

On choice of preconditioner for minimum residual methods for nonsymmetric matrices

Jennifer Pestana

Andrew J. Wathen

Existing convergence bounds for Krylov subspace methods such as GM-RES for nonsymmetric linear systems give little mathematical guidance for the choice of preconditioner. Here, we establish a desirable mathematical property of a preconditioner which guarantees that convergence of a minimum residual method will essentially depend only on the eigenvalues of the preconditioned system, as is true in the symmetric case. Our theory covers only a subset of nonsymmetric coefficient matrices but computations indicate that it might be more generally applicable.

Oxford University Mathematical Institute
Numerical Analysis Group
24-29 St Giles'
Oxford, England OX1 3LB
E-mail: pestana@maths.ox.ac.uk

August, 2010

1 Introduction

For linear systems of equations with a real symmetric (or complex hermitian) coefficient matrix, there are standard iterative methods of choice: for positive definite problems the Conjugate Gradient method [23] and for indefinite problems the MINRES method [36]. (The SYMMLQ method [36] is a lesser used alternative in the indefinite case.) For these methods there exist descriptive convergence theories based on eigenvalues, i.e. convergence bounds given in terms of polynomial approximation problems on the eigenvalue spectrum which are often reasonably tight [20, page 51]. A key consequence is that it is clear what one is trying to achieve to get fast convergence via preconditioning in these cases, namely well distributed eigenvalues or reduced spectral condition number.

By contrast for nonsymmetric matrices, there are many possible and competing methods. GMRES [39] is probably the most popular minimum residual method, but unless the required number of iterations is kept small, the increased work of orthogonalisation in the Arnoldi process and the storage requirements can both grow to infeasible levels. To avoid this increased work and storage, several methods based on the nonsymmetric Lanczos method are also widely employed [16, 17, 44, 45, 51], though none minimises the residual and we do not consider them here; the convergence of these methods is even less well understood than minimum residual methods such as GMRES. For any and all of these methods, preconditioning is almost always employed in the hope of obtaining acceptably rapid convergence.

Given the importance of iterative solution methods for large and sparse or structured linear or linearised systems of equations, it remains a fundamental and largely open mathematical question as to what one is trying to achieve with preconditioning for nonsymmetric matrices: the glib answer ‘fast convergence’ gives no idea nor criteria for making other than heuristically motivated choices. The known results which inform the user are largely negative: Faber and Manteuffel [15] famously proved that, except for trivial cases with few distinct eigenvalues, the set of matrices for which there exist short-term recurrence methods which minimize some relevant quantity is rather small—precisely those diagonalisable matrices which have all of their eigenvalues on a straight line in the complex plane¹ (see [14] for a more recent proof), Nachtigal, Reddy and Trefethen [35] considered three of the competing nonsymmetric iterative methods available in 1992 and identified classes of real $n \times n$ matrices for which each of the methods could outperform the other two methods by a factor of at least $\sqrt{n}$ and Greenbaum, Pták and Strakoš [21] proved that given any complete set of eigenvalues and any monotonically reducing convergence curve that it is possible to construct a matrix with the given eigenvalues and an initial residual (starting iterate) such that GMRES will achieve precisely the given convergence curve². Thus one can certainly achieve stagnation until the n^{th}

¹With a different recurrence than that examined by Faber and Manteuffel it is possible to obtain a minimum residual method with short-term recurrences for matrices with concyclic eigenvalues [25, 26], although the set of matrices possessing this property is also small. For a summary of short-term recurrence methods, see [31].

²The Greenbaum, Pták and Strakoš result built on work in [22] and was followed by a characterisation in [1] of the set of matrices, with prescribed eigenvalues, and right-hand sides for which a given convergence curve is achieved.

iterate even for a matrix all of whose eigenvalues are 1 (see [22] for an example)! This is in stark contrast to the situation for real symmetric matrices.

As is well known ([20, §2.4],[38, §6.5],[12, §4.1.1]) for the linear system $Bx = b$, a minimum residual method such as GMRES computes iterate vectors $\{x_k\}, k = 1, 2, \dots$ from a starting guess x_0 such that $x_k - x_0$ lie in the Krylov subspaces

$$\mathcal{K}_k(B, r_0) = \text{span}\{r_0, Br_0, B^2r_0, \dots, B^{k-1}r_0\}, \quad (1.1)$$

so that the residuals $r_k = b - Bx_k$ satisfy $r_k = p(B)r_0$ where $p \in \Pi_k$, the set of real polynomials of degree k or less, and p satisfies the consistency condition $p(0) = 1$. Assuming that $B \in \mathbb{R}^{n \times n}$ is diagonalisable,

$$B = Z\Lambda Z^{-1},$$

where Λ is a diagonal matrix of eigenvalues and Z has corresponding eigenvectors as its columns, the minimum residual property ensures that

$$\frac{\|r_k\|_2}{\|r_0\|_2} \leq \|Z\|_2 \|Z^{-1}\|_2 \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(B)} |p(\lambda)| \quad (1.2)$$

where $\sigma(B)$ is the eigenvalue spectrum of B . It is clear how this allows the Greenbaum, Pták and Strakoš construction in the nonsymmetric case since even for a diagonalisable matrix B with n distinct eigenvalues arbitrarily close to 1, the smallness of the polynomial term

$$\min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(B)} |p(\lambda)| \quad (1.3)$$

in (1.2) for $k \geq 1$ can still allow for arbitrarily slow convergence for a large enough value of

$$\kappa_2(Z) = \|Z\|_2 \|Z^{-1}\|_2,$$

the condition number of the eigenvector matrix. This is clearly not true in the symmetric case when Z is necessarily orthogonal and thus $\|Z\|_2 = \|Z^{-1}\|_2 = 1$, so that it is essentially the eigenvalues which determine convergence in that case.

In this paper we provide some limited theory which aims at providing some sufficient mathematical conditions for selection of preconditioners in the nonsymmetric case. Specifically for an important subset of nonsymmetric matrices we identify a mathematical condition on a preconditioner which will guarantee that the convergence of a minimum residual method (such as GMRES) will essentially be determined by the eigenvalues alone, without the possible huge term coming from the eigenvectors in (1.2). We comment that other bounds for GMRES convergence are available, notably based on the field of values [10, 11] and pseudospectra [49, Chapter 26] and improvement of these bounds might also follow from the analysis here.

Let us illustrate the point of this paper by reference to a well known example, the 5005×5005 Sherman3 matrix together with its provided right hand side vector from the Harwell-Boeing test set [9] which is now available through Matrix Market (<http://math.nist.gov/MatrixMarket>). This long standing example is certainly well within the

capabilities of excellent sparse direct methods at this point [7, 8], but it is a useful (and often quoted) example in the context of iterative methods [5, 6, 27, 29, 52]. We employ in this example the well known ILU(0) preconditioning technique described by Meijerink and van der Vorst [33]. Figure 1 shows the commonly seen behaviour with such examples: without any preconditioning, the solid line shows the reduction in $\|r_k\|_2$ for increasing numbers of GMRES iterations k . For this nonsymmetric example, the eigenvalues of B are in fact all real and so it is possible to accurately approximate/bound the term (1.3) by employing Chebyshev polynomials as is done in the case of symmetric matrices (see for example [12, §§2.1.1, 6.2.4]). Using such an approach, the convergence bound (1.2) in this case is given by the dashed line which is seen to be nearly horizontal. This is seen not to be such a bad upper bound for the early iterations when little progress is made by GMRES in reducing the Euclidean norm of the residual. The dot-dashed line shows the reduction in the preconditioned residuals (here plotted also in the Euclidean norm) achieved when ILU(0) preconditioning is used in conjunction with GMRES. This rapid convergence is however far from what is described by the bound (1.2) which is indicated by the dotted line on the figure and is computed similarly to the above. The eigenvectors of the original matrix are in fact quite well conditioned for this example (hence the intercept for the dashed line is near 1), however the poor eigenvalue distribution leads to slow convergence without preconditioning. The preconditioner improves the eigenvalue distribution but in fact makes the eigenvector condition number much worse as seen in the large intercept of the dotted line (and in the numbers in Table 1 where the relevant data for this problem are given). The dotted line is nowhere descriptive of the convergence as displayed in the dot-dashed line.

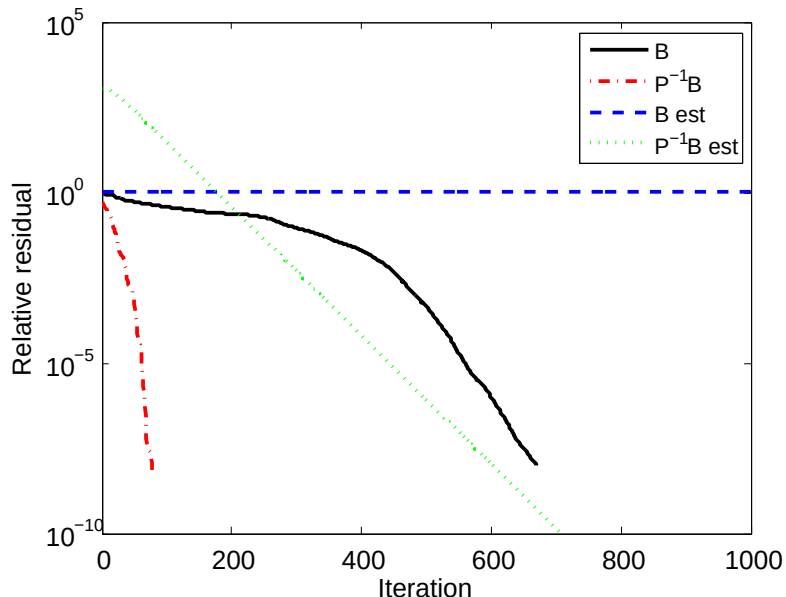


Figure 1: Convergence curves and convergence bounds for ILU(0) preconditioned GMRES applied to Sherman3.

