \sum_{i=1}^{n} \rho_i \boldsymbol{v}_i \in \mathbf{R}^n.$$

Then,

$$p_k(BA)B\boldsymbol{r}_0 = \sum_{i=1}^{r} \rho_i p_k({\sigma_i}^2)\boldsymbol{v}_i + p_k(0)\sum_{i=r+1}^{n} \rho_i \boldsymbol{v}_i.$$

The kth iterate of BA-GMRES satisfies

$$\|B\boldsymbol{r}_k\|_2 = \min_{\boldsymbol{\xi} \in \mathcal{K}_k(BA, B\boldsymbol{r}_0)} \|B\boldsymbol{b} - BA(\boldsymbol{x}_0 + \boldsymbol{\xi})\|_2 = \min_{p_k \in Q_k} \|p_k(BA)B\boldsymbol{r}_0\|_2.$$

Hence,

$$\|B\boldsymbol{r}_k|_{\mathcal{R}(B)}\|_2{}^2 = \min_{p_k\in Q_k} \sum_{i=1}^{r} \rho_i{}^2 p_k(\sigma_i{}^2)^2 \quad \text{and} \quad \|B\boldsymbol{r}_0|_{\mathcal{R}(B)}\|_2{}^2 = \sum_{i=1}^{r} \rho_i{}^2,$$

where $\mathcal{R}(B) = \mathcal{R}(A^{\mathrm{T}})$, so that

$$\begin{aligned} \frac{\|B\boldsymbol{r}_k|_{\mathcal{R}(B)}\|_2}{\|B\boldsymbol{r}_0|_{\mathcal{R}(B)}\|_2} &\le \min_{p_k\in Q_k} \max_{1\le i\le r} |p_k(\sigma_i{}^2)| \\ &\le 2\left[\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right]^k = 2\left(\frac{\sigma_1-\sigma_r}{\sigma_1+\sigma_r}\right)^k, \end{aligned} \tag{4.5}$$

where $\kappa = (\sigma_1/\sigma_r)^2$.

Thus, BA-GMRES minimizes $\|B\boldsymbol{r}_k|_{\mathcal{R}(B)}\|_2$, and we have the following theorem under the same assumptions and notation as in Theorem 4.7.

THEOREM 4.9. *The residual $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$, achieved by the kth step of BA-GMRES, satisfies*

$$\|B\boldsymbol{r}_k|_{\mathcal{R}(B)}\|_2 \le 2\left(\frac{\sigma_1-\sigma_r}{\sigma_1+\sigma_r}\right)^k \|B\boldsymbol{r}_0|_{\mathcal{R}(B)}\|_2.$$

Next, we turn to the CGLS (CGNE) method.

Consider a symmetric positive definite preconditioning matrix $C = (\tilde{L}\tilde{L}^{\mathrm{T}})^{-1} \in \mathbf{R}^{m\times m}$, where $\tilde{L} = LD^{\frac{1}{2}}$, $\tilde{L}\tilde{L}^{\mathrm{T}} = LDL^{\mathrm{T}} \sim AA^{\mathrm{T}}$ as in the RIF method, or $\tilde{L} = (\mathrm{diag}(AA^{\mathrm{T}}))^{\frac{1}{2}}$ in the diagonal scaling. Then, one can precondition $AA^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{b}$ as

$$A'\boldsymbol{y}' = \boldsymbol{b}', \tag{4.6}$$

where $A' = \tilde{L}^{-1}AA^{\mathrm{T}}\tilde{L}^{-\mathrm{T}}$, $\boldsymbol{y}' = \tilde{L}^{\mathrm{T}}\boldsymbol{y}$, $\boldsymbol{b}' = \tilde{L}^{-1}\boldsymbol{b}$, $\boldsymbol{r}' = \tilde{L}^{-1}\boldsymbol{r}$.

Note that (4.6) is equivalent to $AA^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{b}$. From here, assume $\operatorname{rank} A = m$. Then, its solution is given by $\boldsymbol{y}_* = (AA^{\mathrm{T}})^{-1}\boldsymbol{b}$. Let $\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{y}_*$, $\boldsymbol{y}'_* = \tilde{L}^{\mathrm{T}}\boldsymbol{y}_*$, and $\boldsymbol{e}' = \tilde{L}^{\mathrm{T}}\boldsymbol{e}$.

CG applied to (4.6) minimizes $\|\boldsymbol{e}'_k\|_{A'}$.

Since $\tilde{L}A'\tilde{L}^{-1} = AA^{\mathrm{T}}C = AB$, we have $\lambda_i(A') = \lambda_i(AB) = \sigma_i{}^2$ $(i = 1, \ldots, m)$.

From Theorem 4.5, we have the following theorem under the same assumptions and notation as in Theorem 4.7.

THEOREM 4.10. *The residual $\boldsymbol{r} = \boldsymbol{b} - AA^{T}\boldsymbol{y}$ of the kth step of the preconditioned CGLS (CGNE) method (CG applied to* (4.6)*) satisfies*

$$\|\boldsymbol{r}_k\|_{(AA^{T})^{-1}} = \|\boldsymbol{e}_k\|_{AA^{T}} \le 2\left(\frac{\sigma_1-\sigma_m}{\sigma_1+\sigma_m}\right)^k \|\boldsymbol{r}_0\|_{(AA^{T})^{-1}}.$$

Note here that $\|\boldsymbol{r}_k\|_{(AA^{\mathrm{T}})^{-1}} = \|\boldsymbol{\varepsilon}_k\|_2$, where $\boldsymbol{\varepsilon}_k = \boldsymbol{x}_k - \boldsymbol{x}_*$, $\boldsymbol{x}_k = A^{\mathrm{T}}\boldsymbol{y}_k$, $\boldsymbol{x}_* = A^{\mathrm{T}}\boldsymbol{x}_*$, and $\boldsymbol{x}_*$ is the solution to $\min_{A\boldsymbol{x}=\boldsymbol{b}} \|\boldsymbol{x}\|_2$.

The kth approximate solution $\boldsymbol{x}_k$ of AB-GMRES, BA-GMRES, and the preconditioned CGNE (PCGNE) all belong to $\boldsymbol{x}_0 + \mathcal{K}_k(BA, B\boldsymbol{r}_0)$. AB-GMRES minimizes $\|\boldsymbol{r}_k\|_2$, BA-GMRES minimizes $\|B\boldsymbol{r}_k|_{\mathcal{R}(B)}\|_2$, and PCGNE minimizes $\|\boldsymbol{r}_k\|_{(AA^{\mathrm{T}})^{-1}}$.

From the above, we expect AB-GMRES, BA-GMRES, and PCGNE to exhibit similar convergence behavior for the full-rank underdetermined case. This is confirmed by the experiments relative to the reorthogonalized CGNE in section 6.

5. The choice of B. Besides satisfying the conditions $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ and $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$, it is desirable that B satisfy $AB \approx \mathrm{I}_m$ or $BA \approx \mathrm{I}_n$, in order to speed up the convergence.

Simple candidates, as mentioned before, are $C := \{\mathrm{diag}(A^{\mathrm{T}}A)\}^{-1}$ and $B = CA^{\mathrm{T}}$ when $m \geq n = \mathrm{rank}\, A$, and $C := \{\mathrm{diag}(AA^{\mathrm{T}})\}^{-1}$ and $B = A^{\mathrm{T}}C$ when $\mathrm{rank}\, A = m \leq n$.

More sophisticated preconditioners based on incomplete QR decompositions of A may be considered. That is, $A = QR + E$, where $Q \in \mathbf{R}^{m\times n}$ is an (approximately) orthogonal matrix, $R \in \mathbf{R}^{n\times n}$ is an upper triangular matrix, and E is the error matrix.

Usually, R is used as a preconditioner for CGLS, as mentioned in (1.4). Similarly, we may let $B = R^{-1}Q^{\mathrm{T}}$ for AB-GMRES(k) and BA-GMRES(k), for instance, when $m \geq n$, and in the case of BA-GMRES(k) we may apply GMRES(k) to $R^{-1}Q^{\mathrm{T}}A\boldsymbol{x} = R^{-1}Q^{\mathrm{T}}\boldsymbol{b}$.

There are many approaches to constructing an incomplete QR decomposition of A, e.g., the incomplete MGS method [19, 26, 32], the incomplete Householder reflection, and the incomplete Givens rotation [1, 24]. In [18, 14, 15], the IMGS(l) method, an approximation of the MGS method, was applied to CGLS, AB-GMRES(k), and BA-GMRES(k).

