Journal of Computational and Applied Mathematics 290 (2015) 310–318

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Some properties of range restricted GMRES methods

M. Bellalij [a], L. Reichel [b,*], H. Sadok [c]

[a] Laboratoire de Mathématiques et Leurs Applications, Université de Valenciennes, Le Mont Houy, 59313 Valenciennes cedex, France

[b] Department of Mathematical Sciences, Kent State University, Kent, OH 44242, United States

[c] Laboratoire de Mathématiques Pures et Appliquées, Université du Littoral, Centre Universtaire de la Mi-Voix, Batiment H. Poincarré, 50 Rue F. Buisson, BP 699, 62228 Calais cedex, France

ARTICLE INFO

Article history:
Received 11 August 2014
Received in revised form 8 November 2014

Keywords:
Ill-posed problem
Iterative method
Truncated iteration
GMRES
RRGMRES
Shifted GMRES

ABSTRACT

The GMRES method is one of the most popular iterative schemes for the solution of large linear systems of equations with a square nonsingular matrix. GMRES-type methods also have been applied to the solution of linear discrete ill-posed problems. Computational experience indicates that for the latter problems variants of the standard GMRES method, that require the solution to live in the range of a positive power of the matrix of the linear system of equations to be solved, generally yield more accurate approximations of the desired solution than standard GMRES. This paper investigates properties of these variants of GMRES.

1. Introduction

This paper studies properties of a family of minimal residual iterative methods for the computation of approximate solutions of linear systems of equations

$$Ax = b, \quad A \in \mathbb{R}^{n\times n},\ x, b \in \mathbb{R}^n, \tag{1.1}$$

with a large nonsymmetric matrix A whose singular values gradually decay to zero (without a significant gap). In particular, A is severely ill-conditioned and may be singular. Linear systems of equations with a matrix of this kind are commonly referred to as linear discrete ill-posed problems. They arise, for instance, from the discretization of linear ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel.

In many linear discrete ill-posed problems with applications in science and engineering, the right-hand side vector b represents data that is contaminated by error. Thus,

$$b = \widehat{b} + e, \tag{1.2}$$

where $\widehat{b} \in \mathbb{R}^n$ denotes the unknown error-free right-hand side associated with b. We will refer to the error vector $e \in \mathbb{R}^n$ as "noise". The noise vector may stem from measurement inaccuracies and from discretization errors.

We would like to compute the solution of minimal Euclidean norm, $\widehat{x}$, of the linear discrete ill-posed problem with the unknown error-free right-hand side

$$Ax = \widehat{b}.$$

This system is assumed to be consistent.

* Corresponding author.
E-mail addresses: Mohammed.Bellalij@univ-valenciennes.fr (M. Bellalij), reichel@math.kent.edu (L. Reichel), sadok@lmpa.univ-littoral.fr (H. Sadok).

http://dx.doi.org/10.1016/j.cam.2015.05.008
0377-0427/

Since $\widehat{b}$ is not known, we seek to determine an accurate approximation of $\widehat{x}$ by computing an approximate solution of the available linear system of Eqs. (1.1). Due to the severe ill-conditioning of the matrix A and the error e in b, the least-squares solution of minimal Euclidean norm of (1.1) generally does not furnish a useful approximation of $\widehat{x}$.

A popular approach to determine a meaningful approximation of $\widehat{x}$ is to apply an iterative method to the solution of (1.1) and terminate the iterations sufficiently early. This type of solution approach is known as truncated iteration.

GMRES is a popular iterative method for the solution of large linear systems of equations with a nonsymmetric matrix A that arise from the discretization of well-posed problems, such as Dirichlet boundary value problems for elliptic partial differential equations; see, e.g., Saad [1]. Let the initial iterate be $x_0 = 0$. Then the ℓth iterate, x_ℓ, computed by GMRES applied to the solution of (1.1) satisfies

$$\|Ax_\ell - b\| = \min_{x\in\mathbb{K}_\ell(A,b)} \|Ax - b\|, \quad x_\ell \in \mathbb{K}_\ell(A, b),$$

where

$$\mathbb{K}_\ell(A, b) = \text{span}\{b, Ab, \ldots, A^{\ell-1}b\}$$

is a Krylov subspace and $\|\cdot\|$ denotes the Euclidean vector norm. We will tacitly assume that ℓ is sufficiently small so that $\dim(\mathbb{K}_\ell(A, b)) = \ell$; see below for some comments on this assumption.

When applied to the solution of linear discrete ill-posed problems with an error-contaminated right-hand side (1.1), GMRES may determine more accurate approximations of $\widehat{x}$ with less arithmetic work than the conjugate gradient method applied to the normal equations associated with (1.1); see [2–4] for illustrations. For linear discrete ill-posed problems, for which the desired solution $\widehat{x}$ is the discretization of a smooth function, it has been observed in [3,5,6] that the following variants of GMRES, which we refer to as m-shifted GMRES, often deliver more accurate approximations of $\widehat{x}$ than the standard GMRES method.

With the initial iterate $x_0 = 0$, the ℓth iterate, $x_\ell^{(m)}$, determined by m-shifted GMRES satisfies

$$\|Ax_\ell^{(m)} - b\| = \min_{x\in\mathbb{K}_\ell(A,A^m b)} \|Ax - b\|, \quad x_\ell^{(m)} \in \mathbb{K}_\ell(A, A^m b). \tag{1.3}$$

When $m \geq 1$ these are range restricted GMRES (RRGMRES) methods, because the iterates generated are in the range of A, which is denoted by $\mathcal{R}(A)$. Some properties of and computed examples with one-shifted GMRES can be found in [3,7]. The one-shifted GMRES method is in these references referred to as the RRGMRES method. A new implementation of one-shifted GMRES with improved numerical behavior recently has been described in [8,9]. An extension of this implementation to m-shifted GMRES for $m > 1$ is presented in [5].