Name	$\kappa_2(B)$	$\kappa_2(P^{-1}B)$	$\kappa_2(Z_{P^{-1}B})$	$\kappa_2(Z_B)$
sherman3	1.2×10^{19}	2.3×10^3	1.2×10^3	1.002

Table 1: Spectral norm condition numbers ($\kappa_2(\cdot)$) of the Sherman3 matrix B and the ILU(0) preconditioned system $P^{-1}B$ and of the matrix of eigenvectors of each matrix (denoted by Z).

The main results of this paper show why the eigenvector condition number term in (1.2) can essentially be ignored for this example. We indicate situations where this might more generally be done.

In §2 we introduce the class of matrices for which our analysis applies and indicate how the application of minimum residual methods in non-standard inner products could lead to convergence dependent only on eigenvalues. This leads in §3 to the consideration of and relationship between minimum residual methods in different inner products and norms. Our main results regarding preconditioning are given in §4 and the results of many numerical tests are described in §5.

Throughout this manuscript the Parlett convention is adopted, excepting the absolutely standard use of H_k and V_k in the Arnoldi process: any matrix identified by a character with reflectional symmetry will be symmetric, otherwise nonsymmetric. Thus, $A, H, M, \dots$ are symmetric whereas $B, C, D, \dots$ are nonsymmetric.

2 Self-adjointness in nonstandard inner products

Let $H \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Then

$$\langle x, y \rangle_H = x^T H y, \quad x, y \in \mathbb{R}^n, \quad (2.1)$$

is a symmetric bilinear form on $\mathbb{R}^n$. The properties of these forms, at least in the case that H is invertible, are discussed in [18]. If, moreover, H is positive definite, (2.1) defines an inner product, with associated H -norm

$$\|x\|_H = \sqrt{\langle x, x \rangle_H}$$

and induced matrix norm

$$\|B\|_H = \max_{x \neq 0} \frac{\|Bx\|_H}{\|x\|_H}.$$

For example, $\langle \cdot, \cdot \rangle_I$ is the standard Euclidean inner product with induced spectral norm $\|\cdot\|_I$. It is straightforward to show that for any factorisation $H = S^T S$ of a positive definite matrix H ,

$$\|B\|_H = \|SBS^{-1}\|_I. \quad (2.2)$$

An H -orthogonal matrix Z is one for which $Z^T H Z = I$. A matrix B is self-adjoint with respect to $\langle \cdot, \cdot \rangle_H$, or shortly H -self-adjoint, if

$$\langle Bx, y \rangle_H = \langle x, By \rangle_H$$

for all $x, y \in \mathbb{R}^n$. Since H -self-adjointness implies that

$$x^T B^T H y = x^T H B y$$

for all x and y , B is H -self-adjoint if and only if

$$HB = B^T H, \quad (2.3)$$

that is, if HB is symmetric. B is additionally H -positive definite if

$$\langle Bx, x \rangle_H > 0$$

for all nonzero x or, equivalently, if HB is positive definite with respect to the Euclidean inner product. In the case that H is positive definite, and so induces an inner product, B is H -self-adjoint if and only if it is diagonalisable and its eigenvalues are real [47, 48, 50]. Simple algebra shows that B is positive definite with respect to this inner product if and only if these eigenvalues are also positive. See [2, 18, 31] and the references therein for a detailed discussion of H -self-adjointness and H -definiteness.

Taussky [47] proved that every real matrix B can be decomposed as

$$B = H^{-1}T, \quad (2.4)$$

with H and T symmetric. This leads to the, perhaps surprising, result that every real matrix is self-adjoint with respect to some symmetric bilinear form. For example, it is not difficult to show that the Jordan block $J \in \mathbb{R}^{n \times n}$,

$$J = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix},$$

is self-adjoint with respect to indefinite symmetric bilinear forms induced by Hankel matrices of the form

$$H = \begin{bmatrix} & & & a_1 \\ & & \ddots & a_2 \\ & \ddots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_n \end{bmatrix},$$

where $a_1, \dots, a_n$ are any real scalars.

Since (2.4) defines self-adjointness of B with respect to $\langle \cdot, \cdot \rangle_H$, it is possible to characterise the symmetric bilinear forms with respect to which a given matrix B is self-adjoint by employing a theorem of Uhlig's [50]. We quote Theorem 1 from [50] using our notation:

Theorem 1 *Let B be a real matrix with real Jordan normal form J , $Z^{-1}BZ = J = \text{diag}(J_i)$ where Z is real.*

(a) If H^{-1} is a real symmetric factor of B , i.e., $B = H^{-1}T$ for T symmetric, then $H = Z^{-T}C^{-T} \text{diag}(\epsilon_i Y_i) C^{-1} Z^{-1}$ for some $\epsilon_i = \pm 1$, C with $CJ = JC$ and $\dim(Y_i) = \dim(J_i)$ for all i where the Y_i are square matrices of the form

$$\begin{bmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{bmatrix}.$$

(b) If $H = Z^{-T}C^{-T} \text{diag}(\epsilon_i Y_i) C^{-1} Z^{-1}$ for arbitrary $\epsilon_i = \pm 1$ and C with $CJ = JC$ and $\dim(Y_i) = \dim(J_i)$ then H^{-1} is a real symmetric factor of B .

Thus B is self-adjoint with respect to symmetric bilinear forms $\langle \cdot, \cdot \rangle_H$, where $H = Z^{-T}C^{-T} \text{diag}(\epsilon_i Y_i) C^{-1} Z^{-1}$. As previously mentioned, if B has real eigenvalues and diagonalisation $B = Z\Lambda Z^{-1}$ it is self-adjoint with respect to a positive definite inner product. Stoll and Wathen [46] showed that this inner product may take the simpler form $\langle \cdot, \cdot \rangle_{Z^{-T}\Theta Z^{-1}}$, where Θ is diagonal and positive definite. Indeed, these constitute the only inner products with respect to which B is self-adjoint when the eigenvalues of B are distinct, since then C in Theorem 1 commutes with Λ if and only if it is diagonal (see, for example, [24, page 24]). When the eigenvalues are not distinct there is some more freedom in the definition of C . Certainly, however, B is still self-adjoint in the inner product induced by $H = Z^{-T}\Theta Z^{-1}$.

If $H = Z^{-T}\Theta Z^{-1}$ then $H = \tilde{Z}^{-T}\tilde{Z}^{-1}$, where $\tilde{Z} = Z\Theta^{-\frac{1}{2}}$. Additionally, if $B = Z\Lambda Z^{-1}$ is a diagonalisation of B then so is $B = \tilde{Z}\tilde{\Lambda}\tilde{Z}^{-1}$. Thus, Θ can be viewed as a scaling matrix and without loss of generality we consider only inner products of the form $H = \tilde{Z}^{-T}\tilde{Z}^{-1}$ in the following. We drop the tildes for convenience.

Significantly, if B is self-adjoint with respect to a non-standard inner product $\langle \cdot, \cdot \rangle_H$, there exist minimum residual methods in this inner product, which we call H -MR methods, for which convergence depends chiefly on the eigenvalues of B . These methods have been thoroughly explored in, for example, [2, 3, 28] and the references therein, where the authors discuss the particular methods Orthodir, Orthores and Orthomin. The H -MINRES algorithm, a generalisation of the Paige and Saunders method in [36], is discussed in [46]. (For situations where H -Conjugate Gradient methods are applicable see [30].) This property of H -MR methods is a generalisation of the well known result for Euclidean norm minimum residual methods, or I -MR methods, applied to symmetric matrices and is illustrated by the following convergence bound which previously appeared in [4].

Theorem 2 *Let $B = Z\Lambda Z^{-1}$ be a diagonalisation of B and let $H = Z^{-T}Z^{-1}$. Then r_k^H , the k^{th} residual vector of H -MR, satisfies*

$$\frac{\|r_k^H\|_H}{\|r_0\|_H} \leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(B)} |p(\lambda)|. \quad (2.5)$$

Proof By definition, the H -MR residual vector is given by $r_k^H = p_k(B)r_0$, where $p_k(z)$ is the unique polynomial of degree less than or equal to k with $p(0) = 1$ that minimises $\|r_k^H\|_H$ and r_0 is the initial residual. Using (2.2), it follows that

$$\begin{aligned} \|r_k^H\|_H &= \|p_k(B)r_0\|_H \\ &\leq \|Zp_k(\Lambda)Z^{-1}\|_H \|r_0\|_H \\ &= \|Z^{-1}Zp_k(\Lambda)Z^{-1}Z\|_2 \|r_0\|_H \\ &= \|p_k(\Lambda)\|_2 \|r_0\|_H \\ &\leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(B)} |p(\lambda)| \|r_0\|_H \end{aligned}$$

as required. $\blacksquare$

As pointed out by Simoncini [41], since

$$Z^{-1}r_k = p(\Lambda)Z^{-1}r_0,$$

convergence depends essentially on eigenvalues if measured in $\|\cdot\|_H$, where $H = Z^{-T}Z^{-1}$. Thus, although B is not symmetric, its self-adjointness with respect to $\langle \cdot, \cdot \rangle_H$ ensures that the convergence of H -MR, measured in the H -norm, depends mainly on the eigenvalues of B . Of course, the right-hand side vector may affect convergence significantly, as is true even of symmetric positive definite matrices (see [34, page 492] and the references therein). However, we expect that the behaviour of H -MR methods applied to H -self-adjoint matrices will typically be determined by eigenvalues.

This is in contrast to eigenvalue-based convergence bounds for minimum residual methods in any other inner product. If $M = F^T F$ and M -MR is employed then by the same argument used in Theorem 2 we obtain the bound

$$\|r_k^M\|_M = \|p_k(B)r_0\|_M \leq \|FZp_k(\Lambda)Z^{-1}F^{-1}\|_2 \|r_0\|_M$$

and so

$$\frac{\|r_k^M\|_M}{\|r_0\|_M} \leq \kappa_2(FZ) \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(B)} |p(\lambda)|. \quad (2.6)$$

In the case of I -MR methods (for which $F = M = I$) this leads to the standard bound (1.2) that includes the condition number of the eigenvectors. If the eigenvector condition number is large, (1.2) may be a poor estimate of convergence.

