New conditions for non-stagnation of minimal residual methods

Valeria Simoncini and Daniel B. Szyld

Report 07-4-17
April 2007

NEW CONDITIONS FOR NON-STAGNATION OF MINIMAL RESIDUAL METHODS*

VALERIA SIMONCINI† AND DANIEL B. SZYLD‡

Abstract. In the context of the solution of large linear systems, a condition guaranteeing that a minimal residual Krylov subspace method makes some progress, i.e., that it does not stagnate, is that the symmetric part of the coefficient matrix be positive definite. This condition results in a well-established bound due to Elman, for the convergence rate of the iterative method. This bound is usually pessimistic. Nevertheless, it has been extensively used, e.g., to show that for certain preconditioned problems, the convergence of GMRES (or of other minimal residual methods) is independent of the underlying mesh size of the discretized partial differential equation. In this paper we introduce more general non-stagnation conditions on the coefficient matrix, which do not require the symmetric part of the coefficient matrix to be positive definite, and that guarantee, for example, the non-stagnation of restarted GMRES for certain values of the restarting parameter.

1. Introduction. Minimal residual Krylov subspace methods, and in particular in the implementation given in GMRES [27], are routinely employed for the solution of large linear systems of the form $Ax = b$, and especially of those systems arising in the discretization of partial differential equations; see, e.g., [12], [26], [32]. Let x_0 be an initial vector, and x_m be the approximate solution after m iterations, with corresponding residual $r_m = b - Ax_m$. In these methods, the residual norm is non-increasing, i.e., $\|r_m\| \leq \|r_{m-1}\|$. In some instances, though, there is possible stagnation, that is $\|r_m\| = \|r_{m-1}\|$ holds for some m; see, e.g., [8], [19], [30], [38], [39] for examples and discussion of this issue.

Elman [11] studied conditions for non-stagnation of minimal residual methods (and thus applicable to GMRES), and obtained a useful bound on the associated residual norm; see also [10]. Let $H = \mathcal{H}(A) =: (A + A^T)/2$ be the symmetric part of A. If H is positive definite, i.e., if for real vectors x,

$$c = \min_{x \neq 0} \frac{(x, Ax)}{(x, x)} = \min_{x \neq 0} \frac{(x, Hx)}{(x, x)} > 0, \tag{1.1}$$

then, there is no stagnation, and furthermore,

$$\|r_m\| \leq \left(1 - \frac{c^2}{C^2}\right)^{m/2} \|r_0\|, \tag{1.2}$$

where

$$C = \|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}, \qquad c \leq C. \tag{1.3}$$

From (1.1) one has that $\rho := \left(1 - c^2/C^2\right)^{1/2} < 1$. Elman's results indicate that if (1.1) holds, then, the residual norm decreases at each iteration at least by the constant factor ρ. We note that from (1.1) and (1.3), it is immediate that if H is negative definite ($-H$ is positive definite), then the same results apply, i.e., there is no stagnation and (1.2) holds.

---

*This version dated 17 April 2007

†Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, I-40127 Bologna, Italy; and CIRSA, Ravenna (valeria@dm.unibo.it).

‡Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, Pennsylvania 19122-6094, USA (szyld@temple.edu).

The convergence of GMRES is in most cases superlinear (see, e.g., [31], [36]), while the bound (1.2) indicates linear convergence. Thus, it is generally understood that (1.2) may be very pessimistic as a bound. Moreover, if $\rho \approx 1$ the bound may (possibly erroneously) predict a very small residual norm reduction. Nevertheless, this bound is widely used in certain contexts. In particular, when the matrix A represents a discretization of a differential operator, researchers have looked for preconditioners, such that the quantities c and C defined in (1.1), (1.3), can be bounded independently of the mesh size of the discretization; see, e.g., [33, §5.3], [35, §§2.3, 3.6]. These bounds guarantee that a finer discretization does not increase the work per degree of freedom beyond a bounded quantity. It turns out that for the (preconditioned) coefficient matrix to satisfy (1.1), certain conditions on the discretization may need to be imposed, and this limits the applicability of the bound (1.2); see, e.g., [1], [35, §3.6]. In fact, in [6] a simple discretized one-dimensional partial differential equation is presented such that the coefficient matrix obtained with overlapping additive Schwarz preconditioning cannot satisfy (1.1).