It is the purpose of the present paper to analyze the m-shifted GMRES method. Some of the properties we show for m-shifted GMRES are analogues of properties that have been established in [10] for standard GMRES, but have not previously been shown for m-shifted GMRES; other properties we derive are particular for m-shifted GMRES with $m \geq 1$. Our analysis sheds light on the performance of m-shifted GMRES. The good performance of m-shifted GMRES for $m = 2$ and $m = 3$, which is illustrated in Section 3, makes it desirable and important to better understand the features of the method.

One can show that the sensitivity of the iterates $x_\ell^{(m)}$ determined by (1.3) increases with the number of iterations ℓ. The norm $\|x_\ell^{(m)} - \widehat{x}\|$ typically decreases as ℓ increases and is small, but increases with ℓ for ℓ large. This behavior of the iterates is commonly referred to as semiconvergence. The growth of $\|x_\ell^{(m)} - \widehat{x}\|$ for large values of ℓ is caused by severe propagation of the error e in b and of round-off errors introduced during the computations. It is important to terminate the iterations when x_ℓ is close to $\widehat{x}$.

When an estimate of the norm of e is available, the discrepancy principle can be used to determine how many iterations to carry out. Truncated iteration with the standard GMRES method based on the discrepancy principle is analyzed in [11]. We apply the discrepancy principle in the computed examples of this paper. Further discussions on the discrepancy principle can be found in [12].

If no estimate of $\|e\|$ is available, then it may be possible to compute an estimate, which can be used in the discrepancy principle; see [13] for illustrations. Alternatively, one may apply one of the many so-called "heuristic" stopping rules described in the literature for determining a suitable number of iterations; see, e.g., [14–16] for illustrations and discussions.

We remark that GMRES also can be used in a different way than described above to solve linear discrete ill-posed problems. For instance, Matinfar et al. [17] illustrate the application of several GMRES-type methods to the solution of Tikhonov regularized linear discrete ill-posed problems,

$$(A^TA + \mu I)x = A^Tb.$$

Here and below the superscript ^{T} denotes transposition. Regularization is achieved by choosing a suitable value of the regularization parameter $\mu > 0$, not by truncated iteration. Related methods are described in [18–21]. We will not discuss this approach further and instead focus on GMRES-type regularization methods for which regularization is achieved by truncated iteration.

This paper is organized as follows. Section 2 presents an analysis of m-shifted GMRES. A few computed examples are described in Section 3, and concluding remarks can be found in Section 4.

2. Properties of *m*-shifted GMRES

This section shows new results about m-shifted GMRES methods. When solving linear discrete ill-posed problems (1.1) by an iterative method, the initial iterate is generally chosen to be $x_0 = 0$, because this iterate is orthogonal to the null space of the matrix A. However, since our results do not depend on the choice of initial iterate, except that the choice of x_0 may affect when breakdown takes place, we formulate them for an arbitrary initial iterate. The associated residual vector is denoted by $r_0 = b - Ax_0$.

The ℓth iterate $x_\ell^{(m)}$ determined by m-shifted GMRES applied to the solution of

$$Ax = r_0 \tag{2.1}$$

satisfies

$$\|Ax_\ell^{(m)} - r_0\| = \min_{x \in \mathbb{K}_\ell(A, A^m r_0)} \|Ax - r_0\|$$

and

$$x_\ell^{(m)} \in \mathbb{K}_\ell(A, A^m r_0) = \operatorname{span}\{A^m r_0, A^{m+1} r_0, \ldots, A^{m+\ell-1} r_0\}. \tag{2.2}$$

In the problems of interest to us, the matrix A is a low-pass filter. The reason for choosing $m \geq 1$ is to determine a solution subspace (2.2) that is comprised of vectors that can be considered discretizations of smooth functions. Assume for the moment that A has the spectral resolution

$$A = S\Lambda S^{-1}, \qquad \Lambda = \operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n], \qquad S = [s_1, s_2, \ldots, s_n],$$

with the eigenvalues ordered so that $|\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_n| \leq 1$, and unit eigenvectors s_j. Expand r_0 in terms of the eigenvectors

$$r_0 = \sum_{k=1}^{n} \alpha_k s_k.$$

Then

$$A^m r_0 = \sum_{k=1}^{n} \alpha_k \lambda_k^m s_k.$$

Thus, when $|\lambda_k| \ll 1$, the eigenvector component s_k in r_0 is damped considerably by multiplying r_0 by A^m for $m \geq 1$, and the damping is more pronounced the larger m is. For problems of interest to us, eigenvectors associated with eigenvalues of large magnitude can be thought of as discretizations of smooth functions, while eigenvectors associated with eigenvalues of "tiny" magnitude can be considered discretizations of highly oscillatory functions. When the desired solution $\widehat{x}$ is known to be a discretization of a smooth function, it is meaningful to let $m \geq 1$ so that the Krylov subspace (2.2) is well suited to represent the discretization of such a function. In particular, the larger the error e in the given data, the larger m should be chosen. This is illustrated in Section 3. We remark that the above assumption that A be diagonalizable is not essential for the discussion and can be removed.