3 Relationship between the convergence of minimum residual methods in different inner products

In the previous section we showed that nonsymmetric H -self-adjoint matrices enjoy H -MR convergence that essentially depends on eigenvalues. A natural consideration is whether this knowledge gives insight into the convergence of minimum residual methods in the Euclidean, or indeed any other, inner product. Progress in this direction has been made by Essai [13], who related the convergence of GMRES (and FOM) in $\langle \cdot, \cdot \rangle_D$ to that

of GMRES (and FOM) in $\langle \cdot, \cdot \rangle_I$, where D was a diagonal matrix related to the initial residual. The key to Essai's result is the QR decomposition, that we employ below, and that has also been exploited by, for example, Sadok and Szyld [40] in their comparison of CMRH and GMRES and by Simoncini and Szyld [42, Theorem 3.6] in the more general situation that one wishes to relate M -MR convergence, in any inner product $\langle \cdot, \cdot \rangle_M$, to I -MR convergence. Approaches that do not explicitly exploit a QR relationship can also be used to relate M -MR and I -MR convergence as shown in [11, Theorem 4.7]. Similar results for conjugate gradient methods can be found in [2].

Here we extend Essai's results in order to address the broader question of how the convergence of two minimum residual methods M -MR and W -MR, in the respective inner products $\langle \cdot, \cdot \rangle_M$ and $\langle \cdot, \cdot \rangle_W$, compare. We restrict our attention to MR methods for which the basis vectors are orthogonal with respect to the associated inner product. This certainly still encompasses the best known implementations of minimum residual methods, namely MINRES and GMRES. To simplify the exposition the initial iterate is taken to be the zero vector in the following; this loses no generality.

Since M -MR and W -MR are both Krylov subspace methods, at the k^{th} step each selects an iterate from the same k -dimensional Krylov subspace (1.1). However, M -MR minimises the M -norm of the k^{th} residual over the Krylov subspace while W -MR minimises the W -norm and so the iterates of each method will generally differ.

Let the columns of V_k^M be the M -orthogonal basis vectors of the k^{th} Krylov subspace generated by M -MR. Then the standard equation (see, for example, [19, page 548])

$$BV_k^M = V_{k+1}^M H_k^M \quad (3.1)$$

holds, where H_k^M is a $(k+1) \times k$ upper Hessenberg matrix. Similarly, if V_k^W has as its columns the W -orthogonal basis vectors of W -MR then

$$BV_k^W = V_{k+1}^W H_k^W. \quad (3.2)$$

Since the k^{th} iterate x_k^W of W -MR is selected from the k^{th} Krylov subspace (1.1), it can be written as a linear combination of W -orthogonal basis vectors. That is, $x_k^W = V_k^W y_k^W$ where $y_k^W \in \mathbb{R}^k$ is chosen so that $\|r_k^W\|_W$ is minimised. By combining this with (3.2) the usual relationship for the residual can be derived:

$$\begin{aligned} r_k^W &= b - Bx_k^W \\ &= b - BV_k^W y_k^W \\ &= b - V_{k+1}^W H_k^W y_k^W \\ &= V_{k+1}^W (\|r_0\|_W e_1 - H_k^W y_k^W). \end{aligned}$$

Since V_{k+1}^W is W -orthogonal,

$$\|r_k^W\|_W = \min_{y \in \mathbb{R}^k} \|\|r_0\|_W e_1 - H_k^W y\|_2. \quad (3.3)$$

Analogously, the M -norm of the M -MR residual satisfies

$$\|r_k^M\|_M = \min_{z \in \mathbb{R}^k} \|\|r_0\|_M e_1 - H_k^M z\|_2. \quad (3.4)$$

The crucial observation is that M -MR and W -MR form bases of the same nested Krylov subspaces. Consequently, assuming that these methods do not terminate at or before the k^{th} step,

$$V_k^W = V_k^M R_k, \quad (3.5)$$

where $R_k \in \mathbb{R}^{k \times k}$ is upper triangular and nonsingular; this is a QR decomposition of V_k^W in the $\langle \cdot, \cdot \rangle_M$ inner product. From (3.1), (3.2) and (3.5) we obtain the following (see [13, Proposition 1] for details):

$$\begin{aligned} R_k &= (V_k^M)^T M V_k^W, \\ R_k^{-1} &= (V_k^W)^T W V_k^M, \\ H_k^W &= R_{k+1}^{-1} H_k^M R_k. \end{aligned} \quad (3.6)$$

Equations (3.5) and (3.6) provide the links necessary to relate M -MR and W -MR convergence. As a first step we show that it is possible to express the W -MR residual in terms of quantities associated with M -MR.

Lemma 3 *Let r_k^W be the k^{th} residual of W -MR. Then*

$$\|r_k^W\|_W = \min_{\bar{y} \in \mathbb{R}^k} \|\|r_0\|_M e_1 - H_k^M \bar{y}\|_{R_{k+1}^{-T} R_{k+1}^{-1}}, \quad (3.7)$$

where H_k^M is the upper Hessenberg matrix in (3.1) and R_{k+1} is the upper triangular matrix in (3.5).

Proof From (3.3) we have

$$\begin{aligned} \|r_k^W\|_W &= \min_{y \in \mathbb{R}^k} \|\|r_0\|_W e_1 - H_k^W y\|_2 \\ &= \min_{\bar{y} \in \mathbb{R}^k} \|\|r_0\|_W e_1 - H_k^W R_k^{-1} \bar{y}\|_2, \end{aligned} \quad (3.8)$$

where $\bar{y} = R_k y$. Now,

$$r_0 = \|r_0\|_W V_{k+1}^W e_1 = \|r_0\|_M V_{k+1}^M e_1$$

and so, recalling (3.6), we find that

$$\|r_0\|_W e_1 = \|r_0\|_M (V_{k+1}^W)^T W V_{k+1}^M e_1 = \|r_0\|_M R_{k+1}^{-1} e_1.$$

Also from (3.6), $H_k^M = R_{k+1} H_k^W R_k^{-1}$. Combining these results with (3.8) gives

$$\begin{aligned} \|r_k^W\|_W &= \min_{\bar{y} \in \mathbb{R}^k} \|R_{k+1}^{-1} (\|r_0\|_M e_1 - H_k^M \bar{y})\|_2 \\ &= \min_{\bar{y} \in \mathbb{R}^k} \|\|r_0\|_M e_1 - H_k^M \bar{y}\|_{R_{k+1}^{-T} R_{k+1}^{-1}}. \end{aligned}$$

as required. $\blacksquare$

Comparison with the M -MR residual (3.4) shows that the quantities minimised by W -MR and M -MR are the same but the norms with respect to which they are minimised differ. This allows us to relate M -MR and W -MR convergence in the central theorem of this section.

Theorem 4 *The k^{th} W -MR residual r_k^W and the k^{th} M -MR residual r_k^M are related by*

$$\sqrt{\lambda_{\min}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M \leq \|r_k^W\|_W \leq \sqrt{\lambda_{\max}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M, \quad (3.9)$$

where R_{k+1} is defined in (3.5).

Proof Let

$$z_k = \arg \min_z \|\|r_0\|_M e_1 - H_k^M z\|_2,$$

so that

$$\|r_k^M\|_M = \|\|r_0\|_M e_1 - H_k^M z_k\|_2.$$

Then, from (3.7),

$$\begin{aligned} \|r_k^W\|_W &= \min_{\bar{y} \in \mathbb{R}^k} \|\|r_0\|_M e_1 - H_k^M \bar{y}\|_{R_{k+1}^{-T}R_{k+1}^{-1}} \\ &\leq \|\|r_0\|_M e_1 - H_k^M z_k\|_{R_{k+1}^{-T}R_{k+1}^{-1}} \\ &\leq \sqrt{\lambda_{\max}(R_{k+1}^{-T}R_{k+1}^{-1})}\|\|r_0\|_M e_1 - H_k^M z_k\|_2, \end{aligned}$$

where the last line follows from application of the Courant-Fischer theorem (see, for example, [19, Theorem 8.1.2]) to $R_{k+1}^{-T}R_{k+1}^{-1}$. We see from (3.4) that this is equivalent to

$$\|r_k^W\|_W \leq \sqrt{\lambda_{\max}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M.$$

A similar process shows that

$$\sqrt{\lambda_{\min}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M \leq \|r_k^W\|_W,$$

which completes the proof. $\blacksquare$

Information concerning the extreme eigenvalues of $R_{k+1}^{-T}R_{k+1}^{-1}$ would be instructive. However, we are unaware of a means of obtaining this other than by performing both W -MR and M -MR, which would be prohibitively expensive in most cases. Consequently, we bound these extreme eigenvalues by those of $M^{-1}W$ but note that the resulting bounds may be pessimistic.

Theorem 5 *The residuals of W -MR and M -MR satisfy*

$$\sqrt{\lambda_{\min}(M^{-1}W)}\|r_k^M\|_M \leq \|r_k^W\|_W \leq \sqrt{\lambda_{\max}(M^{-1}W)}\|r_k^M\|_M, \quad (3.10)$$

where $\lambda_{\min}(M^{-1}W)$ and $\lambda_{\max}(M^{-1}W)$ are the minimum and maximum eigenvalues of $M^{-1}W$, respectively

Proof We first note that in exact arithmetic both M -MR and W -MR terminate at or before the n^{th} step. In the event of termination at step $k < n$ we can append M -orthogonal basis vectors to V_k^M so that its columns form an M -orthogonal basis of $\mathbb{R}^n$. The matrix V_k^W can be similarly expanded by W -orthogonal basis vectors in such a way that R_n is invertible and upper triangular (by, for example, applying a Gram-Schmidt procedure in $\langle \cdot, \cdot \rangle_W$ to the last $n - k$ columns of V_n^M).