A natural question is whether we can formulate some other conditions for non-stagnation, and an associated convergence bound that is applicable to matrices whose symmetric part is not positive definite, i.e., in the case where (1.1) does not hold. In this paper we answer this question in the affirmative, providing new conditions for non-stagnation which relate the symmetric part of the matrix A, i.e., $H = \mathcal{H}(A)$, with its skew-symmetric part, $S = \mathcal{S}(A) := (A - A^T)/2$. In many cases the new conditions are computable *a priori*, or can be inferred from the nature of the problem.

The rest of the paper is organized as follows. In the next section, we have some preliminary discussion and describe work by other authors studying conditions for non-stagnation, or related to bounds similar to (1.2). As we shall see, most of these bounds require that (1.1) hold, i.e., that $\mathcal{H}(A)$ be positive definite. In section 3, we present our new conditions together with a few elementary examples of their applicability, while in section 4 we discuss the new results and present additional illustrative examples.

In the preceeding expressions, as well as in the rest of this paper, the inner product is the Euclidean one $(v, w) = v^T w$, and the norm is the one associated with this inner product, i.e., the 2-norm $\|v\| = (v^T v)^{1/2}$. Elman's results, as well as all results in this paper carry over to any other inner product, and its induced norm, but for simplicity of the exposition we do not provide the details; cf. [8], [29], [34].

2. Preliminary and related results. In this section, we briefly discuss other results related to non-stagnation and convergence bounds for GMRES.

As already mentioned, the bound in (1.2) may be used to ensure convergence in a *restarted* process. Indeed, for nonsymmetric matrices, optimal minimal residual methods are usually characterized by large computational and memory requirements; these costs increase superlinearly with the number of iterations. For these reasons, methods such as GMRES are often stopped after a fixed number of iterations, and then restarted with the current approximation as initial guess. The estimate in (1.2) ensures that a minimal residual method will be capable to reduce the residual norm even after a very limited number of iterations, regardless of the properties of the initial guess. In this context, it is worth remarking that conditions such as (1.2) try to address *worst-case* scenarios. Indeed, it may be shown that after one iteration of a minimal residual iteration, it holds (see, e.g., [26, §5.3.2])

$$\|r_1\| = \sqrt{1 - \frac{(r_0^T A r_0)^2}{\|Ar_0\|^2\,\|r_0\|^2}}\,\|r_0\|,$$

therefore, for $\|r_1\|$ to be strictly less than $\|r_0\|$, it is sufficient that $r_0^T A r_0 \neq 0$ for the given vector r_0. Clearly, it is quite unlikely, although not impossible, that $r_k^T A r_k$ is exactly zero for some k when A is indefinite. This explains why minimal residual methods rarely show complete (namely at all iterations) stagnation in practice, even in the case of indefinite problems. On the other hand, classes of matrices for which complete stagnation occurs have been analyzed in detail; see [38].

If the matrix A is diagonalizable, other linear convergence bounds of a form similar to (1.2) are available, but including a factor which is the condition number of the eigenvector matrix of A; see. e.g., [10], [18, §3.2], [26, §6.11.14], [32]. See also [23] for improvements of these bounds in certain cases, and [15] for analogous bounds for non-diagonalizable matrices.

Other convergence bounds using the field of values $\mathcal{F}(A)$ were developed, where

$$\mathcal{F}(A) = \{\omega \in \mathbb{C} \;/\; \omega = \frac{(x, Ax)}{(x, x)}, \quad x \in \mathbb{C}^n, x \neq 0\},$$

and n is the order of the matrix. These bounds always assume that $0 \notin \mathcal{F}(A)$; see, e.g., [8, Corollary 6.2], [18, §3.2], [34]. It is precisely for those cases with $0 \in \mathcal{F}(A)$ that we look for new non-stagnation conditions. We refer to [13] for examples where each of the aforementioned bounds is more descriptive than the others. We also mention the paper [24] where bounds are presented in the case that the spectrum of A is contained in a half plane.