Iteration with m-shifted GMRES is said to break down at step $\ell + 1$ if ℓ is the smallest positive integer such that

$$\dim(\mathbb{K}_{\ell+1}(A, A^m r_0)) = \dim(\mathbb{K}_\ell(A, A^m r_0)) = \ell. \tag{2.3}$$

It is well-known that for standard GMRES breakdown at step $\ell + 1$ indicates that the solution of (1.1) lives in $\mathbb{K}_\ell(A, r_0)$; see [1]. We show an analogous result for m-shifted GMRES. It is convenient to introduce the matrix

$$K_\ell^{(m)} = [A^m r_0, A^{m+1} r_0, \ldots, A^{m+\ell-1} r_0] \in \mathbb{R}^{n \times \ell}. \tag{2.4}$$

Proposition 2.1. *Let ℓ be the smallest positive integer such that (2.3) holds. Then there is a matrix $H_{\ell,\ell} \in \mathbb{R}^{\ell \times \ell}$ such that*

$$AK_\ell^{(m)} = K_\ell^{(m)} H_{\ell,\ell}. \tag{2.5}$$

Proof. The assertion follows from the fact that $\mathbb{K}_\ell^{(m)}(A, A^m r_0)$ is an invariant subspace of A. □

Theorem 2.2. *Let A be nonsingular and let ℓ be the smallest positive integer such that (2.3) holds. Then the solution of (2.1) lives in the Krylov subspace $\mathbb{K}_\ell(A, A^m r_0)$. In particular, the solution can be computed at breakdown.*

Proof. The relation (2.5) holds. Since the matrices A and $K_\ell^{(m)}$ are of full rank, it follows that $H_{\ell,\ell}$ is nonsingular.

Consider the linear system of equations

$$Ax = K_\ell^{(m)} e_1,$$

where $e_1 = [1, 0, \ldots, 0]^T$ denotes the first axis vector. The solution of this system is given by $x^{(m)} = A^{m-1}r_0$. We are interested in the vector $A^{-m}x^{(m)}$, the solution of (2.1). It follows from (2.5) that

$$A^{-m}x^{(m)} = A^{-m+1}A^{-1}K_\ell^{(m)}e_1 = A^{-m+1}K_\ell^{(m)}H_{\ell,\ell}^{-1}e_1 = \cdots = K_\ell^{(m)}H_{\ell,\ell}^{-m}e_1.$$

This shows that the solution of (2.1) lives in $\mathbb{K}_\ell(A, A^m r_0)$. □

We turn to the situation when A is singular. Then it might not be possible to solve (1.1). The following result shows that it may be possible to compute a least-squares solution of (1.1) at breakdown. An analogous result for one-shifted GMRES is shown in [7] in a different manner. A discussion on how to continue the computations with standard GMRES in case the dimension of the available Krylov subspace at breakdown is not large enough to represent the desired solution with sufficient accuracy also is provided in [7]. The approach advocated there can be adapted to m-shifted GMRES. The following result depends on the residual vector r_0. Results for standard GMRES that are independent of the residual vector can be found in [22]; see also [7] for further discussions on iterative methods for linear systems of equations with a singular matrix.

Theorem 2.3. *Let the matrix $A \in \mathbb{R}^{n\times n}$ be of rank $\ell < n$. Apply m-shifted GMRES to the solution of (2.1) and assume that the method breaks down at step $\ell + 1$. Let the matrix $K_\ell^{(m)}$ be defined by (2.4). If $\operatorname{rank}(K_\ell^{(m)}) = \ell$, then m-shifted GMRES yields a least-squares solution of (2.1) at breakdown. If, instead, $\operatorname{rank}(K_\ell^{(m)}) < \ell$, then the least-squares solution of (2.1) lives in $\mathbb{K}_\ell(A, A^m r_0) + \mathcal{R}(A^T)$.*

Proof. It follows from Proposition 2.1 that the relation (2.5) holds. Therefore $\mathcal{R}(K_\ell^{(m)}) \subset \mathcal{R}(A)$. By assumption $\operatorname{rank}(K_\ell^{(m)}) = \operatorname{rank}(A)$. Hence,

$$\mathcal{R}(K_\ell^{(m)}) = \mathcal{R}(A). \tag{2.6}$$

Assume first that $\operatorname{rank}(AK_\ell^{(m)}) = \ell$. Then it follows from (2.5) that $H_{\ell,\ell}$ is nonsingular. A vector $\widetilde{x} \in \mathbb{R}^n$ is a least-squares solution of (2.1) if and only if the associated residual vector is orthogonal to $\mathcal{R}(A)$, i.e., if and only if

$$(K_\ell^{(m)})^T(r_0 - A\widetilde{x}) = 0. \tag{2.7}$$

At breakdown, we solve

$$\min_{y\in\mathbb{R}^\ell} \|AK_\ell^{(m)}y - r_0\|.$$

Using (2.5) the unique solution can be written as

$$\widetilde{y} = H_{\ell,\ell}^{-1}((K_\ell^{(m)})^T K_\ell^{(m)})^{-1}(K_\ell^{(m)})^T r_0. \tag{2.8}$$

It follows from (2.5) that $\widetilde{x} = K_\ell^{(m)}\widetilde{y}$ satisfies (2.7).

Now consider the situation when $\operatorname{rank}(AK_\ell^{(m)}) < \ell$. Similarly as above, we have that the relation (2.6) holds and, therefore, $\widetilde{x} \in \mathbb{R}^n$ is a least-squares solution of (2.1) if and only if (2.7) holds. Since $H_{\ell,\ell}$ is singular, the right-hand side of (2.8) does not exist. We circumvent this difficulty as follows. First observe that the leading principal $(\ell-1)\times(\ell-1)$ submatrix of $H_{\ell,\ell}$ is nonsingular. Therefore, $\operatorname{rank}(H_{\ell,\ell}) = \ell - 1$.