Since $R_{k+1}^{-T} R_{k+1}^{-1}$ is symmetric, by Cauchy's interlacing theorem [19, Theorem 8.1.7],

$$\begin{aligned}\lambda_{\max}(R_{k+1}^{-T} R_{k+1}^{-1}) &\leq \lambda_{\max}(R_n^{-T} R_n^{-1}), \\ \lambda_{\min}(R_n^{-T} R_n^{-1}) &\leq \lambda_{\min}(R_{k+1}^{-T} R_{k+1}^{-1}), \quad k < n.\end{aligned}$$

Additionally, since V_n^M and V_n^W are invertible, (3.6) implies that

$$R_n = (V_n^M)^T M V_n^W \text{ and } R_n^{-1} = (V_n^W)^T W V_n^M.$$

Accordingly,

$$R_n^{-T} R_n^{-1} = (V_n^M)^{-1} M^{-1} W V_n^M,$$

from which we conclude that $R_n^{-T} R_n^{-1}$ is similar to $M^{-1}W$.

We see, therefore, that the maximum eigenvalue of $R_{k+1}^{-T} R_{k+1}^{-1}$ is bounded from above by the maximum eigenvalue of $M^{-1}W$. Similarly, the minimum eigenvalue of $R_{k+1}^{-T} R_{k+1}^{-1}$ is bounded from below by the minimum eigenvalue of $M^{-1}W$.

Thus, (3.9) above can be reformulated as

$$\sqrt{\lambda_{\min}(M^{-1}W)} \|r_k^M\|_M \leq \|r_k^W\|_W \leq \sqrt{\lambda_{\max}(M^{-1}W)} \|r_k^M\|_M$$

as required. $\blacksquare$

It is possible, however, that the extreme eigenvalues of $R_{k+1}^{-T} R_{k+1}^{-1}$ may not be close to those of $M^{-1}W$ when k is small. If this is the case, the initial convergence of M -MR and W -MR will be closer than (3.10) suggests. That is, the less accessible bound in Theorem 4 will be more descriptive than that in Theorem 5.

Subsequent results make use of the upper bound in (3.10) and so we briefly examine the nearness of (3.9) and (3.10) for a particular class of matrix. We construct $B \in \mathbb{R}^{n \times n}$ so that it has diagonalisation $B = Z \Lambda Z^{-1}$, where the elements of Z and the diagonal elements of Λ are from the standard normal distribution. Thus B is H -self-adjoint, where $H = Z^{-T} Z^{-1}$. The right-hand side vector also has normally distributed elements. Both I -GMRES and H -GMRES are applied to this system. The residuals of H -GMRES, measured in the H -norm, and the residuals of I -GMRES, measured in the I -norm, given in Figure 2 (a).

We then set $M = H$ and $W = I$ in the above and compute explicitly the bounds in (3.9) and (3.10). The results, which are typical for this class of matrix, are shown in Figure 2 (b). From this we see that the two upper bounds are close, although the two lower bounds differ by a order of magnitude until about the 70th iteration. Thus, for this class of matrix not much is lost by using the more easily computed bound based on the eigenvalues of $M^{-1}W$. However, we do not know whether this is true of self-adjoint matrices generally.

4 Preconditioning

We have seen that when minimum residual methods in $\langle \cdot, \cdot \rangle_H$ are applied to an H -self-adjoint matrix convergence primarily depends on its eigenvalues (except for certain right-hand side vectors for which convergence may be better). Unless these eigenvalues

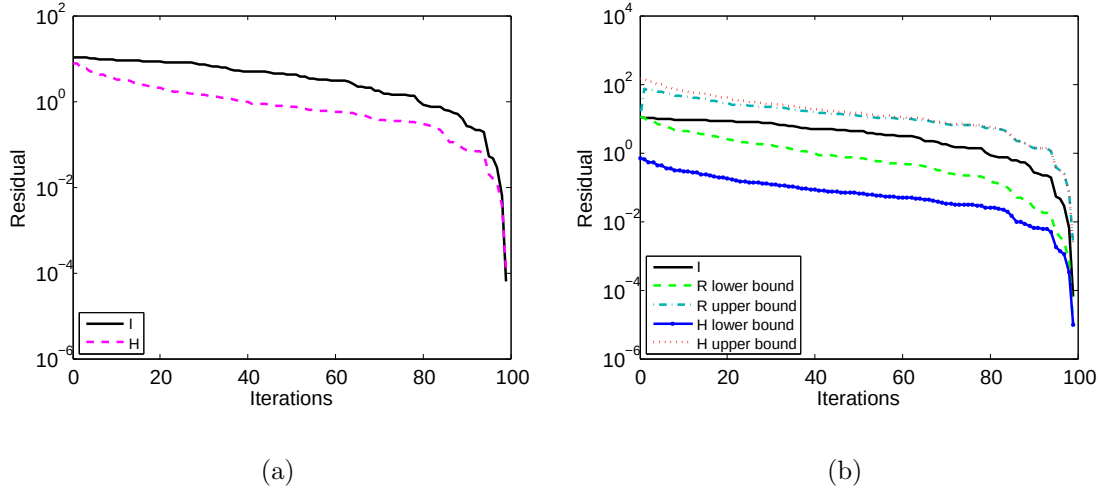


Figure 2: (a) H -GMRES and I -GMRES convergence for the matrix $B = Z\Lambda Z^{-1} \in \mathbb{R}^{100 \times 100}$ described above and (b) I -GMRES convergence and the upper and lower bounds based on the eigenvalues of $R_{k+1}^{-T} R_{k+1}^{-1}$ and of H .

are favourably distributed, convergence will be poor and preconditioning will be required. Here we investigate how the results of the previous sections may assist in the construction of effective left preconditioners, P . It seems not unlikely that similar results hold for right preconditioning, although we have not investigated this here.

If the preconditioned system $P^{-1}B$ is itself self-adjoint with respect to an inner product, the analysis of §2 implies that a minimum residual method in this inner product can be applied. Its convergence depends on the spectrum of $P^{-1}B$ and will, therefore, be rapid if the eigenvalues are well distributed. Inner products with respect to which a preconditioned matrix is self-adjoint have been characterised for certain matrices; see [43, §13] and the references therein. In the case that B is H -self-adjoint but not symmetric, we show that self-adjointness can be preserved by applying a preconditioner that is also H -self-adjoint and is additionally H -positive definite.

Suppose B is H -self-adjoint and that P is H -self-adjoint and H -positive definite. Then, since

$$(HP)(P^{-1}B) = HB = B^T H = B^T P^{-T} P^T H = (P^{-1}B)^T (HP), \quad (4.1)$$

HP induces an inner product with respect to which $P^{-1}B$ is self-adjoint (cf. (2.3)). Application of Theorem 2 shows that convergence of HP-MR applied to this system depends largely on eigenvalues. Specifically, the preconditioned residuals, $P^{-1}r_k^{HP}$, satisfy

$$\frac{\|P^{-1}r_k^{HP}\|_{HP}}{\|P^{-1}r_0\|_{HP}} \leq \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|.$$

Thus, in the situation that B and P are H -self-adjoint and P is H -positive definite, there exist minimum residual methods in $\langle \cdot, \cdot \rangle_{HP}$ for which convergence is essentially

determined by the eigenvalues of the preconditioned system. Additionally, it is possible to find a diagonalisation $P^{-1}B = S\Lambda S^{-1}$ such that $HP = S^{-T}S^{-1}$. Unfortunately, for many applications it is infeasible to compute a H -self-adjoint preconditioner. If, however, we can obtain a preconditioner that is self-adjoint and positive definite with respect to an inner product that is close to H , say $\mathcal{H}$, we might expect that we could find a diagonalisation $P^{-1}B = (\mathcal{S} - E)\Lambda(\mathcal{S} - E)^{-1}$ such that $\mathcal{H}P = \mathcal{S}^{-T}\mathcal{S}^{-1}$, with E being a small perturbation. If this is the case then convergence of $\mathcal{H}P$ -MR is still governed chiefly by eigenvalues, as the following theorem illustrates.

Theorem 6 *Let B be self-adjoint with respect to $\langle \cdot, \cdot \rangle_H$ and P be self-adjoint and positive definite with respect to $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. Let*

$$P^{-1}B = (\mathcal{S} - E)\Lambda(\mathcal{S} - E)^{-1} \quad (4.2)$$

be a diagonalisation of $P^{-1}B$, where

$$\mathcal{H}P = \mathcal{S}^{-T}\mathcal{S}^{-1}. \quad (4.3)$$

If $\|\mathcal{S}^{-1}E\|_2 < 1$ the preconditioned residual $P^{-1}r_k^{\mathcal{H}P}$ of any minimum residual method in $\langle \cdot, \cdot \rangle_{\mathcal{H}P}$ satisfies

$$\frac{\|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P}}{\|P^{-1}r_0\|_{\mathcal{H}P}} \leq \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|. \quad (4.4)$$

Proof We start from the polynomial representation of the residual, $P^{-1}r_k^{\mathcal{H}P} = p_k(P^{-1}B)P^{-1}r_0$. Taking $\mathcal{H}P$ -norms gives

$$\|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P} = \|p_k(P^{-1}B)P^{-1}r_0\|_{\mathcal{H}P}.$$

Using (4.2), we obtain

$$\begin{aligned} \|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P} &\leq \|(\mathcal{S} - E)p_k(\Lambda)(\mathcal{S} - E)^{-1}\|_{\mathcal{H}P} \|P^{-1}r_0\|_{\mathcal{H}P} \\ &= \|\mathcal{S}^{-1}(\mathcal{S} - E)p_k(\Lambda)(\mathcal{S} - E)^{-1}\mathcal{S}\|_2 \|P^{-1}r_0\|_{\mathcal{H}P}, \end{aligned}$$

where the last line follows from (2.2). Thus,

$$\|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P} \leq \|I - \mathcal{S}^{-1}E\|_2 \|(I - \mathcal{S}^{-1}E)^{-1}\|_2 \|p_k(\Lambda)\|_2 \|P^{-1}r_0\|_{\mathcal{H}P}.$$

If $\|\mathcal{S}^{-1}E\|_2 < 1$ then

$$\|(I - \mathcal{S}^{-1}E)^{-1}\|_2 \leq \frac{1}{1 - \|\mathcal{S}^{-1}E\|_2}$$

(see, for example, [24, page 301]) and

$$\frac{\|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P}}{\|P^{-1}r_0\|_{\mathcal{H}P}} \leq \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|.$$

■

When $\|\mathcal{S}^{-1}E\|_2$ is small, the factor $(1 + \|\mathcal{S}^{-1}E\|_2)/(1 - \|\mathcal{S}^{-1}E\|_2)$ is close to 1 and the convergence of any $\mathcal{H}P$ -MR method is dominated by the eigenvalues of the preconditioned matrix. This is clearly illustrated by the following corollary, which shows that convergence can be bounded by a quantity that depends only on eigenvalues to first order in $\|\mathcal{S}^{-1}E\|_2$.