The RIF is an attractive method with low memory requirements [2]. It has been applied to CGLS for sparse least squares problems [3]. The method is based on an $A^{\mathrm{T}}A$-orthogonalization procedure and never breaks down. In the complete factorization, the method gives an upper triangular matrix Z or a lower triangular matrix L, and a diagonal matrix $D = \mathrm{diag}(\boldsymbol{d})$ such that

$$A^{\mathrm{T}}A = LDL^{\mathrm{T}} \quad \text{and} \quad A^{\mathrm{T}}A = Z^{-\mathrm{T}}DZ^{-1},$$

where $A \in \mathbf{R}^{m\times n}$ with $m \geq n = \mathrm{rank}\, A$.

In the following, we approximate $(A^{\mathrm{T}}A)^{-1}$ by $ZD^{-1}Z^{\mathrm{T}}$, so that the extra work necessary for each iteration of the preconditioned algorithm is only two matrix-vector multiplications and a vector scaling. With a drop tolerance τ, the number of fill-ins of the matrix Z can be controlled. For $1 \leq j \leq n$, let $\mathbf{z}_j$, $\mathbf{e}_j$, and d_j be the jth column vector of the matrix Z, the jth unit basis vector, and the jth entry of the vector $\boldsymbol{d}$, respectively. The method is described as follows.

Method 5.1 (the robust incomplete factorization (RIF) method).

Let $Z = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n] = \mathrm{I}_n$.
for $j = 1, \ldots, n$ *do:*
  Compute $\mathbf{u}_j = A\mathbf{z}_j$.
  Compute $d_j = (\mathbf{u}_j, \mathbf{u}_j)$.
  for $i = j+1, \ldots, n$ *do:*
    Compute $\mathbf{v}_i = A\mathbf{e}_i$.
    Compute $\theta_{ij} = \frac{(\mathbf{v}_i, \mathbf{u}_j)}{d_j}$.
    if $|\theta_{ij}| > \tau$
      Compute $\mathbf{z}_i = \mathbf{z}_i - \theta_{ij}\mathbf{z}_j$.
    endif
    Drop the elements in $\mathbf{z}_i$ *that are smaller than* τ.
  enddo
enddo

If $\tau = 0$ in the above method, the complete factorization of the matrix $A^{\mathrm{T}}A$ is obtained, so that $A^{\mathrm{T}}A = Z^{-\mathrm{T}}DZ^{-1}$. If we let $\widetilde{L} = Z^{-\mathrm{T}}D^{1/2}$, we have $A^{\mathrm{T}}A = \widetilde{L}\widetilde{L}^{\mathrm{T}}$,

which means that the RIF is equivalent to the Cholesky factorization of the matrix of the normal equation.

The drop tolerance parameter τ plays an important role in the RIF method, as it determines not only how the sparse matrix Z approximates the corresponding complete factorization matrix, but also the amount of computation and storage required. The relative drop tolerance can also be used by replacing τ with $\tau\|\boldsymbol{a}_i\|_2$, where $\boldsymbol{a}_i$ is the ith column of A [3]. In the numerical experiments in section 6, we use the relative drop tolerance.

The matrix $ZD^{-1}Z^{\mathrm{T}}$ is guaranteed to be positive definite when A has full column rank, since $d_j > 0$ for $1 \le j \le n$, and Z is a sparse upper triangular matrix whose diagonal elements are one. Therefore, let $C := ZD^{-1}Z^{\mathrm{T}}$ and $B = CA^{\mathrm{T}}$ when $m \ge n = \operatorname{rank} A$. Since C is nonsingular, B satisfies $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ and $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ (cf. section 3.3 and Theorem 3.18). The approximation $BA \approx I_n$ improves as τ approaches 0.

For the underdetermined case where $\operatorname{rank} A = m \le n$, we can also construct matrices $\widehat{Z}$ and $\widehat{D}$ using RIF, where $AA^{\mathrm{T}} \approx \widehat{Z}^{-\mathrm{T}}\widehat{D}\widehat{Z}^{-1}$. Then, let $B = A^{\mathrm{T}}C$, where $C := \widehat{Z}\widehat{D}^{-1}\widehat{Z}^{\mathrm{T}}$. B satisfies the conditions in Theorem 3.7 and $AB \approx I_m$.

6. Numerical experiments. Finally, we present numerical experiment results to show the performance of AB-GMRES and BA-GMRES. We compare them to the preconditioned CGLS [4] (also called CGNR in [27]) and LSQR [23] method for overdetermined problems, and also compare them to the preconditioned CGNE [27] and LSQR method for underdetermined problems.

All of the computations were run on a Dell Precision 690, where the CPU is 3 GHz and the memory is 16 GB, and the programming language and compiling environment was MATLAB 6.5 or GNU C/C++ 3.4.3 (for CPU-time measurements) in Redhat Linux.

6.1. Overdetermined case. In this section, we consider the overdetermined problem

$$\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2, \qquad A \in \mathbf{R}^{m\times n} \qquad (m \ge n)$$

and show the performance of BA-GMRES by three classes of problems.

First, in order to observe how the condition number and the distribution of singular values of A influence the convergence of CGLS, RCGLS (reorthogonalized CGLS based on Method 1.1 with $l = \infty$), and BA-GMRES, we test a class of matrices A of the form $A = U\Sigma V^{\mathrm{T}}$, where $U \in \mathbf{R}^{m\times m}$ and $V \in \mathbf{R}^{n\times n}$ are random orthogonal matrices, and $\Sigma \in \mathbf{R}^{m\times n}$ is a rectangular diagonal matrix with diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_n$, where the ith smallest singular value is

$$\sigma_{n-i+1} = \sigma_n + \frac{i-1}{n-1}(\sigma_1 - \sigma_n)\rho^{n-i}, \quad i = 1, \ldots, n.$$

Here, the parameter ρ is chosen between 0 and 1. For $\rho = 1$, the singular values are uniformly spaced. For $\rho < 1$, the singular values are tightly clustered towards the smallest singular value σ_n and are far apart towards the largest singular value σ_1. The idea of generating this kind of matrix is from [29, 13].

Letting $m = 50, n = 25$, $\sigma_1 = 1$, $\sigma_n = 0.01$ or 0.0001, and $\rho = 1, 0.8, 0.6, 0.4$, we form inconsistent least squares problems where the elements of the right-hand side vector $\boldsymbol{b}$ were generated randomly following the normal distribution with mean zero and variance 1, using the MATLAB code randn(m,1). The same $\boldsymbol{b}$ was used for all

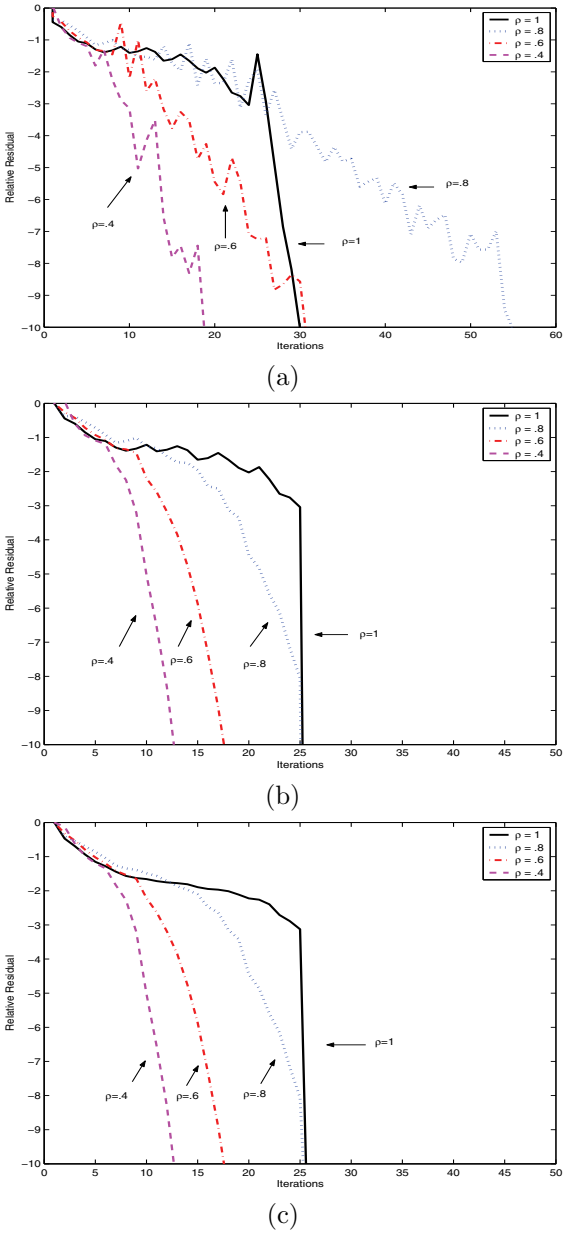

FIG. 6.1. $\|A^T\boldsymbol{r}\|_2/\|A^T\boldsymbol{b}\|_2$ *vs. iterations for* (a) *CGLS,* (b) *RCGLS, and* (c) *BA-GMRES for* $\sigma_n = 0.01$ *and* $\rho = 1$ *solid,* $\rho = 0.8$ *dotted,* $\rho = 0.6$ *dash-dot, and* $\rho = 0.4$ *dashed, inconsistent problem.*

the cases. The initial guess is $\boldsymbol{x}_0 = (0, 0, \ldots, 0)^{\mathrm{T}}$. Experiments were performed in MATLAB with machine precision $\epsilon \approx 1.1 \times 10^{-16}$.