If A is normal, i.e., if $AA^T = A^T A$, then its field of values coincides with the convex hull of its set of eigenvalues $\sigma(A)$. For a non-normal matrix, the field of values can be, and it usually is, much larger than the convex hull of $\sigma(A)$; see, e.g., [20].

Following [16], we say that a matrix A is positive (negative) definite, if $x^T Ax > 0$ for all nonzero real vectors x (if $-A$ is positive definite). In this case, its field of values is completely contained in the right-half (left-half) plane $\mathbb{C}^+$ ($\mathbb{C}^-$), and Elman's bound (1.2) holds.

A minimal residual Krylov subspace method proceeds by finding at the mth iteration an approximation x_m, so that $x_m - x_0 \in \mathcal{K}_m = \mathrm{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{m-1} r_0\}$, and such that $\|r_m\| \leq \|b - Ax\|$ for all $x - x_0 \in \mathcal{K}_m$. Equivalently, letting $\mathcal{P}_m$ be the set of polynomials p of degree m satisfying $p(0) = 1$, we can write $r_m = p_m(A)r_0$, for $p_m \in \mathcal{P}_m$ such that $\|p_m(A)r_0\| \leq \|p(A)r_0\|$, for all $p \in \mathcal{P}_m$. This polynomial p_m is called the GMRES residual polynomial. It follows from this standard characterization that stagnation is avoided as soon as m is large enough so that p_m satisfies $\|p_m(A)r_0\| < \|r_0\|$. For a more intuitive argument, assume that A is not positive definite but that a power of it is, say A^k. Using Elman's results, a minimal residual method applied to the system $A^k x = r_0$ with zero initial guess would not stagnate. Denoting by q_1 the hypothetical residual polynomial after one minimal residual iteration with A^k, it would hold that $\|q_1(A^k)r_0\| < \|r_0\|$. We notice that $q_1 = q_1(\eta)$ applied to η^k has degree k and satisfies $q_1(0) = 1$. Therefore the residual after k iterations of a minimal residual method applied to A satisfies

$$\|r_k\| = \|p_k(A)r_0\| \leq \|q_1(A^k)r_0\| \leq \left(1 - \frac{c_k^2}{C_k^2}\right)^{1/2} \|r_0\| := \rho_k \|r_0\|,$$

where $c_k = \min\big(x, \mathcal{H}(A^k)x\big)/(x,x) > 0$, with $x \in \mathbb{R}^n$, $x \neq 0$, and $C_k = \|A^k\|$. Thus, after $j \cdot k$ iterations, the residual r_{jk} of the minimal residual method applied to A satisfies

$$\|r_{jk}\| \leq \rho_k \|r_{(j-1)k}\| \leq \rho_k^j \|r_0\|. \tag{2.1}$$

Therefore, if A^k is positive definite, the minimal residual method does not stagnate for more than k iterations, and after k iterations, the residual norm decreases by at least a factor ρ_k. The same argument can be used to show that if A^k is positive definite, then, the restarted version of a minimal residual vector with restart parameter k does not stagnate; this goes back to [11]. In summary, we have the following result.

PROPOSITION 2.1. *Let A^k be positive or negative definite for some $k \geq 1$. Then, a minimal residual method to solve $Ax = b$ does not stagnate for more than k iterations, and its residual satisfies* (2.1).

This result is in fact a special case of [17, Theorem 1], stated in the next section, where it holds using any polynomial of degree k, and also a special case of [39, Theorem 2.2] where it is shown for complex matrices.