Corollary 6A *Let $\|\mathcal{S}^{-1}E\|_2 < 1$, where $\mathcal{S}$ and E are as defined in equations (4.2) and (4.3). Then*

$$\frac{\|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P}}{\|P^{-1}r_0\|_{\mathcal{H}P}} \leq (1 + O(\|\mathcal{S}^{-1}E\|_2)) \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|. \quad (4.5)$$

Proof The result follows from the expansion of $(1 - \|\mathcal{S}^{-1}E\|_2)^{-1}$ as a power series. $\blacksquare$

Remark 7 This bound sheds some light on how to precondition a H -self-adjoint matrix. For a preconditioner P that is self-adjoint with respect to an inner product that is close to $\langle \cdot, \cdot \rangle_H$, Corollary 6A indicates that $\mathcal{H}P$ -MR convergence will essentially depend on the eigenvalues of $P^{-1}B$. If, further, these eigenvalues have a more favourable distribution than those of the original system, convergence should be faster. This is analogous to the well understood convergence of symmetric matrices.

Applying a $\mathcal{H}P$ -MR method involves inner products with $\mathcal{H}P$, which may be costly to compute even if this matrix is known. Fortunately, the analysis of §3 allows us to relate the convergence of any $\mathcal{H}P$ -MR method to that of any I -MR method. We emphasise as in §3 that in our experience the bound (4.5) may be more descriptive of what is actually seen even when preconditioned residuals are computed with an I -MR method, such as I -GMRES, and are measured in the Euclidean norm; conversion of the bound (4.4) to the I -norm always introduces a condition number that often makes the bound in the theorem below pessimistic in practice. For a discussion of this topic see [42].

Theorem 8 *Let P , B , $\mathcal{H}$, $\mathcal{S}$ and E be as defined in Theorem 6. Then the residuals of any I -MR method are bounded by*

$$\frac{\|P^{-1}r_k^I\|_2}{\|P^{-1}r_0\|_2} \leq \sqrt{\kappa_2(\mathcal{H}P)} \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|. \quad (4.6)$$

Proof From Theorem 6 we know that the preconditioned residuals of $\mathcal{H}P$ -MR are bounded by

$$\frac{\|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P}}{\|P^{-1}r_0\|_{\mathcal{H}P}} \leq \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|.$$

Then, using (3.10) with $M = \mathcal{H}P$ and $W = I$ we obtain

$$\begin{aligned} \|P^{-1}r_k^I\|_2 &\leq \sqrt{\lambda_{\max}((\mathcal{H}P)^{-1})} \|P^{-1}r_k^{\mathcal{H}P}\|_{\mathcal{H}P} \\ &\leq \frac{1}{\sqrt{\lambda_{\min}(\mathcal{H}P)}} \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)| \|P^{-1}r_0\|_{\mathcal{H}P} \\ &\leq \frac{\sqrt{\lambda_{\max}(\mathcal{H}P)}}{\sqrt{\lambda_{\min}(\mathcal{H}P)}} \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)| \|P^{-1}r_0\|_2, \end{aligned}$$

where the last line follows from the Courant-Fischer theorem applied to $\mathcal{H}P$. Thus,

$$\frac{\|P^{-1}r_k^I\|_2}{\|P^{-1}r_0\|_2} \leq \sqrt{\kappa_2(\mathcal{H}P)} \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p \in \Pi_k, p(0)=1} \max_{\lambda \in \sigma(P^{-1}B)} |p(\lambda)|$$

as required. $\blacksquare$

The factor $\sqrt{\lambda_{\max}((\mathcal{H}P)^{-1})}$ is an upper bound on $\sqrt{\lambda_{\max}(R_{k+1}^{-T}R_{k+1}^{-1})}$, where R_{k+1} is defined in (3.5). At the end of §3 it was found that for the particular class of matrix examined there was little difference between the two. However, as previously discussed we do not know whether this is true of general matrices. Indeed, our experience over a wide range of problems is that if $\mathcal{H}P$ -MR convergence is rapid then I -MR performance is also good. In particular, if the eigenvalues of $P^{-1}B$ are well distributed and H is close to $\mathcal{H}$ a significant transient phase, i.e., stagnation, is not observed.

5 Numerical experiments

In this section we apply the results of the previous section to several examples. Throughout we use I -GMRES, that is, the standard (unrestarted) GMRES method of Saad and Schultz [39]. In Examples 1 and 2 the unpreconditioned relative residuals are measured. For the remaining problems, we employ Matlab's inbuilt GMRES routine that measures relative preconditioned residuals in a recursive fashion. Additionally, in the first two examples, elements of matrices and vectors come from either the standard normal distribution or the uniform distribution on $(0,1)$. For brevity, these are referred to as normally distributed and randomly distributed elements hereafter.

Example 1 Let $B = Z\Lambda Z^{-1} \in \mathbb{R}^{100 \times 100}$, where the elements of Z and of the right-hand size vector are normally distributed and the diagonal entries of Λ are uniformly distributed. It follows that B is self-adjoint (and positive definite) with respect to the inner product induced by $H = Z^{-T}Z^{-1}$. We consider three preconditioners, P_{exact} , $P_{perturbed}$ and P_{randn} , each of which has the same matrix of eigenvectors $\mathcal{Z} = Z + \epsilon N$, where $\epsilon \in [10^{-3}, 1]$ is a scaling factor and N has normally distributed elements. The eigenvalues of P_{exact} are identical to those of B , so that $P = \mathcal{Z}\Lambda\mathcal{Z}^{-1}$. $P_{perturbed}$ has diagonalisation $P_{perturbed} = \mathcal{Z}\tilde{\Lambda}\mathcal{Z}^{-1}$, with $\tilde{\Lambda} = \Lambda + 10^{-1}\text{diag}(\nu)$, where $\nu \in \mathbb{R}^n$ is a vector with uniformly distributed elements. The final preconditioner has diagonalisation $P_{rand} = \mathcal{Z}\Gamma\mathcal{Z}^{-1}$, where Γ is a diagonal matrix with uniformly distributed diagonal entries. By construction, each preconditioner has positive real eigenvalues and is therefore self-adjoint and positive definite with respect to the inner product induced by $\mathcal{H} = \mathcal{Z}^{-T}\mathcal{Z}^{-1}$.

Here, the parameter ϵ determines the nearness of H and $\mathcal{H}$. Theorem 8 implies that as these matrices become more alike the convergence of the preconditioned system should depend more heavily on its eigenvalues; if these have a better distribution than the eigenvalues of B , convergence should improve. This is corroborated by the results in Table 2, which show the number of GMRES iterations required for convergence, averaged over 100 random matrices. When $\epsilon = 1$ the perturbation of Z is of the order of the entries themselves. In this case H and $\mathcal{H}$ cannot be expected to be close and it is not surprising that convergence is poor even when the eigenvalues of the preconditioner are identical to those of B . However, as ϵ is decreased and H and $\mathcal{H}$ get nearer, the influence of eigenvalues on convergence is striking. The number of iterations decreases

with ϵ when P_{exact} and $P_{perturbed}$, each of which improves the eigenvalues of B , are applied. In contrast, P_{rand} takes no account of the eigenvalues of B and consequently does not accelerate convergence. Indeed, as might be expected, it impedes it.

ϵ	B	$P_{exact}^{-1}B$	$P_{perturbed}^{-1}B$	$P_{rand}^{-1}B$
1	58	100	100	100
10^{-1}	58	40	40	100
10^{-2}	58	14	20	100
10^{-3}	58	8	19	100

Table 2: GMRES convergence for the unpreconditioned matrix B and preconditioned systems discussed in Example 1, averaged over 100 matrices.

This dependence on eigenvalues when H and $\mathcal{H}$ are close is also clear from Figure 3, which demonstrates behaviour typical of this type of example. Together with the relative norms of the unpreconditioned residuals for $\epsilon = 10^{-2}$ we plot the convergence of (unpreconditioned) I -GMRES applied to T , where T is constructed so that it has the same eigenvalues as $P_{perturbed}^{-1}B$ but is symmetric. Consequently, we expect I -GMRES convergence for T to be governed chiefly by eigenvalues; in fact in exact arithmetic GMRES is really an expensive implementation of MINRES when applied to this symmetric matrix. Thus, by comparing the convergence of I -GMRES for the preconditioned systems to that of T we are able to assess the extent to which eigenvalues dictate the convergence of I -GMRES for $P_{perturbed}^{-1}B$ and, to a lesser extent, that of I -GMRES for $P_{exact}^{-1}B$. Figure 3 shows that the I -GMRES convergence for $P_{perturbed}^{-1}B$ and $P_{exact}^{-1}B$ are each very similar to that of T . This indicates that in this case, with $\mathcal{H}$ near to H , convergence is dominated by eigenvalues.

The convergence of the preconditioned systems is better than that predicted by Theorem 8. We see that, as discussed at the end of §3, the factor $\sqrt{\kappa_2(\mathcal{H}P)}$ is pessimistic. In particular the preconditioned systems $P_{exact}^{-1}B$ and $P_{perturbed}^{-1}B$ do not suffer from any significant initial stagnation.