Figures 6.1 and 6.2 show for $\sigma_n = 0.01$ and 0.0001, respectively, the relative residual $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$ vs. the number of iterations for CGLS and RCGLS without any preconditioning, as well as BA-GMRES with $B = A^{\mathrm{T}}$.

The above results indicate that the convergence of CGLS in finite-precision arith-

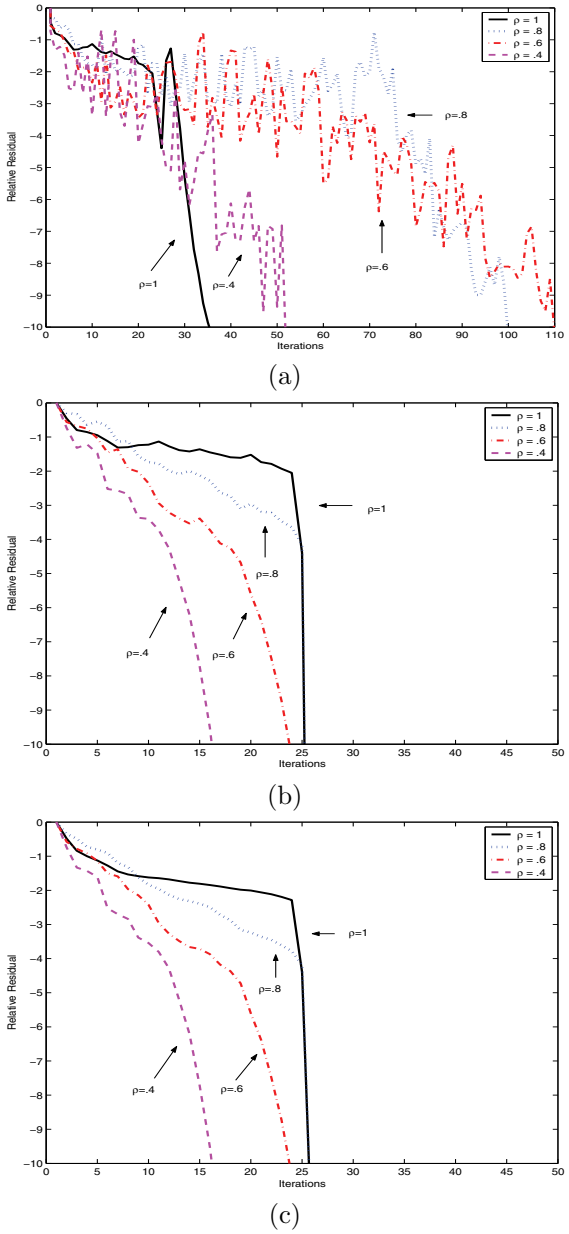

FIG. 6.2. $\|A^T\boldsymbol{r}\|_2/\|A^T\boldsymbol{b}\|_2$ *vs. iterations for* (a) *CGLS,* (b) *RCGLS, and* (c) *BA-GMRES for* $\sigma_n = 0.0001$ *and* $\rho = 1$ *solid,* $\rho = 0.8$ *dotted,* $\rho = 0.6$ *dash-dot, and* $\rho = 0.4$ *dashed, inconsistent problem.*

TABLE 6.1
The condition numbers of the test matrices.

Name	$\kappa(A)$
RANDL1	1.9×10
RANDL2	1.6×10^2
RANDL3	1.3×10^3
RANDL4	2.0×10^4
RANDL5	1.3×10^5
RANDL6	1.3×10^6
RANDL7	1.3×10^7

metic is very much influenced by the distribution of the singular values as well as the condition number of A (cf. [29, 13]). This may be due to the loss of orthogonality of the vectors due to the use of three-term recurrence.

On the other hand, the convergence of RCGLS and BA-GMRES were less influenced by the condition number and the distribution of the singular values. This may be explained by the fact that both apply the Gram–Schmidt process to explicitly orthogonalize the vectors, so that the loss of orthogonality due to rounding error that occurs in CGLS is suppressed.

Note also that the convergence behaviors of RCGLS and BA-GMRES are similar, as predicted in section 4.1.

Next, we generated a class of matrices by the MATLAB routine "sprandn" with $m = 30,000, n = 3,000$, and density 0.1%. The condition numbers $\kappa(A)$ of these matrices are given in Table 6.1. In this experiment, we specified the density (the ratio of nonzero elements) and $\kappa(A)$. The nonzero element values were generated by a random number generator following the normal distribution, and the pattern of the nonzero elements was also determined by a random number generator.

In these experiments, we compare the CGLS, LSQR, RCGLS (fully reorthogonalized CGLS based on Method 1.1 with $l = \infty$), and BA-GMRES, with the diagonal scaling preconditioner and the RIF preconditioner. The methods are denoted by CGLS-diag., LSQR-diag., RCGLS-diag., BA-GMRES-diag., and CGLS-RIF, LSQR-RIF, RCGLS-RIF, BA-GMRES-RIF, respectively. We let $B = CA^{\mathrm{T}}$ for BA-GMRES with $C = \{\mathrm{diag}(A^{\mathrm{T}}A)\}^{-1}$ for diagonal scaling and $C = ZD^{-1}Z^{\mathrm{T}}$ for RIF. In reorthogonalized CGLS, the (preconditioned) residual vectors were reorthogonalized with respect to all previous (preconditioned) residual vectors.

We set the initial approximate solution to $\boldsymbol{x}_0 = \boldsymbol{0}$, and the convergence was monitored by observing the relative residual (of the range component): $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$, where $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$ is the residual. When the computation times were compared, the convergence criterion for the relative residual was set to 10^{-6}. For the right-hand side vector $\boldsymbol{b}$, each of its components was generated by a random number generator following the normal distribution, so that $\boldsymbol{b} \in \mathbf{R}^m$ but $\boldsymbol{b} \notin \mathcal{R}(A)$.

For BA-GMRES, $\|B\boldsymbol{r}\|_2 = \|CA^{\mathrm{T}}\boldsymbol{r}\|_2$ is readily available in the GMRES process, so it is practical to use it to judge convergence. However, in the experiments below, we have compared all the methods by $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$, and the extra time to compute $\|A^{\mathrm{T}}\boldsymbol{r}\|_2$ was neglected in the CPU times for BA-GMRES.