If A is normal, one can characterize $\sigma(A)$ for which $\sigma(A^k) \subset \mathbb{C}^+$, and thus its convex hull is also contained in $\mathbb{C}^+$. It can be shown that if

$$\sigma(A) \subset \left\{ \omega \in \mathbb{C} \;/\; \omega = |\omega| e^{i\theta},\, \theta \in \bigcup_{j=0}^{k-1} [(-\pi + 4j)/2k, (\pi + 4j)/2k] \right\},$$

then $\sigma(A^k) \subset \mathbb{C}^+$, see, e.g., [39]. For the special case of $k = 2$, this is [18, Exercise 2.8] where it is shown that for A normal, A^2 is positive definite if $|Re(\lambda)| > |Im(\lambda)|$ for all $\lambda \in \sigma(A)$. In particular, if the matrix is symmetric indefinite, then A^2 is always positive definite and thus stagnation of minimal residual methods can only take place in not more than two consecutive iterations.

Finally, we refer to [2], [3], [22] and to [28], for convergence results that also require $\mathcal{H}(A)$ to be positive definite.

3. The new conditions.

We begin by stating the already mentioned result of Grcar [17].

THEOREM 3.1. *Let q be a polynomial of degree at most k, with $q(0) = 0$, and such that $\mathcal{H}(q(A))$ is positive or negative definite. Then for every x_0, the affine space $x_0 + \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$ contains a vector x for which $\|b - Ax\| \leq \rho \|r_0\|$, where*

$$\rho = \left(1 - \frac{\hat{c}^2}{\hat{C}^2}\right)^{1/2} < 1,$$

with $\hat{c} = \min\{|\lambda|\ ,\ \lambda \in \sigma\left(\mathcal{H}(q(A))\right)\}$ and $\hat{C} = \|q(A)\|$.

We observe that for the special case of $q(A) = A$, the hypothesis is that $\mathcal{H}(A)$ is definite, and one recovers the result (1.2). Similarly for $q(A) = A^k$ one has Proposition 2.1.

Our first result gives conditions so that $\mathcal{H}(q(A))$ is either positive or negative definite for the case $q(\eta) = \eta^2$. Thus, using Theorem 3.1 we then conclude that the GMRES residual does not stagnate for more than two iterations, and that GMRES(2), the restarted GMRES method with restarting parameter $k = 2$, does not stagnate.

THEOREM 3.2. *Let $H = \mathcal{H}(A)$ and $S = \mathcal{S}(A)$. Then, the following holds:*
1. For all real vectors x,

$$x^T A^2 x = \|Hx\|^2 - \|Sx\|^2. \tag{3.1}$$

2. If H is nonsingular, then $\mathcal{H}(A^2)$ is positive definite if and only if

$$\|SH^{-1}\| < 1. \tag{3.2}$$

3. If S is nonsingular, then $\mathcal{H}(A^2)$ is negative definite if and only if

$$\|HS^{-1}\| < 1. \tag{3.3}$$

Proof. We have $A = H+S$, so that $A^2 = H^2+HS+SH+S^2$. Observe that since $S^T = -S$, then, $HS+SH$ is skew-symmetric. We then have for all real vectors x,

$$x^T A^2 x = x^T(H^2 + S^2)x = x^T H^T Hx - x^T S^T Sx = \|Hx\|^2 - \|Sx\|^2,$$

which is (3.1). To show the second statement, let

$$H^{-T} S^T S H^{-1} = Q\Lambda Q^T, \quad \text{with } Q^TQ = I, \quad \text{and} \quad \Lambda \geq O \quad \text{diagonal}; \tag{3.4}$$

let $y = Hx$, and $y = Qz$ for some z. We have then that for all real vectors x

$$\begin{aligned} x^T A^2 x &= x^T H^T Hx - x^T S^T Sx = y^T y - y^T H^{-T} S^T S H^{-1} y \\ &= z^T z - z^T \Lambda z = z^T (I - \Lambda) z. \end{aligned} \tag{3.5}$$

It follows from (3.4) that the diagonal entries of Λ are the squares of the singular values of SH^{-1}, that is $\lambda_i = \sigma_i^2$, $i = 1, \ldots, n$. Thus, from (3.5) we have that for all real vectors x, $x^T A^2 x = z^T (I - \Lambda) z > 0$ if and only if $\lambda_i < 1$, which in turn is equivalent to requiring that $\sigma_i < 1$ for all i, that is $\|SH^{-1}\| < 1$.