Example 2 Although the theory in §4 does not apply when B and P are self-adjoint with respect to symmetric bilinear forms that do not satisfy the positivity axiom for an inner product, our experience is that similar behaviour can be observed in practice. To demonstrate this we repeat the above experiment but let B and its preconditioners have complex eigenvalues. Specifically, the elements of $B \in \mathbb{R}^{100 \times 100}$ are normally distributed so that, provided B has at least one complex conjugate pair of eigenvalues, there is no inner product with respect to which B is self-adjoint. The size of B makes this situation highly likely. Let the real Jordan form of B be $B = ZJZ^{-1}$. Then, by Theorem 1, B is self-adjoint with respect to the symmetric bilinear form induced by $H = Z^{-T} \text{diag}(Y_i)Z^{-1}$, where the Y_i are the same size as the corresponding diagonal blocks of J . We again choose a right-hand side vector with normally distributed elements.

As in the first example, three preconditioners are applied: P_{exact} , $P_{perturbed}$ and P_{rand} . The first, $P_{exact} = \mathcal{Z}J\mathcal{Z}^{-1}$, has the same real Jordan matrix as B but with $\mathcal{Z} = Z + \epsilon N$,

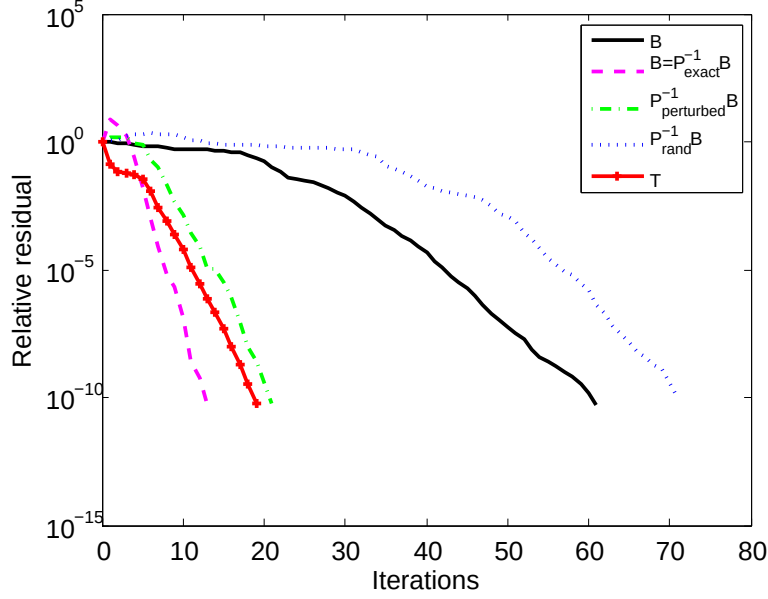


Figure 3: GMRES convergence for the preconditioned systems discussed in Example 1 with $\epsilon = 10^{-2}$ and the symmetric matrix T that has the same eigenvalues as $P_{perturbed}^{-1}B$.

where ϵ and N are defined in Example 1. The second, $P_{perturbed} = \mathcal{Z}\tilde{\mathcal{J}}\mathcal{Z}^{-1}$, is constructed so that its real Jordan matrix is slightly perturbed from that of B . That is, $\tilde{\mathcal{J}} = J + 0.1\Delta$, where Δ has the same block structure as J and its nonzero elements are normally distributed with 2×2 blocks satisfying

$$\Delta_i = \begin{bmatrix} \delta_{real} & \delta_{imag} \\ -\delta_{imag} & \delta_{real} \end{bmatrix}, \quad (5.1)$$

with $\delta_r, \delta_i \in \mathbb{R}$. The real Jordan form of the final preconditioner is $P_{randn} = \mathcal{Z}\Gamma\mathcal{Z}^{-1}$, where Γ has the same block structure as J . Its 2×2 blocks are of the form (5.1) and its nonzero entries are normally distributed. Each preconditioner is self-adjoint with respect to the symmetric bilinear form induced by $\mathcal{H} = \mathcal{Z}^{-T} \text{diag}(Y_i)\mathcal{Z}^{-1}$.

We see from Table 3 and Figure 4 that the results of this experiment show similar trends to those of Example 1. This indicates that even when the symmetric bilinear forms with respect to which B and P are self-adjoint are indefinite, their nearness may determine whether eigenvalues are important. Again, when ϵ is large, so that $\mathcal{Z}$ is far from Z and $\mathcal{H}$ is far from H , it appears that factors other than eigenvalues affect convergence. However, as ϵ decreases and H and $\mathcal{H}$ become more alike, eigenvalues appear to play a larger role; when the preconditioner takes eigenvalues into account, as P_{exact} and $P_{perturbed}$ do, convergence is fast. Convergence is not improved, however, when the eigenvalues of the preconditioned system are no better than those of B , as is the case when P_{randn} is applied.

The effect of eigenvalues is also clear from Figure 4, for which $\epsilon = 10^{-2}$ and T , as in Example 1, has the same eigenvalues as $P_{perturbed}^{-1}B$ but is symmetric. Again, relative

ϵ	B	$P_{exact}^{-1}B$	$P_{perturbed}^{-1}B$	$P_{randn}^{-1}B$
1	100	100	100	100
10^{-1}	100	100	100	100
10^{-2}	100	54	54	100
10^{-3}	100	11	11	100

Table 3: GMRES convergence for the unpreconditioned matrix B and preconditioned systems discussed in Example 2, averaged over 100 matrices.

unpreconditioned residuals are measured. The convergence of I -GMRES for $P_{exact}^{-1}B$ and of $P_{perturbed}^{-1}B$ is similar to that of I -GMRES (which, again, is in essence MINRES) for the symmetric matrix T . Since convergence of I -GMRES applied to T is governed primarily by eigenvalues, this implies that when P_{exact} and $P_{perturbed}$ are applied as preconditioners, the eigenvalues of the preconditioned system essentially determine convergence.

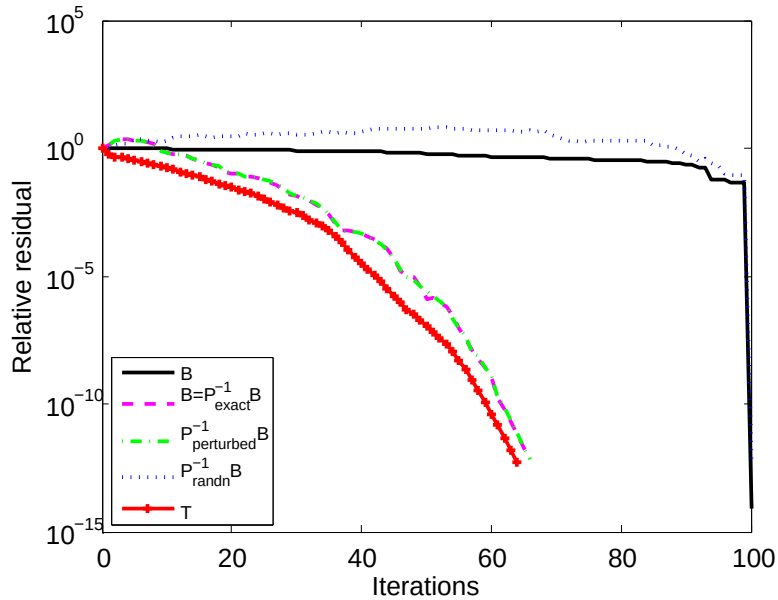


Figure 4: GMRES convergence for the unpreconditioned matrix B , preconditioned systems discussed in Example 2 with $\epsilon = 10^{-2}$ and a matrix with the same eigenvalues as $P_{perturbed}^{-1}B$.

Example 3 Our next examples involve SUPG finite element (FE) discretisations of an advection-diffusion operator [12, Chapters 3 and 4]. Consider the coefficient matrix B produced by a piecewise quadratic (Q_2) FE discretisation of an advection-diffusion equation and the matrix P formed from a piecewise linear (Q_1) discretisation with the same nodes. The matrices P and B are the same size but may appear otherwise dissimilar. However, since both are formed from discretisations of the same operator, we might expect that the inner products with respect to which P and B are self-adjoint

are close. If this is the case, preconditioning B with P should result in convergence that depends chiefly on the eigenvalues of the preconditioned system.

We begin first with a simple one-dimensional example in order to illustrate phenomena that may be present in more complicated problems:

$$-\epsilon u_{xx} + u_x = 0, \quad u(0) = 1, \quad u(1) = 0. \quad (5.2)$$

With $\epsilon = 0.1$ and 50 quadratic elements (so that $P, B \in \mathbb{R}^{101 \times 101}$), the mesh Peclet number is 0.1 and, using the parameter choice as in [12, Equation (3.44)], the SUPG method reduces to a Galerkin FE method. For this small mesh Peclet number the eigenvalues of both P and B are real and positive. Thus, P and B are self-adjoint and positive definite with respect to the inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\langle \cdot, \cdot \rangle_H$, respectively. Although it would be advantageous to determine the matrix E in (4.2), this is difficult. However, it is possible to gauge the size of E indirectly. Recall from (4.2) that there exists a factorisation $\mathcal{H}P = \mathcal{S}^{-T}\mathcal{S}^{-1}$ such that $(\mathcal{S} - E)^{-1}P^{-1}B(\mathcal{S} - E) = \Lambda$. If E were small we would expect, therefore, that the off-diagonal elements of $G = \mathcal{S}P^{-1}B\mathcal{S}$ would also be small. Thus, as an indirect indicator of the size of E we measure $\|G - \text{diag}(\text{diag}(G))\|_2$, the spectral norm of the off-diagonal elements. For this preconditioned system we have $\|G - \text{diag}(\text{diag}(G))\|_2 = 0.1931$, which is small relative to $n = 100$. One would infer, therefore, that convergence of I -GMRES applied to the preconditioned system should depend mainly on the (well distributed) eigenvalues of the preconditioned system. This appears to be the case. When I -GMRES is applied (using the inbuilt GMRES implementation in Matlab) to both the original and preconditioned systems, we see from Figure 5 that convergence of I -GMRES for the preconditioned system is rapid and that no stagnation is observed.