Figure 6.3 shows $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$ vs. the number of iterations for AB-GMRES, BA-GMRES ($k = \infty$), and PRCGLS (preconditioned, fully reorthogonalized CGLS), with diagonal scaling and RIF preconditioning for the problem RANDL3. Here the relative drop tolerance for RIF was set to value $\tau = 0.8$ (determined experimentally

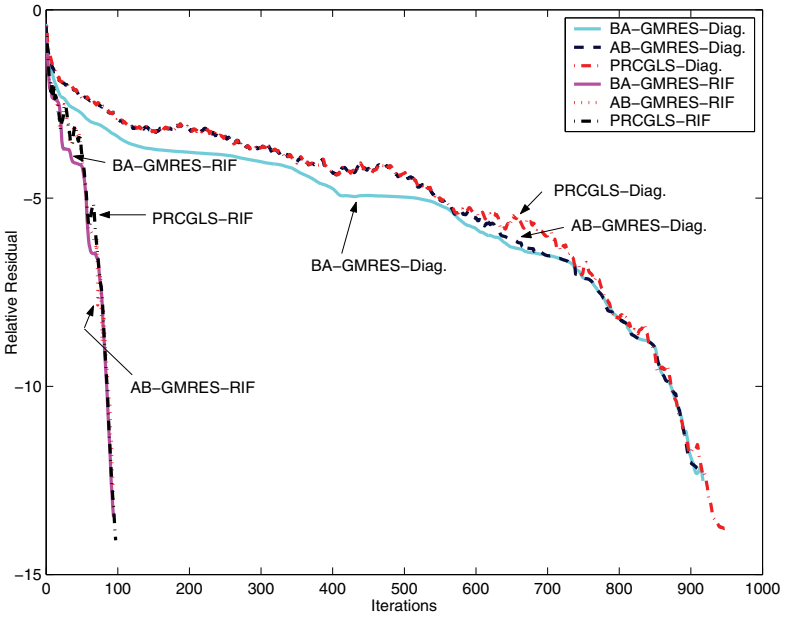

FIG. 6.3. *Comparison of AB-GMRES and BA-GMRES (RANDL3, $\tau = 0.8$ for RIF, $\|A^T\boldsymbol{r}\|_2/\|A^T\boldsymbol{b}\|_2$ vs. iterations, inconsistent problem).*

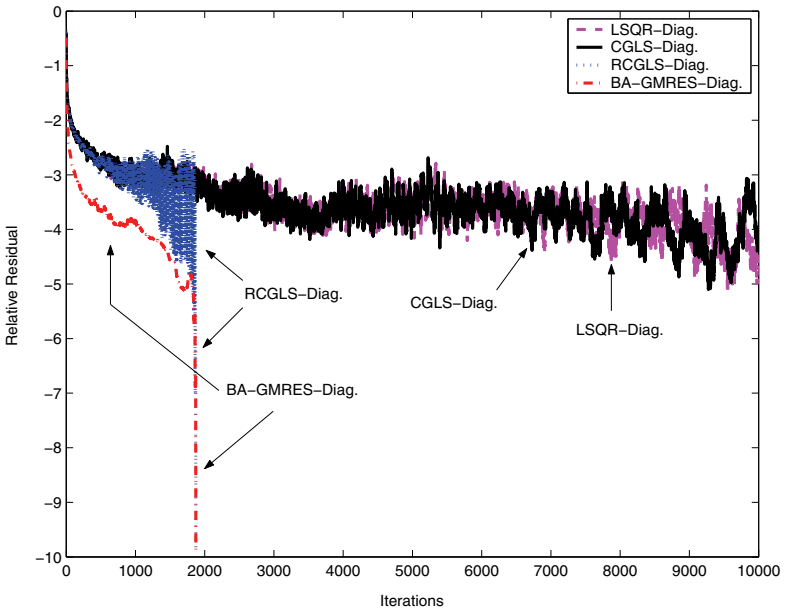

FIG. 6.4. *$\|A^T\boldsymbol{r}\|_2/\|A^T\boldsymbol{b}\|_2$ vs. iterations (RANDL6, diagonal scaling, inconsistent problem).*

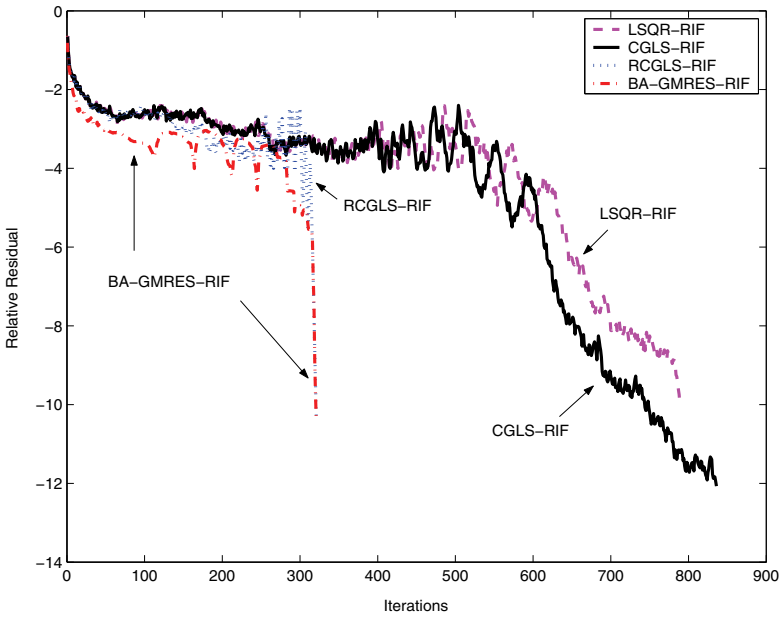

FIG. 6.5. $\|A^T\boldsymbol{r}\|_2/\|A^T\boldsymbol{b}\|_2$ *vs. iterations (RANDL6, RIF,* $\tau = 0.07$*, inconsistent problem).*

to be optimal).

AB-GMRES and PRCGLS show almost identical convergence, and BA-GMRES shows similar convergence, with both the diagonal scaling and the RIF preconditioner. This is in accordance with the convergence analysis in section 4.1. However, as mentioned in section 3.3, the amount of computation per iteration for BA-GMRES is less than AB-GMRES when $m \geq n$. Therefore, we compare BA-GMRES with preconditioned CGLS, LSQR, and RCGLS in the following experiments.

For BA-GMRES, we made experiments changing the restart cycle and observed that both the number of iterations and the computation time were the minimum for the full GMRES without restarts, although this requires more storage to store the orthogonal vectors. Hence, in the experiments below, we adopt the full BA-GMRES.

For the relative drop tolerance τ for the RIF preconditioner, we use the optimal τ for each problem, which minimizes the sum of preconditioning time and iteration time. When $\tau = 0$, the $A^{\mathrm{T}}A$ orthogonal factorization is complete and the iterative methods act as direct methods and converge in only one iteration. The preconditioned Krylov subspace iterative methods were faster than this "direct method."

Figures 6.4 and 6.5 show $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$ vs. the number of iterations for CGLS, LSQR, RCGLS, and BA-GMRES, with diagonal scaling and RIF preconditioners, respectively, for problem RANDL6 where the optimal τ is 0.07.

Figure 6.6 shows the relative error $\|x - x_*\|_2/\|x_*\|_2$ vs. the number of iterations for the RIF preconditioned methods for the same problem, where x_* is the converged solution.

The figures show that CGLS and LSQR are slow to converge compared to BA-GMRES and RCGLS for this ill-conditioned problem.

BA-GMRES and RCGLS show similar convergence behaviors. This is in accordance with the convergence analysis in section 4.1, where we obtained similar upper bounds for BA-GMRES and the similarly preconditioned CGLS in the absence of

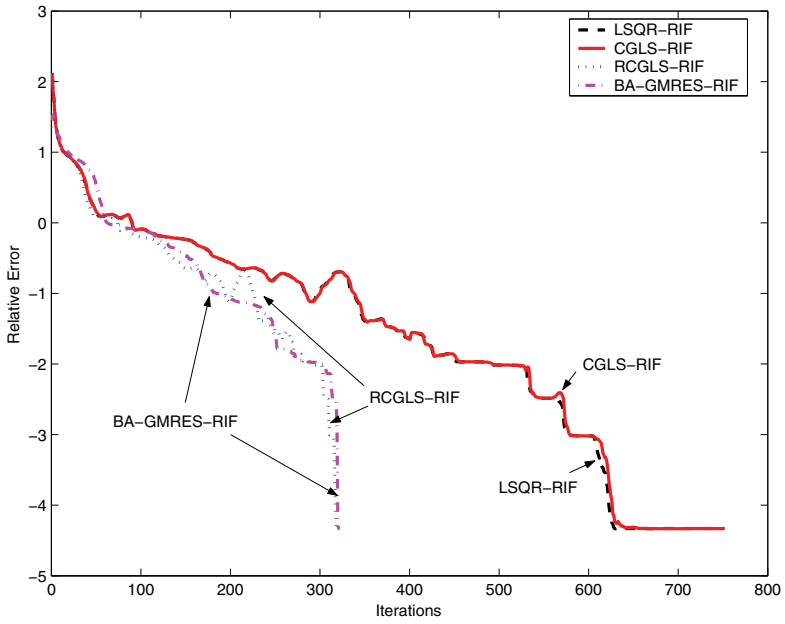

FIG. 6.6. $\|x - x_*\|_2/\|x_*\|_2$ *vs. iterations (RANDL6, RIF,* $\tau = 0.07$*, inconsistent problem).*

TABLE 6.2
Comparison of the iterative methods (inconsistent problem).