Let $u \in \mathbb{R}^\ell$ be a unit vector that is orthogonal to $\mathcal{R}(H_{\ell,\ell})$ and define $v = A^\dagger K_\ell^{(m)}u \in \mathcal{R}(A^T)$, where $A^\dagger$ denotes the Moore–Penrose pseudoinverse. It follows from (2.6) that

$$Av = K_\ell^{(m)}u. \tag{2.9}$$

We seek a least-squares solution of (2.1) of the form

$$\widetilde{x} = K_\ell^{(m)}\widetilde{y} + v\widetilde{\eta}, \quad \widetilde{y} \in \mathbb{R}^\ell,\ \widetilde{\eta} \in \mathbb{R}. \tag{2.10}$$

Substituting this expression into (2.7), and using (2.5) and (2.9), yields

$$\begin{aligned} 0 &= (K_\ell^{(m)})^T(r_0 - AK_\ell^{(m)}\widetilde{y} - Av\widetilde{\eta}) \\ &= (K_\ell^{(m)})^T(r_0 - K_\ell^{(m)}H_{\ell,\ell}\widetilde{y} - K_\ell^{(m)}u\widetilde{\eta}) \\ &= (K_\ell^{(m)})^T r_0 - (K_\ell^{(m)})^T K_\ell^{(m)}[H_{\ell,\ell}, u]\begin{bmatrix}\widetilde{y}\\ \widetilde{\eta}\end{bmatrix}, \end{aligned}$$

which is equivalent to

$$[H_{\ell,\ell}, u]\begin{bmatrix}\widetilde{y}\\ \widetilde{\eta}\end{bmatrix} = ((K_\ell^{(m)})^T K_\ell^{(m)})^{-1}(K_\ell^{(m)})^T r_0.$$

Since the matrix $[H_{\ell,\ell}, u]$ is of full rank, the above system has a solution $[\widetilde{y}^T, \widetilde{\eta}]^T$. It defines a least-squares solution (2.10) of (2.1) of the desired form. □

We do not form the matrix $K_\ell^{(m)}$ in actual computations, because it is very ill-conditioned for many matrices A; see, e.g., [23] for a discussion on the conditioning of related matrices. Instead, solution methods for (2.1) are based on variants of the standard Arnoldi process; see [5,8] for details.

Proposition 2.4. *Let $x_0 \in \mathbb{R}^n$ denote the initial iterate and determine subsequent iterates, $x_\ell^{(m)}$, by m-shifted GMRES. Define the associated residual vectors $r_0 = b - Ax_0$ and $r_\ell^{(m)} = b - Ax_\ell^{(m)}$, $\ell = 1, 2, 3, \ldots$. Then*

$$x_\ell^{(m)} - x_0 \in \mathbb{K}_\ell(A, A^m r_0) \tag{2.11}$$

and

$$(A^{m+i} r_0, r_\ell^{(m)}) = 0, \quad i = 1, 2, \ldots, \ell, \tag{2.12}$$

where $(u, v) = u^T v$ denotes the standard inner product in $\mathbb{R}^n$.

Proof. By the projection theorem, the minimum residual condition (1.3) is equivalent to the Petrov–Galerkin condition applied to the residual expressed by the orthogonality condition (2.12). □

It is convenient to introduce the matrices

$$\begin{aligned} W_\ell^{(m)} &= K_\ell^{(m+1)}, \\ Z_{\ell+1}^{(m)} &= [r_0, W_\ell^{(m)}], \end{aligned} \tag{2.13}$$

for integers $\ell \geq 1$ and $m \geq -1$. It follows from (2.11) that

$$x_\ell^{(m)} - x_0 = W_\ell^{(m-1)} y_\ell^{(m)} \quad \text{and} \quad r_\ell^{(m)} = r_0 - W_\ell^{(m)} y_\ell^{(m)}$$

for some $y_\ell^{(m)} \in \mathbb{R}^\ell$, which is determined by the orthogonality conditions (2.12). We obtain the system of linear equations

$$(W_\ell^{(m)})^T W_\ell^{(m)} y_\ell^{(m)} = (W_\ell^{(m)})^T r_0.$$

Assume that the matrix $W_\ell^{(m)}$ is of full rank. Then the properties (2.11) and (2.12) allow us to express $x_\ell^{(m)}$ and $r_\ell^{(m)}$ in the forms

$$\begin{aligned} x_\ell^{(m)} &= x_0 + W_\ell^{(m-1)} (W_\ell^{(m)})^\dagger r_0, \\ r_\ell^{(m)} &= P_\ell^{(m)} r_0, \qquad P_\ell^{(m)} = I - W_\ell^{(m)} (W_\ell^{(m)})^\dagger, \end{aligned} \tag{2.14}$$

for $m = 1, 2, 3, \ldots$, where as above $(W_\ell^{(m)})^\dagger$ denotes the Moore–Penrose pseudoinverse of $W_\ell^{(m)}$. We see that $r_\ell^{(m)}$ is an orthogonal projection of r_0 onto the complement of $\mathbb{K}_\ell(A, A^{m+1} r_0)$.

The following lemma summarizes classical results about matrix inversion in block form. We will need them below.