Preconditioning with the exact Q_1 matrix P may not be much easier than working with the Q_2 matrix B , but the above results are indicative of what happens when practical approximations are used. For example, when the exact matrix P in the above problem is replaced by a geometric multigrid process, rapid convergence is maintained: instead of computing P exactly we employ the multigrid process for piecewise linear elements in [37] with Gauss-Seidel smoothing. Four pre- and post-smoothing steps are used and 2 multigrid cycles are performed each time the preconditioner is required. The convergence of I -GMRES applied to a system with B and this multigrid preconditioner is shown in Figure 6. We see that rapid convergence of I -GMRES is achieved for the preconditioned system. Unfortunately, since the preconditioner is not explicitly available when multigrid is applied, we cannot comment more on the theory in §4 in this case.

Example 4 Next, we apply an SUPG finite element discretisation to (5.2) but increase the advection by setting $\epsilon = 0.001$. As a result the mesh Peclet number increases to 10, so that streamline upwinding is required. Additionally, B has imaginary eigenvalues and is therefore not self-adjoint with respect to an inner product. Thus, our theory does not apply. However, similarly to our observations in Example 2, we see from Figure 7 that I -GMRES for B preconditioned with the exactly computed P shows rapid convergence without stagnation. The absence of stagnation suggests that it is the (fairly

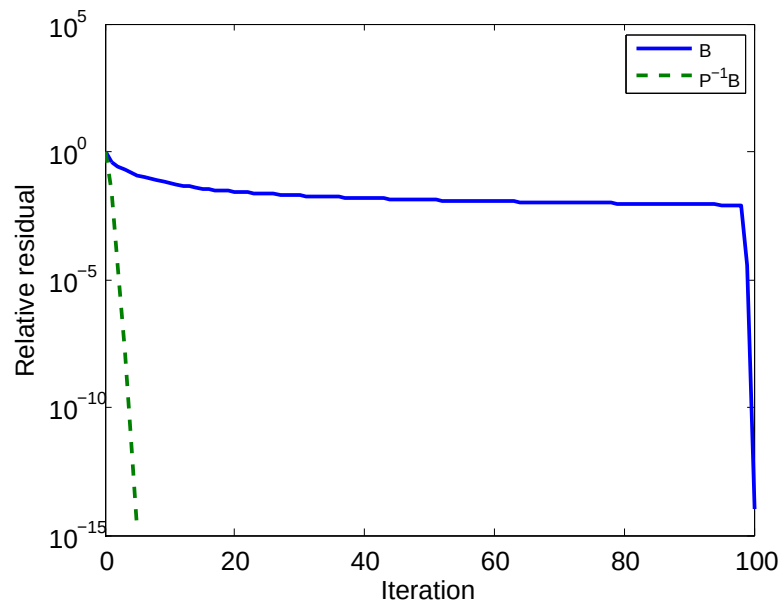


Figure 5: 1D example: Relative residuals for the quadratic finite element matrix and relative preconditioned residuals for the preconditioned system with $n = 101$ and $\epsilon = 0.1$.

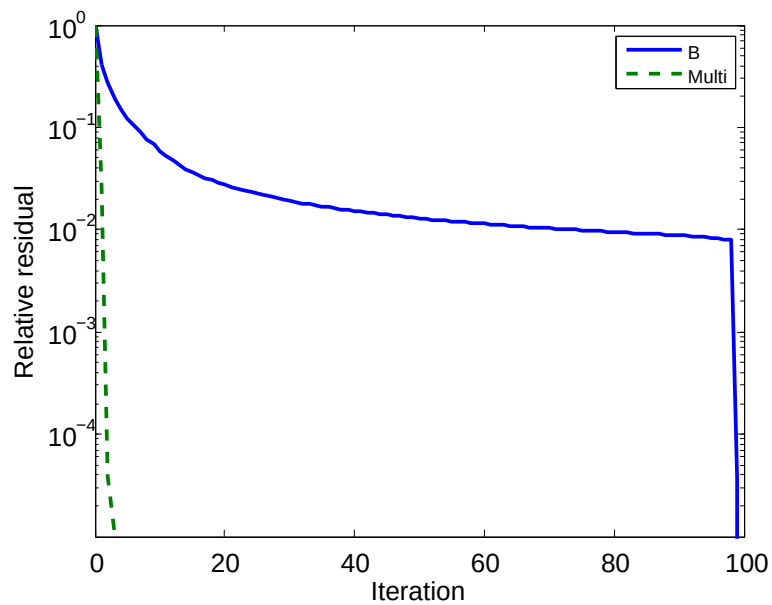


Figure 6: 1D example: Relative residuals for the quadratic finite element matrix and relative preconditioned residuals for the system with geometric multigrid preconditioning with $n = 101$ and $\epsilon = 0.1$.

well distributed) eigenvalues that are important in determining the convergence of I -GMRES for this preconditioned system.

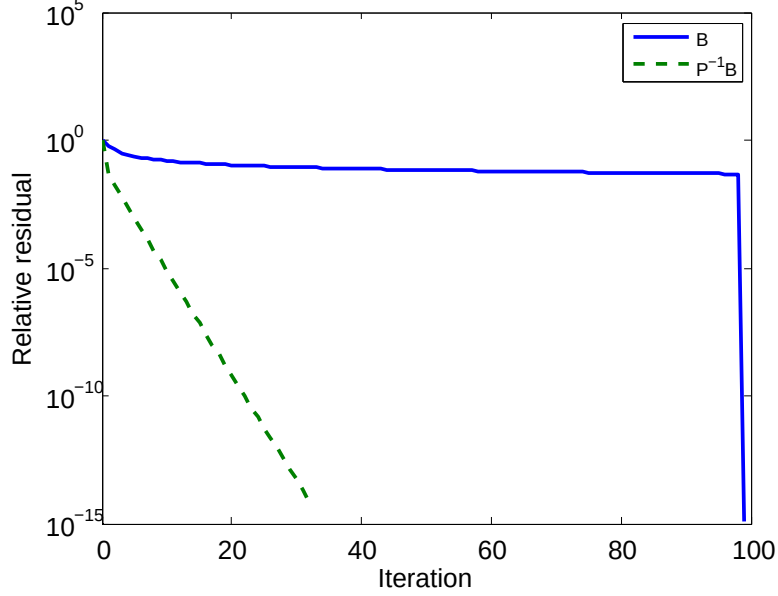


Figure 7: 1D example: Relative residuals for the quadratic finite element matrix and preconditioned relative preconditioned residuals for the preconditioned system with $n = 101$ and $\epsilon = 0.001$.

This fast convergence is certainly not suggested given the differences between the two matrices P and B . While the upper and lower bandwidths of B are both 3 (see [19, §1.2.1] for definitions of upper and lower bandwidths), P is effectively bidiagonal since its superdiagonal elements are of the order of machine epsilon (approximately 10^{-16}). Additionally, the diagonal elements of P are all 1 and its subdiagonal is the vector $(0, -1, \dots, -1, 0)^T$; the part of the matrix corresponding to interior nodes is essentially a transposed and scaled Jordan block. We conjecture that P and B are self-adjoint with respect to symmetric bilinear forms that are nearby, and that this explains the rapid convergence displayed.

Example 5 Our final advection-diffusion problem is the two-dimensional example 3.3 from [12] in the domain $\Omega = [-1, 1] \times [-1, 1]$:

$$-\epsilon \nabla u + \mathbf{w} \cdot \nabla u = 0,$$

where $\mathbf{w} = (-\sin \frac{\pi}{6}, \cos \frac{\pi}{6})^T$, with Dirichlet boundary conditions

$$\begin{aligned} u(-1, y) &= 0 & u(1, y) &= 1 \\ u(x, -1) &= \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} & u(x, 1) &= 1. \end{aligned}$$

With $\epsilon = 0.05$ the Q_1 mesh Peclet number is 0.72 and we apply Galerkin finite elements on a 16×16 grid to obtain the Q_2 finite element discretisation. Similarly to the one-dimensional case, the Q_2 biquadratic finite element matrix is denoted B and the Q_1 bilinear finite element matrix is P . Both P and B are of size 1089×1089 and both have complex eigenvalues. Thus, neither matrix is self-adjoint with respect to an inner product and Theorem 8 is not applicable. However, as in Example 4, P reduces significantly the number of I -GMRES iterations required as can be seen in Figure 8. In particular, the preconditioned system does not suffer the initial stagnation phase of unpreconditioned I -GMRES as analysed in [32]. The condition number of the matrix of eigenvectors of this preconditioned system is of the order of 10^7 , indicating that (1.2) is a poor estimate in this case.

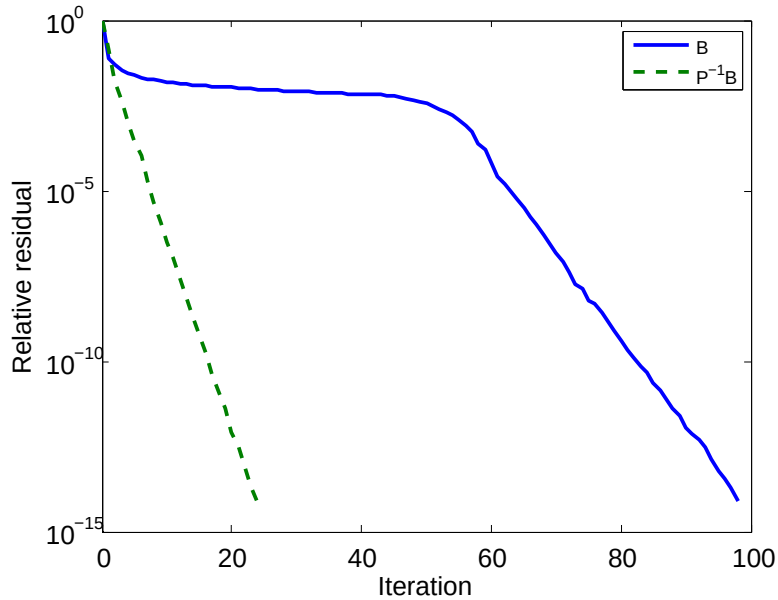


Figure 8: 2D example: Relative residuals for the quadratic finite element matrix and relative preconditioned residuals for the preconditioned system with $n = 1089$ and $\epsilon = 0.05$.