	CGLS		LSQR		RCGLS		BA-GMRES	
	-diag.	-RIF	-diag.	-RIF	-diag.	-RIF	-diag.	-RIF
RANDL1	35	14	35	14	35	14	35	14
($\tau = 0.5$)	∗0.10	4.98	∗0.10	4.98	0.16	4.99	0.14	4.99
RANDL2	214	21	214	21	208	21	193	21
($\tau = 0.7$)	∗0.61	5.10	∗0.61	5.10	2.77	5.13	2.28	5.11
RANDL3	742	72	740	73	697	70	622	60
($\tau = 0.8$)	2.08	5.38	∗2.07	5.38	26.30	5.62	20.70	5.47
RANDL4	1,147	85	1,154	85	1,062	84	1,069	82
($\tau = 0.5$)	∗3.22	6.38	3.38	6.38	59.35	7.17	59.41	6.62
RANDL5	4,897	470	5,064	401	1,521	305	1,522	299
($\tau = 0.9$)	13.74	12.78	14.03	∗11.79	119.70	14.42	118.87	13.91
RANDL6	10,551	615	11,088	645	1,861	317	1,862	318
($\tau = 0.07$)	29.65	33.67	30.21	34.15	177.94	26.93	176.93	∗26.73
RANDL7	32,143	1,951	35,034	2,443	1,914	371	1,899	362
($\tau = 0.02$)	89.93	102.28	91.31	128.40	195.63	40.07	183.90	∗37.26

First row: number of iterations, second row: computation time (seconds).
Convergence criterion: $\|A^T r\|_2/\|A^T b\|_2 < 10^{-6}$.

rounding errors. However, the BA-GMRES shows smoother convergence compared to RCGLS in $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$, which is a natural and practical measure of convergence for (inconsistent) least squares problems. This is because BA-GMRES minimizes $\|B\boldsymbol{r}\|_2 = \|CA^{\mathrm{T}}\boldsymbol{r}\|_2$, which is closer to $\|A^{\mathrm{T}}\boldsymbol{r}\|_2$, whereas PCGLS minimizes $\|\boldsymbol{r}|_{\mathcal{R}(A)}\|_2$. The convergence of the relative error is also shown to be a little smoother for the former.

RIF preconditioning significantly improves convergence over diagonal scaling.

In Table 6.2, we compare the methods for the problems in Table 6.1. The first row in each cell gives the number of iterations required for convergence, and the second

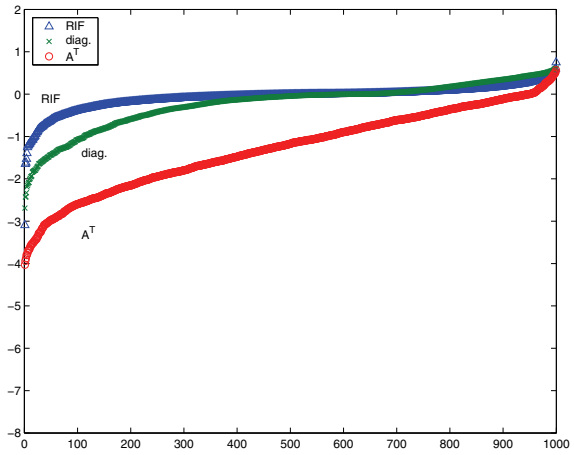

FIG. 6.7. *Eigenvalues* ($\log_{10}$ *scale) of* BA *for* $B = A^T$, $\mathrm{diag}(A^TA)^{-1}A^T$, *and* ZDZ^TA^T *with RIF* ($\tau = 0.1$) *for RANDM2.*

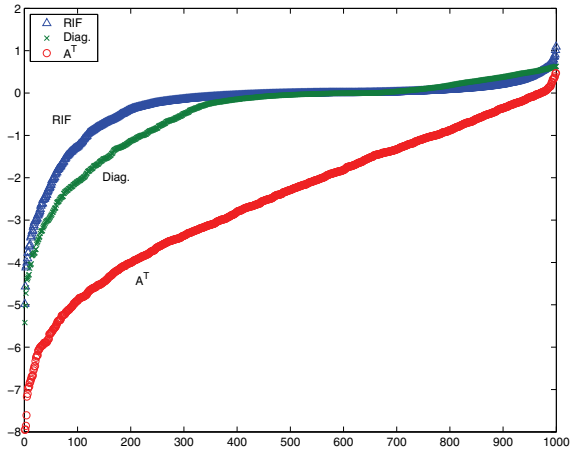

FIG. 6.8. *Eigenvalues* ($\log_{10}$ *scale) of* BA *for* $B = A^T$, $\mathrm{diag}(A^TA)^{-1}A^T$, *and* ZDZ^TA^T *with RIF* ($\tau = 0.1$) *for RANDM4.*

row gives the total computation time in seconds. The optimal relative drop tolerance τ for RIF preconditioning is also indicated for each problem. The fastest method for each problem is indicated by $*$.

For problems RANDL1 to RANDL4, CGLS or LSQR with diagonal scaling were the fastest. As the condition number increases, the number of iterations for CGLS (and its stabilized version, LSQR) increases much more rapidly than the correspondingly preconditioned BA-GMRES and RCGLS. For RANDL5, LSQR with RIF preconditioning is the fastest.

For the strongly ill-conditioned problems RANDL6 and RANDL7, BA-GMRES with RIF requires far less iterations than the CGLS with RIF, so that it is the fastest method with respect to computation time as well as the number of iterations. RCGLS performs similarly to BA-GMRES.

Figures 6.4, 6.5, and 6.6 showed that CGLS and LSQR are slow to converge

TABLE 6.3
Eigenvalues and condition number of BA.

Matrix	Preconditioner	$\lambda_1(BA)$	$\lambda_{50}(BA)$	$\lambda_{1,000}(BA)$	$\kappa(BA)$
RANDM2	A^{T}	1.16×10^{-4}	7.33×10^{-4}	4.01	3.46×10^{4}
	diag.	3.07×10^{-3}	5.00×10^{-2}	3.75	1.22×10^{3}
	RIF	1.22×10^{-2}	2.98×10^{-1}	5.35	4.39×10^{2}
RANDM4	A^{T}	1.01×10^{-8}	4.47×10^{-6}	3.42	3.39×10^{8}
	diag.	1.81×10^{-6}	1.30×10^{-3}	4.37	2.41×10^{6}
	RIF	2.07×10^{-6}	7.40×10^{-3}	11.80	4.45×10^{6}

TABLE 6.4
Number of iterations for each method (inconsistent problem).

	Preconditioner	CGLS	RCGLS	BA-GMRES
RANDM2	A^{T}	779	455	424
	diag.	180	174	176
	RIF	45	42	42
RANDM4	A^{T}	33,085	961	949
	diag.	2,003	887	884
	RIF	1,226	484	488

compared to BA-GMRES and RCGLS for the ill-conditioned problem RANDL6, and Table 6.2 shows that this phenomenon becomes more pronounced as the condition number of A increases.

To explain this phenomenon, we generated smaller random matrices RANDM2 and RANDM4 in a similar manner as the RANDLn matrices with $m = 10,000, n = 1,000$, density 1%, and $\kappa(A) = 1.75 \times 10^2$ and 1.58×10^4, respectively.

Figures 6.7 and 6.8 show the eigenvalue distribution of BA where $B = A^{\mathrm{T}}$, $\mathrm{diag}(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$, and $ZDZ^{\mathrm{T}}A^{\mathrm{T}}$ with RIF ($\tau = 0.1$), for RANDM2 and RANDM4, respectively. The eigenvalues coincide with the eigenvalues of the corresponding preconditioned CGLS matrices.

Table 6.3 gives the ith smallest eigenvalue $\lambda_i(BA)$ $(i = 1, 50, 1000)$ and the condition number $\kappa(BA)$ of BA for each case. Figures 6.7 and 6.8 and Table 6.3 show that as the condition number of A increases, the condition number of BA also increases, and at the same time, the eigenvalues of BA become more clustered towards the smallest eigenvalue λ_1, although both effects were substantially controlled by the use of RIF for B.