Lemma 2.5. *Let $s \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n \times q}$ with $1 \leq q < n$ be such that the matrix $N = [s, M] \in \mathbb{R}^{n \times (q+1)}$ is of full rank. Let the orthogonal projector P_M be defined by $P_M = I - MM^\dagger$. Then*

$$(N^T N)^{-1} = \begin{bmatrix} \dfrac{1}{\alpha} & u^T \\ u & \Gamma \end{bmatrix}, \tag{2.15}$$

where

$$\begin{aligned} \alpha &= s^T P_M s, \\ u &= -\frac{(M^T M)^{-1} M^T s}{\alpha}, \\ \Gamma &= (M^T M)^{-1} + \frac{(M^T M)^{-1} M^T s s^T M (M^T M)^{-1}}{\alpha}. \end{aligned}$$

Moreover, the orthogonal projector $P_N = I - NN^\dagger$ can be expressed as

$$P_N = P_M - \frac{P_M s s^T P_M}{\alpha}. \tag{2.16}$$

Proof. The results of the lemma can be shown with the aid of the Sherman–Morrison–Woodbury formula described in, e.g., [24]. We omit the details. □

Theorem 2.6. *Assume that the matrices $W_\ell^{(m)}$ are of full rank for $1 \le \ell \le \nu_m$ and $m = -1, 0, \ldots, q$, where ν_m is the grade of the vector $A^{m+1}r_0$. Then the residual vectors $r_\ell^{(m)}$ satisfy*

$$\|r_\ell^{(m)}\|^2 = \frac{1}{e_1^T\,((Z_{\ell+1}^{(m)})^T Z_{\ell+1}^{(m)})^{-1}\,e_1}, \tag{2.17}$$

where $Z_{\ell+1}^{(m)}$ is defined by (2.13).

Let the orthogonal projector $P_\ell^{(m)}$ be given by (2.14) and define the vector $t_\ell^{(m)} = P_\ell^{(m)} A^m r_0$. Then

$$r_{\ell+1}^{(m-1)} - r_\ell^{(m)} = -\frac{(t_\ell^{(m)})^T r_0}{(t_\ell^{(m)})^T t_\ell^{(m)}}\, t_\ell^{(m)} = -\frac{(t_\ell^{(m)})^T r_\ell^{(m)}}{(t_\ell^{(m)})^T t_\ell^{(m)}}\, t_\ell^{(m)} \tag{2.18}$$

and

$$\|r_{\ell+1}^{(m-1)} - r_\ell^{(m)}\| = \frac{|(t_\ell^{(m)})^T r_0|}{\|t_\ell^{(m)}\|},$$
$$\|t_\ell^{(m)}\|^2 = \frac{1}{e_1^T\,((W_{\ell+1}^{(m-1)})^T W_{\ell+1}^{(m-1)})^{-1} e_1}.$$

Moreover,

$$\|r_{\ell+1}^{(m-1)} - r_\ell^{(m)}\| \le \|r_\ell^{(m)}\|. \tag{2.19}$$

Proof. By (2.14), we have $r_\ell^{(m)} = P_\ell^{(m)} r_0$ and $\|r_\ell^{(m)}\|^2 = r_0^T r_0 - r_0^T W_\ell^{(m)} W_\ell^{(m)\dagger} r_0$. To obtain (2.17), it is easy to see that by (2.15), $\|r_\ell^{(m)}\|^{-2}$ is the (1, 1) block of $((Z_{\ell+1}^{(m)})^T Z_{\ell+1}^{(m)})^{-1}$.

The relation (2.18) can be deduced from (2.16). Observing that $W_{\ell+1}^{(m-1)} = [s, M]$, where $s = A^m r_0$ and $M = W_\ell^{(m)}$, allows us to write

$$P_{\ell+1}^{(m-1)} = P_\ell^{(m)} - \frac{P_\ell^{(m)} (A^m r_0)(A^m r_0)^T P_\ell^{(m)}}{(A^m r_0)^T P_\ell^{(m)} (A^m r_0)}.$$

Therefore, since

$$r_i^{(j)} = P_i^{(j)} r_0, \qquad (P_\ell^{(m)})^2 = P_\ell^{(m)},$$

and

$$t_\ell^{(m)} (t_\ell^{(m)})^T r_0 = t_\ell^{(m)} (t_\ell^{(m)})^T r_\ell^{(m)} = ((t_\ell^{(m)})^T r_0) t_\ell^{(m)} = ((t_\ell^{(m)})^T r_\ell^{(m)}) t_\ell^{(m)},$$

we obtain that

$$\begin{aligned} r_{\ell+1}^{(m-1)} - r_\ell^{(m)} &= -\frac{P_\ell^{(m)} (A^m r_0)(A^m r_0)^T P_\ell^{(m)} r_0}{(A^m r_0)^T P_\ell^{(m)} (A^m r_0)} \\ &= -\frac{(t_\ell^{(m)})^T r_0}{(t_\ell^{(m)})^T t_\ell^{(m)}}\, t_\ell^{(m)} = -\frac{(t_\ell^{(m)})^T r_\ell^{(m)}}{(t_\ell^{(m)})^T t_\ell^{(m)}}\, t_\ell^{(m)}. \end{aligned}$$

Moreover, we have

$$\begin{aligned} e_1^T\,((W_{\ell+1}^{(m-1)})^T W_{\ell+1}^{(m-1)})^{-1}\,e_1 &= \frac{1}{s^T P_M s} \\ &= \frac{1}{(A^m r_0)^T P_\ell^{(m)} (A^m r_0)} = \frac{1}{\|t_\ell^{(m)}\|^2}. \end{aligned}$$

Finally, application of the Cauchy–Schwarz inequality yields (2.19). □

The property (2.19) is interesting, because it relates residual errors associated with different m- and ℓ-values. However, there is no simple relation between the errors in the associated iterates. In particular, the error in $x_{\ell+1}^{(m-1)}$ may be much larger than the error in $x_\ell^{(m)}$.

3. Computed examples

We apply shifted and standard GMRES methods to the solution of several standard test problems to illustrate the performance of these methods. All computations are carried out in MATLAB with about 15 significant decimal digits.