Example 6 Finally, we return to the Sherman3 example in §1 and explain the rapid convergence of I -GMRES applied to the ILU(0) preconditioned system shown in Figure 1. The Sherman3 matrix B has real eigenvalues and so is self-adjoint in an inner product $\langle \cdot, \cdot \rangle_H$. The ILU(0) preconditioner P is also self-adjoint with respect to an inner product, $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. If these inner products are close we would expect E in (4.2) to be small and thus, using our theory, that convergence is dominated by eigenvalues. As in Example 4 we cannot obtain the matrix E in (4.2) directly but we can use the same indirect measure of the size of E that was used in that example. That is, with $\mathcal{H}P = \mathcal{S}^{-T}\mathcal{S}^{-1}$, where $\mathcal{S}$ is again computed by Matlab's `sqrtrm` function, and $G = \mathcal{S}^{-1}P^{-1}B\mathcal{S}$, we find that $\|G - \text{diag}(\text{diag}(G))\|_2 = 0.9775$ which is small relative to $n = 5005$; we conclude that E is

small. Consequently, the preceding theory indicates that the convergence of I -GMRES essentially depends chiefly on the (well distributed) eigenvalues of the preconditioned matrix.

6 Conclusions

In the context of preconditioning for the iterative solution of nonsymmetric linear systems, for a class of nonsymmetric matrices we have identified a mathematical property of preconditioning matrices which ensures that convergence of any minimum residual method (such as GMRES) depends essentially only on the eigenvalues of the preconditioned system. This is in contrast to previous/existing convergence bounds which allow for the often observed initial stagnation, and is more consistent with convergence properties for symmetric systems. More generally, though we have not yet been able to prove this, we observe similar behaviour for nonsymmetric systems which do not lie in the required class when appropriate preconditioning is used.

Acknowledgment. We are extremely grateful to Michele Benzi, Jörg Liesen, Miro Rozložník, Valeria Simoncini and Zdeněk Strakoš for their helpful comments and suggestions that greatly improved the manuscript.

References

- [1] M. ARIOLI, V. PTÁK, AND Z. STRAKOŠ, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.
- [2] S. F. ASHBY, M. J. HOLST, T. A. MANTEUFFEL, AND P. E. SAYLOR, *The role of the inner product in stopping criteria for conjugate gradient iterations*, BIT, 41 (2001), pp. 26–52.
- [3] S. F. ASHBY, T. A. MANTEUFFEL, AND P. E. SAYLOR, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.
- [4] M. BENZI AND V. SIMONCINI, *On the eigenvalues of a class of saddle point matrices*, Numer. Math., 103 (2006), pp. 173–196.
- [5] M. BENZI AND M. TŮMA, *A comparative study of sparse approximate inverse preconditioners*, Appl. Numer. Math., 30 (1999), pp. 305–340.
- [6] A. CHAPMAN, Y. SAAD, AND L. WIGTON, *High-order ILU preconditioners for CFD problems*, Int. J. Numer. Meth. Fluids, 33 (2000), pp. 767–788.
- [7] T. A. DAVIS, *Direct Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2006.
- [8] I. S. DUFF, A. M. ERISMAN, AND J. K. REID, *Direct Methods for Sparse Matrices*, Oxford University Press, Oxford, 1989.

- [9] I. S. DUFF, R. G. GRIMES, AND J. G. LEWIS, *Sparse matrix test problems*, ACM Trans. Math. Softw., 15 (1989), pp. 1–14.
- [10] M. EIERMANN, *Fields of values and iterative methods*, Linear Algebra Appl., 180 (1993), pp. 167–197.
- [11] M. EIERMANN AND O. G. ERNST, *Geometric aspects in the theory of Krylov subspace methods*, Acta Numer., 10 (2001), pp. 251–312.
- [12] H. C. ELMAN, D. J. SILVESTER, AND A. J. WATHEN, *Finite Elements and Fast Iterative Solvers: with applications in incompressible fluid dynamics*, Oxford University Press, Oxford, 2005.
- [13] A. ESSAI, *Weighted FOM and GMRES for solving nonsymmetric linear systems*, Numer. Algorithms, 18 (1998), pp. 277–292.
- [14] V. FABER, J. LIESEN, AND P. TICHÝ, *The Faber-Manteuffel theorem for linear operators*, SIAM J. Numer. Anal., 46 (2008), pp. 1323–1337.
- [15] V. FABER AND T. MANTEUFFEL, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.
- [16] R. FLETCHER, *Conjugate gradient methods for indefinite systems*, in Numerical Methods Dundee, G. A. Watson, ed., Springer-Verlag, New York, 1976, pp. 73–89.
- [17] R. W. FREUND AND N. M. NACHTIGAL, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.
- [18] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Indefinite Linear Algebra and Applications*, Birkhäuser Verlag, Basel, 2005.
- [19] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The John Hopkins University Press, Maryland, USA, third ed., 1996.
- [20] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.
- [21] A. GREENBAUM, V. PTÁK, AND Z. STRAKOŠ, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.
- [22] A. GREENBAUM AND Z. STRAKOŠ, *Matrices that generate the same Krylov varieties*, in Recent Advances in Iterative Methods, G. Golub, A. Greenbaum, and M. Luskin, eds., vol. 60 of IMA Volumes in Mathematics and its Applications, New York, 1994, Springer, pp. 95–118.
- [23] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436.

- [24] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.
- [25] C. JAGELS AND L. REICHEL, *The isometric Arnoldi process and an application to iterative solution of large linear systems*, in *Iterative Methods in Linear Algebra*, R. Beauwens and P. de Groen, eds., Elsevier, Amsterdam, 1992, pp. 361–369.
- [26] C. F. JAGELS AND L. REICHEL, *A fast minimal residual algorithm for shifted unitary matrices*, *Numer. Linear Algebra Appl.*, 1 (1994), pp. 555–570.
- [27] Z. JIA AND B. ZHU, *A power sparse approximate inverse preconditioner for large sparse linear systems*, *Numer. Linear Algebra Appl.*, 16 (2009), pp. 259–299.
- [28] W. D. JOUBERT AND D. M. YOUNG, *Necessary and sufficient conditions for the simplification of generalized conjugate gradient algorithms*, *Linear Algebra Appl.*, 88/89 (1987), pp. 449–485.
- [29] H.-B. LI, T.-Z. HUANG, Y. ZHANG, X.-P. LIU, AND H. LI, *On some new approximate factorization methods for block tridiagonal matrices suitable for vector and parallel processors*, *Math. Comput. Simulat.*, 79 (2009), pp. 2135–2147.
- [30] J. LIESEN AND B. PARLETT, *On nonsymmetric saddle point matrices that allow conjugate gradient iterations*, *Numer. Math.*, 108 (2008), pp. 605–624.
- [31] J. LIESEN AND Z. STRAKOŠ, *On optimal short recurrences for generating orthogonal Krylov subspace bases*, *SIAM Rev.*, 50 (2008), pp. 485–503.
- [32] J. LIESEN AND Z. STRAKOŠ, *GMRES convergence analysis for a convection-diffusion model problem*, *SIAM J. Sci. Comput.*, 26 (2005), pp. 1989–2009.
- [33] 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.
- [34] G. MEURANT AND Z. STRAKOŠ, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, *Acta Numer.*, 15 (2006), pp. 471–542.
- [35] N. M. NACHTIGAL, S. C. REDDY, AND L. N. TREFETHEN, *How fast are non-symmetric matrix iterations?*, *SIAM J. Matrix Anal. Appl.*, 13 (1992), pp. 778–795.
- [36] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, *SIAM J. Numer. Anal.*, 12 (1975), pp. 617–629.
- [37] A. RAMAGE, *A multigrid preconditioner for stabilised discretisations of advection-diffusion problems*, *J. Comput. Appl. Math.*, 110 (1999), pp. 187–203.
- [38] Y. SAAD, *Iterative methods for sparse linear systems*, SIAM, Philadelphia, 2003.

- [39] Y. SAAD AND M. H. SCHULTZ, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. and Stat. Comput., 7 (1986), pp. 856–869.
- [40] H. SADOK AND D. B. SZYLD, *A new look at CMRH and its relation to GMRES*, Tech. Report 10-04-25, Department of Mathematics, Temple University, Apr. 2010.
- [41] V. SIMONCINI. Personal communication, 2010.
- [42] V. SIMONCINI AND D. B. SZYLD, *The effect of non-optimal bases on the convergence of Krylov subspace methods*, Numer. Math., 100 (2005), pp. 711–733.
- [43] V. SIMONCINI AND D. B. SZYLD, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59.
- [44] G. L. G. SLEIJPEN AND D. R. FOKKEMA, *Bi-CGStab(ℓ) for linear equations involving unsymmetric matrices with complex spectrum*, Elec. Trans. Numer. Anal., 1 (1993), pp. 11–32.
- [45] P. SONNEVELD, *CGS, a fast Lanczos-type solver for nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 36–52.
- [46] M. STOLL AND A. J. WATHEN, *Combination preconditioning and the Bramble-Pasciak⁺ preconditioner*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 582–608.
- [47] O. TAUSKY, *The role of symmetric matrices in the study of general matrices*, Linear Algebra Appl., 5 (1972), pp. 147–154.
- [48] O. TAUSKY AND W. V. PARKER, *Problem 4846*, Amer. Math. Monthly, 67 (1960), pp. 192–193.
- [49] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behaviour of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, 2005.
- [50] F. UHLIG, *On the nonsingular symmetric factors of a real matrix*, Linear Algebra Appl., 8 (1974), pp. 351–354.
- [51] H. A. VAN DER VORST, *Bi-CGStab: A fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.
- [52] H. A. VAN DER VORST AND Q. YE, *Residual replacement strategies for Krylov subspace iterative methods for the convergence of true residuals*, SIAM J. Sci. Comput., 22 (2000), pp. 835–852.