Next, Table 6.4 gives the number of iterations required for CGLS, RCGLS, and BA-GMRES with $B = A^{\mathrm{T}}$, diagonal scaling, and RIF to achieve $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2 < 10^{-6}$. Here, $\boldsymbol{x}_0 = \boldsymbol{0}$ and the elements of $\boldsymbol{b}$ were chosen randomly.

For RANDM4, the condition number of BA for diagonal scaling is less than that of RIF. However, RIF converges faster than diagonal scaling. This may be better explained by the fact that RIF makes the eigenvalues less clustered towards λ_1 and more clustered around 1.

The slow convergence of CGLS compared to RCGLS and BA-GMRES may be due to the loss of orthogonality of the vectors, which becomes pronounced when the preconditioned matrix becomes ill-conditioned and, moreover, when its eigenvalues become clustered towards λ_1 [29, 13].

Similarly, the slow convergence of CGLS and LSQR compared to RCGLS and BA-GMRES in the RANDLn matrices may be explained by the fact that CGLS and LSQR rely on three-term recurrence and suffer from loss of orthogonality of the

TABLE 6.5
Properties of matrices from the University of Florida Sparse Matrix Collection.

Matrix	m	n	nnz	Rank
lp_brandy	303	220	2,202	193
lp_standgub	1,383	361	3,338	360
lp_cycle	3,371	1,903	21,234	1,875
lpi_greenbea	5,596	2,393	31,074	2,390
nemspmm1	8,903	2,372	55,867	2,362
lpi_gosh	13,455	3,792	99,953	3,790
stormg2-27	37,485	14,441	94,274	14,387
stormg2-125	172,431	66,185	433,256	65,935

TABLE 6.6
Comparison of the iterative methods (inconsistent problem).

Matrix	CGLS	LSQR	RCGLS	BA-GMRES
lp_brandy	591	587	91	88
	0.05	0.05	$*$ 0.03	$*$ 0.03
lp_standgub	191	186	71	68
	$*$ 0.03	$*$ 0.03	0.05	0.04
lp_cycle	20,448	20,417	397	374
	16.94	16.32	5.15	$*$ 4.56
lpi_greenbea	2,361	2,342	440	418
	$*$ 2.23	2.68	8.35	7.22
nemspmm1	1,797	2,078	220	187
	2.64	3.59	2.05	$*$ 1.58
lpi_gosh	6,529	6,788	431	404
	15.99	19.47	12.56	$*$ 11.04
stormg2-27	962	947	146	141
	$*$ 4.76	6.18	5.70	5.21
stormg2-125	953	947	137	130
	25.41	33.42	25.47	$*$ 22.94

First row: number of iterations, second row: computation time (seconds).
Convergence criterion: $\|A^{\mathrm{T}}r\|_2/\|A^{\mathrm{T}}b\|_2 < 10^{-8}$.

vectors due to rounding errors when the (preconditioned) system is ill-conditioned and its eigenvalues are clustered towards λ_1, whereas BA-GMRES and RCGLS are more robust against loss of orthogonality because they perform explicit orthogonalization by the MGS procedure and reorthogonalization, respectively.

Finally, we test a class of rectangular matrices from the University of Florida Sparse Matrix Collection [12]. We construct overdetermined systems by transposing the original matrix, where the number of rows m, columns n, and nonzero elements nnz, as well as the rank, are given in Table 6.5.

The matrices are rank-deficient, although the rank deficiency is due to zero columns of the matrix, except for problem "lp_cycle," where the number of columns after deleting all the zero columns is $1,890$, which is larger than the rank of the original matrix, $1,875$.

For these problems, we compare the convergence of CGLS, LSQR, RCGLS, and BA-GMRES without any preconditioners; i.e., $C = \mathrm{I}_n$ and $B = A^{\mathrm{T}}$, which satisfies the condition in Corollary 3.19. The right-hand side vectors were chosen to be vectors whose elements are all one, and the problems are inconsistent. The numerical results are given in Table 6.6, where the first row in each cell gives the number of iterations required for convergence, and the second row gives the total computation time in seconds.

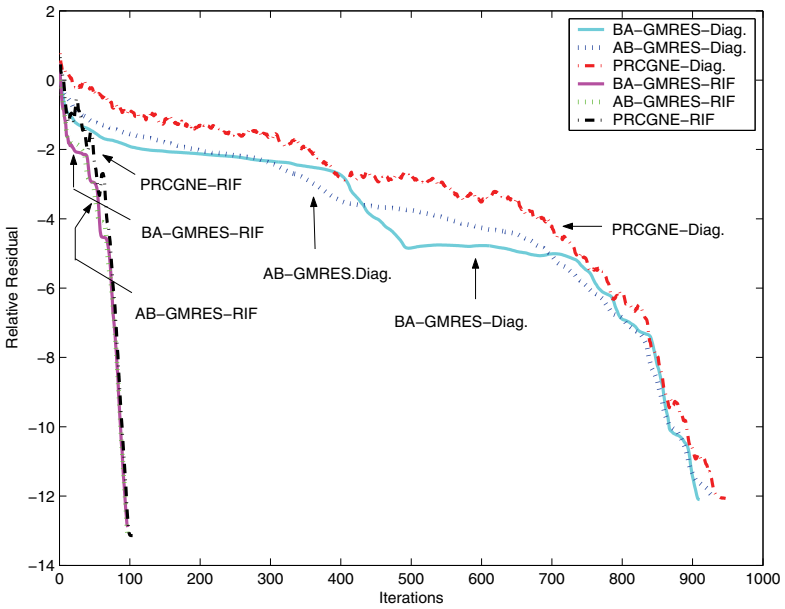

FIG. 6.9. *Comparison of AB-GMRES and BA-GMRES (RANDL3T, $\tau = 0.8$ for RIF, $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ vs. iterations, consistent problem).*

Table 6.6 shows that BA-GMRES requires fewer iterations than CGLS, LSQR, and RCGLS. Also, BA-GMRES is fastest with respect to the total CPU time for problems "lp_cycle," "nemspmm1," "lpi_gosh," and "stormg2-125."

6.2. Underdetermined problems. Next, we show numerical experiment results for the underdetermined least squares problem

$$\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2, \qquad A \in \mathbf{R}^{m\times n} \qquad (m < n).$$

Here, we compare the AB-GMRES (BA-GMRES) methods with $B = A^{\mathrm{T}}C$, where C is obtained by diagonal scaling or RIF, with the corresponding preconditioned CGNE, LSQR, and RCGNE (reorthogonalized CGNE based on Method 1.2 with $l = \infty$) methods.

The reason we compare with PRCGNE instead of PRCGLS is because, for the reorthogonalized version, PRCGNE requires less memory and work per iteration compared to PRCGLS, since the former works in m-dimensional space, whereas the latter works in n-dimensional ($m < n$.) Also, for $m < n$, $A^{\mathrm{T}}A$ is singular (positive semidefinite), and some preconditioners may not be defined for RCGLS, when it can be defined for RCGNE. However, PRCGLS has the advantage that the residual norm $\|r\|$ decreases monotonically.

The matrices A were obtained by transposing the matrices RANDLn in Table 6.1 for the overdetermined problems, and are denoted by RANDLnT. The density and condition number of RANDLnT is the same for the corresponding RANDLn.

Since $\operatorname{rank} A = m < n$, the systems are consistent. The right-hand side $\boldsymbol{b}$ was given by $\boldsymbol{b} = A\boldsymbol{x}^*$, where $\boldsymbol{x}^* = (1, \ldots, 1)^{\mathrm{T}}$, and the initial approximate solution was set to $\boldsymbol{x}_0 = \boldsymbol{0}$. The convergence of the methods were judged by $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ where $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}$.