Table 3.1
Example 3.1: Iterates determined by the discrepancy principle and best iterates computed by m-shifted GMRES methods for $m = 0, 1, \ldots, 4$. The noise level is $\nu = 1 \cdot 10^{-2}$.

Shifted GMRES	Discrepancy principle		Best iterate	
Shift	Iteration number	Relative error	Iteration number	Relative error
0	4	$1.0 \cdot 10^{-1}$	10	$2.9 \cdot 10^{-2}$
1	4	$2.4 \cdot 10^{-2}$	8	$2.0 \cdot 10^{-2}$
2	5	$2.5 \cdot 10^{-2}$	9	$1.7 \cdot 10^{-2}$
3	6	$2.5 \cdot 10^{-2}$	9	$1.8 \cdot 10^{-2}$
4	6	$2.5 \cdot 10^{-2}$	9	$1.7 \cdot 10^{-2}$

Example 3.1. Let the matrix A be obtained by discretizing the integral equation of the first kind

$$\int_{-6}^{6} \kappa(t,s)x(s)ds = b(t), \quad -6 \le t \le 6,$$

discussed by Phillips [25]. Its solution, kernel, and right-hand side are given by

$$x(s) = \begin{cases} 1 + \cos\left(\frac{\pi}{3}s\right), & \text{if } |s| < 3, \\ 0, & \text{otherwise}, \end{cases}$$
$$\kappa(t,s) = x(t-s),$$
$$b(t) = (6 - |t|)\left(1 + \frac{1}{2}\cos\left(\frac{\pi}{3}t\right)\right) + \frac{9}{2\pi}\sin\left(\frac{\pi}{3}|t|\right).$$

We discretize this integral equation by a Nyström method based on a composite trapezoidal quadrature rule with 1000 equidistant nodes. This gives the nonsymmetric matrix $A \in \mathbb{R}^{1000\times 1000}$. A discretization of the exact solution defines $\widehat{x} \in \mathbb{R}^{1000}$ from which we determine $\widehat{b} = A\widehat{x}$. A vector $e \in \mathbb{R}^{1000}$ with normally distributed random entries with zero mean simulates noise; it is scaled to correspond to a specified noise level

$$\nu = \frac{\|e\|}{\|\widehat{b}\|}.$$

We determine the contaminated right-hand side in (1.1) from (1.2) and scale the error vector e so that $\nu = 1 \cdot 10^{-2}$ or $\nu = 1 \cdot 10^{-3}$.

The initial iterate is in all computed examples in this section chosen to be $x_0 = 0$. Since the m-shifted GMRES methods are minimal residual methods, the norm of the residual error $r_k^{(m)} = b - Ax_k^{(m)}$ is a decreasing function of the iteration number k. We assume that the noise level ν is known, and then can determine a suitable iterate $x_k^{(m)}$ with the aid of the discrepancy principle, which prescribes that $k \ge 0$ be the smallest iteration number such that $\|r_k^{(m)}\| \le \nu\|\widehat{b}\|$. Table 3.1 reports the iteration numbers k determined by the discrepancy principle as well as the relative errors $\|x_k^{(m)} - \widehat{x}\|/\|\widehat{x}\|$ in the associated computed approximate solution $x_k^{(m)}$ of (1.1) for the shifts $m = 0, 1, \ldots, 4$. The shift $m = 0$ corresponds to the standard GMRES method. Table 3.1 also shows iteration numbers and relative errors for the iterates that best approximate $\widehat{x}$ for $m = 0, 1, \ldots, 4$. These iterates are difficult to compute when $\widehat{x}$ is not known, but inclusion of results for these iterates in the table sheds light on how much the error in the iterates determined by the discrepancy principle could be reduced if we knew how to compute the best iterates.

Table 3.1 shows the m-shifted GMRES methods to yield more accurate approximations of the desired solution $\widehat{x}$ for $m \ge 1$ than standard GMRES both when the iterates are determined with the aid of the discrepancy principle and when the best achievable iterates are known. Reducing the noise level to $\nu = 1 \cdot 10^{-3}$ yields Table 3.2. Also for this lower noise level, the m-shifted GMRES methods with $m \ge 1$ provide more accurate approximations of $\widehat{x}$ than standard GMRES. More precisely, the best iterate with the smallest relative error is achieved for $m = 2$ when the noise level is $\nu = 1 \cdot 10^{-2}$ (Table 3.1), and for $m = 1$ when the noise level is reduced to $\nu = 1 \cdot 10^{-3}$ (Table 3.2). In our experience, it generally holds that when the noise level is large, it is beneficial to choose a larger shift m than when the noise level is small. When the right-hand side b in (1.1) is error-free, the most accurate approximation of $\widehat{x}$ is typically obtained for $m = 0$. □

Example 3.2. The matrix of this example is obtained by discretizing the integral equation of the first kind

$$\int_{-\pi/2}^{\pi/2} \kappa(\tau,\sigma)x(\sigma)d\sigma = b(\tau), \quad -\frac{\pi}{2} \le \tau \le \frac{\pi}{2},$$

where

$$\kappa(\sigma,\tau) = (\cos(\sigma) + \cos(\tau))\left(\frac{\sin(\xi)}{\xi}\right)^2, \quad \xi = \pi(\sin(\sigma) + \sin(\tau)).$$

Table 3.2
Example 3.1: Iterates determined by the discrepancy principle and best iterates for m-shifted GMRES methods for $m = 0, 1, \ldots, 4$. The noise level is $\nu = 1 \cdot 10^{-3}$.