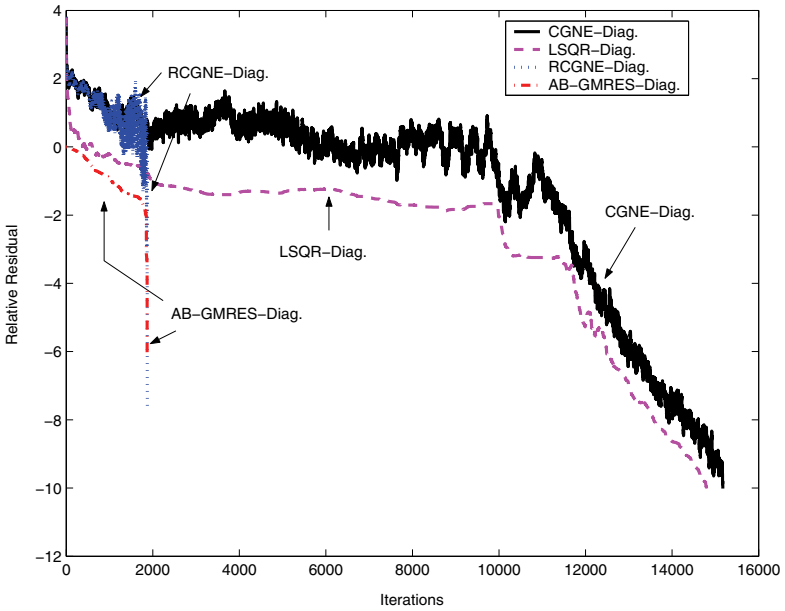

FIG. 6.10. $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ *vs. iterations (RANDL6T, diagonal scaling, consistent problem).*

Note that for AB-GMRES, $\|\boldsymbol{r}\|_2$ is available at each iteration without extra computational cost, unlike the BA-GMRES.

Figure 6.9 shows $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ vs. the number of iterations for AB-GMRES, BA-GMRES ($k = \infty$), and PRCGNE (preconditioned, fully reorthogonalized CGNE) with diagonal scaling and RIF preconditioning, for problem RANDL3T. τ for RIF was set to the optimal value 0.8. In the figure, AB-GMRES, BA-GMRES, and PRCGNE show similar convergence behavior, as predicted in section 4.2.

In the experiments below, we compare AB-GMRES with CGNE, LSQR, and RCGNE. The reason AB-GMRES is used instead of BA-GMRES is because the former gives the minimum-norm solution, and requires less computation per iteration when $m < n$ (see section 3.3). Full AB-GMRES without restart was used. The optimal τ was chosen for the RIF preconditioner for all the problems.

With all the methods, the approximate solutions converged to the minimum-norm least squares solution.

Figures 6.10 and 6.11 show the relative residual $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ vs. the number of iterations for the different methods for problem RANDL6T. τ for RIF was set to the optimal value 0.01. The figures show that AB-GMRES and RCGNE converge faster than the corresponding CGNE and LSQR, for diagonal scaling and RIF. AB-GMRES shows smoother convergence in $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ compared to RCGNE. This is because AB-GMRES minimizes $\|\boldsymbol{r}\|_2$, whereas RCGNE minimizes $\|\boldsymbol{r}\|_{(AA^{\mathrm{T}})^{-1}}$ (cf. section 4.2).

Table 6.7 compares the methods for underdetermined systems RANDLnT. The first row in each box gives the number of iterations required for convergence, and the second row gives the total computation time in seconds. It is observed that the AB-GMRES-RIF is the fastest method for the ill-conditioned problems RANDL6T and RANDL7T, and RCGNE-RIF performs similarly.

6.3. Required memory and work. One drawback of the GMRES based methods is that they require increasingly more memory with the number of iterations or

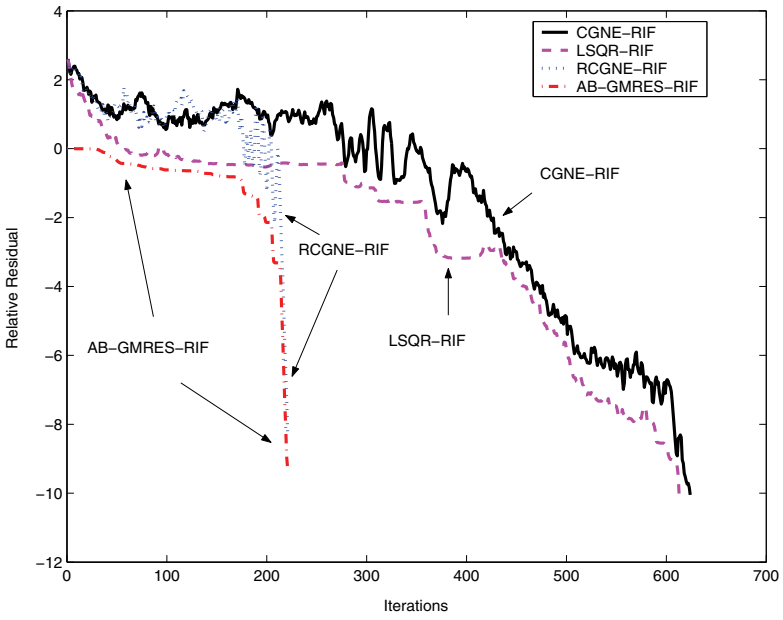

FIG. 6.11. $\|\boldsymbol{r}\|_2/\|\boldsymbol{b}\|_2$ *vs. iterations (RANDL6T, RIF,* $\tau = 0.01$*, consistent problem).*

TABLE 6.7
Comparison of the iterative methods (consistent problem).

	CGNE		LSQR		RCGNE		AB-GMRES	
	-diag.	-RIF	-diag.	-RIF	-diag.	-RIF	-diag.	-RIF
RANDL1T	37	12	37	12	37	12	37	12
($\tau = 0.4$)	∗0.15	5.01	∗0.15	5.02	0.23	5.02	0.17	5.01
RANDL2T	259	25	256	25	252	25	247	25
($\tau = 0.7$)	∗1.06	5.16	∗1.06	5.16	4.22	5.17	3.75	5.17
RANDL3T	838	81	823	80	783	77	754	75
($\tau = 0.8$)	∗3.43	5.53	∗3.43	5.53	33.87	5.81	30.56	5.67
RANDL4T	1,464	116	1,407	114	1,223	106	1,187	106
($\tau = 0.5$)	5.97	6.94	∗5.96	6.94	79.67	7.45	73.67	7.21
RANDL5T	5,548	544	5,414	539	1,535	322	1,533	322
($\tau = 0.9$)	22.61	13.88	22.71	∗13.86	123.80	15.58	121.58	15.11
RANDL6T	12,837	514	12,486	502	1,873	219	1,871	218
($\tau = 0.01$)	52.38	39.04	50.10	38.99	182.51	27.43	179.85	∗26.99
RANDL7T	38,397	3,078	37,792	2,979	2,240	451	2,238	450
($\tau = 0.04$)	156.01	94.19	152.07	91.69	366.88	44.49	353.92	∗43.76

First row: number of iterations, second row: computation time (seconds).
Convergence criterion: $\|r\|_2/\|b\|_2 < 10^{-6}$.

the restarting cycle k, whereas the CG based methods (without reorthogonalization) require constant memory. This is because the GMRES based methods require storing the orthonormal vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ in the MGS process, as well as the Hessenberg matrix.

Table 6.8 shows the memory required other than A and the preconditioner for the kth iteration for each method. $\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{e}$, and $\boldsymbol{y}$ are the intermediate vectors, and V denotes the k orthonormal vectors in the MGS process. For LSQR, we used the notation of variables according to [23].

In addition, Table 6.9 shows the work for the kth iteration of the unpreconditioned

TABLE 6.8
Intermediate memory required for the kth iteration for each method.

	dim(m)	dim(n)	dim(k)	Total
CGLS(CGNE)	$\boldsymbol{r}, A\boldsymbol{p}$	$\boldsymbol{x}, \boldsymbol{p}, A^{\mathrm{T}}\boldsymbol{r}$		$2m+3n$
LSQR	$\boldsymbol{u}$	$\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{x}$		$m+3n$
RCGLS	$\boldsymbol{r}, A\boldsymbol{p}$	$\boldsymbol{x}, \boldsymbol{p}, \{A^{\mathrm{T}}\boldsymbol{r}_i\}_{i=1}^{k+1}$		$2m+(k+3)n$
RCGNE	$\{\boldsymbol{r}_i\}_{i=1}^{k+1}, A\boldsymbol{p}$	$\boldsymbol{x}, \boldsymbol{p}, A^{\mathrm{T}}\boldsymbol{r}$		$(k+2)m+3n$
AB-GMRES	$\{\boldsymbol{v}_i\}_{i=1}^{k}, \boldsymbol{w}$	$\boldsymbol{x}$	$\boldsymbol{y}, \boldsymbol{e}, H_k$	$(k+1)m+n+k^2/2+2k$
BA-GMRES		$\{\boldsymbol{v}_i\}_{i=1}^{k}, \boldsymbol{w}, \boldsymbol{x}$	$\boldsymbol{y}, \boldsymbol{e}, H_k$	$(k+2)n+k^2/2+2k$

TABLE 6.9
Work for the kth iteration for each method.

	MV	NRM2	SAXPY
CGLS(CGNE)	2	2	3
LSQR	2	2	4
RCGLS(RCGNE)	2	$k+1$	$k+2$
BA-GMRES(AB-GMRES)	2	$k+1$	$k+1$

CGLS (CGNE), LSQR, RCGLS (RCGNE), and full BA-GMRES (AB-GMRES), where MV, NRM2, and SAXPY denote the computational work of a matrix-vector multiplication, a norm or inner product of a vector, and an operation computing a constant α times a vector $\boldsymbol{x}$ plus a vector $\boldsymbol{y}$, respectively. These operations are implemented as a basic routine in BLAS (basic linear algebra subprograms). Here, we omit the computation of scalar operations, e.g., scalar division in CGLS and computation of Givens rotations in LSQR.

For the preconditioned case, assuming that $M = LL^{\mathrm{T}}$ and $B = MA^{\mathrm{T}} = LL^{\mathrm{T}}A^{\mathrm{T}}$, two additional matrix-vector multiplications are introduced for all the methods.

If one can keep the number of iterations k of the BA-(AB-)GMRES and RCGLS (RCGNE) sufficiently small compared to m or n with the use of an efficient preconditioner like RIF, they may be faster compared to CGLS (CGNE) or LSQR, and the required memory may not be prohibitive.

7. Conclusions. We proposed two methods for applying GMRES to linear least squares problems with $m \times n$ matrix A, using an $n \times m$ matrix B. AB-GMRES applies GMRES to $\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$, while BA-GMRES applies GMRES to $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|B\boldsymbol{b} - BA\boldsymbol{x}\|_2$.

Then, we showed that if $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ and $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$, AB-GMRES and BA-GMRES determine a least squares solution of $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ for all $\boldsymbol{b} \in \mathbf{R}^m$ and all $\boldsymbol{x}_0 \in \mathbf{R}^n$ without breakdown.

Next, we showed that, theoretically, one may expect similar convergence behaviors for the AB- and BA-GMRES as well as the corresponding CGLS type methods.

Further, we proposed using the RIF method for B in the GMRES methods, as well as for CGLS and CGNE.

Numerical experiments on overdetermined problems showed that BA-GMRES and the fully reorthogonalized CGLS (RCGLS) with the RIF preconditioner was faster than the RIF preconditioned CGLS and LSQR for ill-conditioned problems. BA-GMRES gave smoother convergence compared to preconditioned RCGLS in $\|A^{\mathrm{T}}\boldsymbol{r}\|_2$, which is a natural norm for monitoring convergence, especially for inconsistent problems.

For underdetermined problems with full row rank, AB-GMRES and the fully reorthogonalized CGNE (RCGNE) with RIF preconditioning were faster than the RIF

preconditioned CGNE and LSQR for ill-conditioned problems. AB-GMRES gave smoother convergence compared to preconditioned RCGNE in $\|\boldsymbol{r}_k\|_2$. The preconditioned RCGLS would give monotonic decrease of $\|\boldsymbol{r}_k\|_2$ when the preconditioner can be defined, with the expense of more memory and computation per iteration.

Another advantage of the BA- or AB-GMRES over preconditioned RCGLS or RCGNE is that the former have the freedom of employing a nonsymmetric matrix C for the preconditioner as $B = CA^{\mathrm{T}}$, whereas the latter require a symmetric matrix C for the preconditioner.

Acknowledgments. We would like to thank Professors Shao-Liang Zhang, Kokichi Sugihara, Yimin Wei, Yohsuke Hosoda, Kunio Tanabe, Takashi Tsuchiya, Michael Eiermann, Masaaki Sugihara, Zhong-Zhi Bai, Michele Benzi, Lothar Reichel, Gene Golub, Michael Saunders, and Yang-Feng Su for valuable remarks and discussions.

We would also like to thank the referees for valuable remarks which helped to greatly improve this paper.

REFERENCES

[1] Z.-Z. Bai, I. S. Duff, and A. J. Wathen, *A class of incomplete orthogonal factorization methods*. I: *Methods and theories*, BIT, 41 (2001), pp. 53–70.

[2] M. Benzi and M. Tůma, *A robust incomplete factorization preconditioner for positive definite matrices*, Numer. Linear Algebra Appl., 10 (2003), pp. 385–400.

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

[4] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[5] Å. Björck, T. Elfving, and Z. Strakoš, *Stability of conjugate gradient and Lanczos methods for linear least squares problems*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 720–736.

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

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

[8] X. Cui and K. Hayami, *Greville's Method for Preconditioning Least Squares Problems*, NII Technical Reports, National Institute of Informatics, Tokyo, NII-2008-008E (2008), pp. 1–26.

[9] X. Cui, K. Hayami, and J.-F. Yin, *Greville's method for preconditioning least squares problems*, in Proceedings of the ALGORITMY 2009, Podbanské, Slovakia, 2009, pp. 440–448.

[10] X. Cui, K. Hayami, and J.-F. Yin, *Greville's method for preconditioning least squares problems*, Adv. Comput. Math., accepted.

[11] G. Dahlquist, B. Sjöberg, and P. Svensson, *Comparison of the method of averages with the method of least squares*, Math. Comp., 22 (1968), pp. 833–845.

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

[13] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers Appl. Math. 17, SIAM, Philadelphia, 1997.

[14] K. Hayami and T. Ito, *Application of the GMRES method to singular systems and least squares problems*, in Proceedings of the Seventh China-Japan Seminar on Numerical Mathematics and Scientific Computing, Z.-C. Shi and H. Okamoto, eds., Zhangjiajie, 2004, Science Press, Beijing, 2006, pp. 31–44.

[15] K. Hayami and T. Ito, *Solution of least squares problems using GMRES methods*, Proc. Inst. Statist. Math., 53 (2005), pp. 331–348 (in Japanese).

[16] K. Hayami and M. Sugihara, *A geometric view of Krylov subspace methods on singular systems*, NII Technical Reports, National Institute of Informatics, NII-2009-007E, Tokyo, pp. 1–28 (revised version to appear in Numer. Linear Algebra Appl.).

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

[18] T. Ito and K. Hayami, *Preconditioned GMRES methods for least squares problems*, Japan. J. Indust. Appl. Math., 25 (2008), pp. 185–207.

[19] A. Jennings and M. A. Ajiz, *Incomplete methods for solving $A^TAx = b$*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 978–987.
[20] R. M. Larsen, *Lanczos Bidiagonalization with Partial Reorthogonalization*, Technical Report DAIMI PB-357, Department of Computer Science, Aarhus University, Aarhus, Denmark, 1998.
[21] J. A. Meijerink and H. A. van der Vorst, *An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix*, Math. Comp., 31 (1977), pp. 148–162.
[22] C. C. Paige, M. Rozložník, and Z. Strakoš, *Modified Gram-Schmidt (MGS), least squares, and backward stability of MGS-GMRES*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 264–284.
[23] C. C. Paige and M. A. Saunders, *LSQR: An algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software, 8 (1982), pp. 43–71.
[24] A. T. Papadopoulos, I. S. Duff, and A. J. Wathen, *A class of incomplete orthogonal factorization methods*. II: *Implementation and results*, BIT, 45 (2005), pp. 159–179.
[25] L. Reichel and Q. Ye, *Breakdown-free GMRES for singular systems*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 1001–1021.
[26] Y. Saad, *Preconditioning techniques for nonsymmetric and indefinite linear systems*, J. Comput. Appl. Math., 24 (1988), pp. 89–105.
[27] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[28] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual method for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[29] Z. Strakoš, *On the real convergence rate of the conjugate gradient method*, Linear Algebra Appl., 154-156 (1991), pp. 535–549.
[30] K. Tanabe, *Characterization of linear stationary iterative processes for solving a singular system of linear equations*, Numer. Math., 22 (1974), pp. 349–359.
[31] K. Vuik, A. G. J. Sevink, and G. C. Herman, *A preconditioned Krylov subspace method for the solution of least squares problems in inverse scattering*, J. Comput. Phys., 123 (1996), pp. 330–340.
[32] X. Wang, K. A. Gallivan, and R. Bramley, *CIMGS: An incomplete orthogonal factorization preconditioner*, SIAM J. Sci. Comput., 18 (1997), pp. 516–536.
[33] S.-L. Zhang and Y. Oyanagi, *Orthomin(k) method for linear least squares problem*, J. Inform. Process., 14 (1991), pp. 121–125.