Shifted GMRES	Discrepancy principle		Best iterate	
Shift	Iteration number	Relative error	Iteration number	Relative error
0	9	$2.8 \cdot 10^{-2}$	8	$1.0 \cdot 10^{-2}$
1	10	$6.4 \cdot 10^{-3}$	10	$6.4 \cdot 10^{-3}$
2	10	$8.0 \cdot 10^{-3}$	12	$6.4 \cdot 10^{-3}$
3	10	$8.3 \cdot 10^{-3}$	11	$7.9 \cdot 10^{-3}$
4	10	$8.6 \cdot 10^{-3}$	12	$6.4 \cdot 10^{-3}$

Table 3.3
Example 3.2: Iterates determined by the discrepancy principle and best iterates for m-shifted GMRES methods for $m = 0, 1, \ldots, 4$. The noise level is $\nu = 1 \cdot 10^{-2}$.

Shifted GMRES	Discrepancy principle		Best iterate	
Shift	Iteration number	Relative error	Iteration number	Relative error
0	7	$2.2 \cdot 10^{-1}$	6	$1.3 \cdot 10^{-1}$
1	6	$2.1 \cdot 10^{-1}$	5	$1.3 \cdot 10^{-1}$
2	6	$1.4 \cdot 10^{-1}$	7	$6.5 \cdot 10^{-2}$
3	6	$8.1 \cdot 10^{-2}$	7	$6.5 \cdot 10^{-2}$
4	6	$1.3 \cdot 10^{-1}$	7	$6.5 \cdot 10^{-2}$

Table 3.4
Example 3.2: Iterates determined by the discrepancy principle and best iterates for m-shifted GMRES methods for $m = 0, 1, \ldots, 4$. The noise level is $\nu = 1 \cdot 10^{-3}$.

Shifted GMRES	Discrepancy principle		Best iterate	
Shift	Iteration number	Relative error	Iteration number	Relative error
0	7	$5.6 \cdot 10^{-2}$	7	$5.6 \cdot 10^{-2}$
1	6	$5.6 \cdot 10^{-2}$	7	$5.6 \cdot 10^{-2}$
2	7	$5.2 \cdot 10^{-2}$	7	$5.2 \cdot 10^{-2}$
3	7	$5.2 \cdot 10^{-2}$	7	$5.2 \cdot 10^{-2}$
4	7	$5.2 \cdot 10^{-2}$	7	$5.2 \cdot 10^{-2}$

The right-hand side function $b(\tau)$ is chosen so that the solution $x(\sigma)$ is the sum of two Gaussian functions. This integral equation is discussed by Shaw [26]. We discretize it by a Nyström method based on the composite trapezoidal rule with $n = 1000$ equidistant nodes. This yields the nonsymmetric matrix $A \in \mathbb{R}^{1000\times 1000}$ and the discretized solution $\widehat{x} \in \mathbb{R}^{1000}$ from which we determine $\widehat{b} = A\widehat{x}$. The contaminated right-hand sides $b \in \mathbb{R}^{1000}$ are defined analogously as in Example 3.1. The noise levels are $\nu = 1 \cdot 10^{-2}$ and $\nu = 1 \cdot 10^{-3}$.

Table 3.3 is analogous to Table 3.1. The noise level is $\nu = 1 \cdot 10^{-2}$. Table 3.3 illustrates that it may be beneficial to let $m \geq 2$ for some linear discrete ill-posed problems. In fact, the best iterate and the iterate determined by the discrepancy principle have minimal relative errors for the shift $m = 3$. Table 3.4 is analogous to Table 3.3 with the noise level reduced to $\nu = 1 \cdot 10^{-3}$. Table 3.4 shows the best iterate as well as the iterate determined by the discrepancy principle to have minimal residual errors for $m = 2$. Thus, it is beneficial to let m grow with the noise level. □

Example 3.3. The Fredholm integral equation of the first kind

$$\int_0^{\pi} \kappa(\sigma, \tau)x(\tau)d\tau = b(\sigma), \quad 0 \leq \sigma \leq \frac{\pi}{2}, \tag{3.1}$$

with kernel $\kappa(\sigma, \tau) = \exp(\sigma \cos(\tau))$, right-hand side function $b(\sigma) = 2\sinh(\sigma)/\sigma$, and solution $x(\tau) = \sin(\tau)$ is discussed by Baart [27]. We use the MATLAB code baart from [28] to discretize (3.1) by a Galerkin method with 1000 orthonormal box functions as test and trial functions. The code produces the nonsymmetric matrix $A \in \mathbb{R}^{1000\times 1000}$ and the scaled discrete approximation $\widehat{x} \in \mathbb{R}^{1000}$ of $x(\tau)$. The noise-free right-hand side is given by $\widehat{b} = A\widehat{x}$. The entries of the noise vector $e \in \mathbb{R}^{1000}$ are generated in the same way as in Example 3.1 with noise level $\nu = 1 \cdot 10^{-2}$. The contaminated right-hand side is defined by (1.2). Table 3.5 reports the computed results. The table shows m-shifted GMRES with $m \geq 1$ to give more accurate approximations of $\widehat{x}$ than standard GMRES. □

4. Conclusion

This paper is concerned with properties of m-shifted GMRES methods. Our analysis sheds light on some aspects of these methods. Computed examples illustrate the benefit in improved quality of the computed approximate solutions that can be

Table 3.5
Example 3.3: Iterates determined by the discrepancy principle and best iterates for m-shifted GMRES methods for $m = 0, 1, \ldots, 4$. The noise level is $\nu = 1 \cdot 10^{-2}$.

Shifted GMRES	Discrepancy principle		Best iterate	
Shift	Iteration number	Relative error	Iteration number	Relative error
0	3	$3.1 \cdot 10^{-1}$	3	$3.1 \cdot 10^{-1}$
1	3	$3.5 \cdot 10^{-2}$	3	$3.5 \cdot 10^{-2}$
2	3	$3.5 \cdot 10^{-2}$	3	$3.5 \cdot 10^{-2}$
3	3	$3.5 \cdot 10^{-2}$	3	$3.5 \cdot 10^{-2}$
4	3	$3.5 \cdot 10^{-2}$	3	$3.5 \cdot 10^{-2}$

achieved by using a positive shift m. Extensive computational experience suggests that the choices $m = 2$ and $m = 3$ often are suitable. However, the value of m that yields iterates that best approximate the desired solution $\widehat{x}$ depends both on the noise level and on the problem to be solved.

Acknowledgment

The second author's research was supported in part by NSF grant DMS-1115385.

References

[1] Y. Saad, Iterative Methods for Sparse Linear Systems, second ed., SIAM, Philadelphia, 2003.
[2] D. Calvetti, B. Lewis, L. Reichel, Restoration of images with spatially variant blur by the GMRES method, in: F.T. Luk (Ed.), Advanced Signal Processing Algorithms, Architectures, and Implementations X, in: Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 4116, The International Society for Optical Engineering, Bellingham, WA, 2000, pp. 364–374.
[3] D. Calvetti, B. Lewis, L. Reichel, On the choice of subspace for iterative methods for linear discrete ill-posed problems, Int. J. Appl. Math. Comput. Sci. 11 (2001) 1069–1092.
[4] D. Calvetti, B. Lewis, L. Reichel, GMRES, L-curves, and discrete ill-posed problems, BIT 42 (2002) 44–65.
[5] L. Dykes, L. Reichel, A family of range restricted iterative methods for linear discrete ill-posed problems, Dolomit. Res. Notes Approx. 6 (2013) 27–36.
[6] P.C. Hansen, T.K. Jensen, Noise propagation in regularizing iterations for image deblurring, Electron. Trans. Numer. Anal. 31 (2008) 204–220.
[7] L. Reichel, Q. Ye, Breakdown-free GMRES for singular systems, SIAM J. Matrix Anal. Appl. 26 (2005) 1001–1021.
[8] A. Neuman, L. Reichel, H. Sadok, Implementations of range restricted iterative methods for linear discrete ill-posed problems, Linear Algebra Appl. 436 (2012) 3974–3990.
[9] A. Neuman, L. Reichel, H. Sadok, Algorithms for range restricted iterative methods for linear discrete ill-posed problems, Numer. Algorithms 59 (2012) 325–331.
[10] H. Sadok, Analysis of the convergence of the minimal and orthogonal residual methods, Numer. Algorithms 40 (2005) 201–216.
[11] D. Calvetti, B. Lewis, L. Reichel, On the regularizing properties of the GMRES method, Numer. Math. 91 (2002) 605–625.
[12] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.
[13] T.A. Hearn, L. Reichel, Application of denoising methods to regularization of ill-posed problems, Numer. Algorithms 66 (2014) 761–777.
[14] C. Brezinski, G. Rodriguez, S. Seatzu, Error estimates for linear systems with applications to regularization, Numer. Algorithms 49 (2008) 85–104.
[15] S. Kindermann, Convergence analysis of minimization-based noise levelfree parameter choice rules for linear ill-posed problems, Electron. Trans. Numer. Anal. 38 (2011) 233–257.
[16] L. Reichel, G. Rodriguez, Old and new parameter choice rules for discrete ill-posed problems, Numer. Algorithms 63 (2013) 65–87.
[17] M. Matinfar, H. Zareamoghaddam, M. Eslami, M. Saeidy, GMRES implementations and residual smoothing techniques for solving ill-posed linear systems, Comput. Math. Appl. 63 (2012) 1–13.
[18] D. Calvetti, S. Morigi, L. Reichel, F. Sgallari, Tikhonov regularization and the L-curve for large, discrete ill-posed problems, J. Comput. Appl. Math. 123 (2000) 423–446.
[19] S. Gazzola, P. Novati, M.R. Russo, On Krylov projection methods and Tikhonov regularization, Electron. Trans. Numer. Anal. 44 (2015) 83–123.
[20] S. Gazzola, P. Novati, Automatic parameter setting for Arnoldi–Tikhonov methods, J. Comput. Appl. Math. 256 (2014) 180–195.
[21] M.E. Hochstenbach, N. McNinch, L. Reichel, Discrete ill-posed least-squares problems with a solution norm constraint, Linear Algebra Appl. 436 (2012) 3801–3818.
[22] P.N. Brown, H.F. Walker, GMRES on (nearly) singular systems, SIAM J. Matrix Anal. Appl. 18 (1997) 37–51.
[23] B. Philippe, L. Reichel, On the generation of Krylov subspace bases, Appl. Numer. Math. 62 (2012) 1171–1186.
[24] G.H. Golub, C.F. Van Loan, Matrix Computations, fourth ed., Johns Hopkins University Press, Baltimore, 2013.
[25] D.L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, J. ACM 9 (1962) 84–97.
[26] C.B. Shaw Jr., Improvements of the resolution of an instrument by numerical solution of an integral equation, J. Math. Anal. Appl. 37 (1972) 83–112.
[27] M.L. Baart, The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least-squares problems, IMA J. Numer. Anal. 2 (1982) 241–247.
[28] P.C. Hansen, Regularization tools version 4.0 for Matlab 7.3, Numer. Algorithms 46 (2007) 189–194.