The third statement is shown in a similar manner. ◻

We note that in the conditions (3.2) or (3.3) one can interchange the order of the factors. This is because for any symmetric matrix H and any skew-symmetric matrix S, it holds that $\|HS\| = \|(HS)^T\| = \|S^T H^T\| = \|-SH\| = \|SH\|$.

It is very easy to construct examples where (3.2) or (3.3) hold, but (1.1) does not. Two such cases follow.

EXAMPLE 3.3. Any matrix $A = H + S$ with H indefinite and nonsingular, and S skew-symmetric and orthogonal, cannot be used in the context of (1.1), but satisfy, say, (3.2), if the eigenvalues of H are greater than one in modulus. Indeed, in this case, $\|SH^{-1}\| = \|H^{-1}\| < 1$.

EXAMPLE 3.4. We next consider the following non-diagonalizable matrix

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 4 & 1/2 \\ 0 & 1/2 & 4 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} = H + S$$

Here H is indefinite with eigenvalues $\{-1, 7/2, 9/2\}$, while $\|SH^{-1}\| = 1/7 < 1$ and thus Theorem 3.2 applies.

Intuitively, the second result of Theorem 3.2 says that if H is nonsingular and if it "dominates" S, then A^2 is positive definite. This fact is made more explicit in the following result, which gives a simple sufficient condition for $\|SH^{-1}\| < 1$ to hold.

Corollary 3.5. *Let $\lambda_i(M)$ be the ith eigenvalue of the matrix M. If $|\lambda_i(H)| > |\lambda_j(S)|$ for $i,j=1,\ldots,n$, then $\mathcal{H}(A^2)$ is positive definite.*

Proof. For any real vector x, we have

$$\|Hx\|^2 \geq \lambda_{\min}(H)^2\|x\|^2 > \lambda_{\max}(S)^2\|x\|^2 \geq \|Sx\|^2.$$

In view of (3.1) the result follows. □

A corresponding result holds when $|\lambda_i(H)| < |\lambda_j(S)|$ for $i,j=1,\ldots,n$. Thus, if S is nonsingular and if it "dominates" H, then A^2 is negative definite.

We now obtain conditions for $\mathcal{H}(A^4)$ to be either positive or negative definite, and this would imply that the GMRES residual does not stagnate for more than four iterations.

Theorem 3.6. *Let $H=\mathcal{H}(A)$ and $S=\mathcal{S}(A)$. Then, the following holds:*

1. *If H^2+S^2 is nonsingular, then $\mathcal{H}(A^4)$ is positive definite if and only if $\|(HS+SH)(H^2+S^2)^{-1}\| < 1$.*
2. *If $HS+SH$ is nonsingular, then $\mathcal{H}(A^4)$ is negative definite if and only if $\|(H^2+S^2)(HS+SH)^{-1}\| < 1$.*

Proof. As we have seen, $A^2=(H^2+S^2)+(HS+SH)=\mathcal{H}(A^2)+\mathcal{S}(A^2)$. Thus, the result follows applying Theorem 3.2 to A^2. □

Example 3.7. Let

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & -10 \\ 0 & 10 & 0 \end{bmatrix}.$$

It is easy to see that both $H=\mathcal{H}(A)$ and $S=\mathcal{S}(A)$ are singular, and thus, neither condition in Theorem 3.2 is satisfied. On the other hand, we have that (H^2+S^2) is nonsingular, and $\|(HS+SH)(H^2+S^2)^{-1}\| \approx 0.329$. Thus Theorem 3.6 applies.

Remark 3.8. *We can continue in the same manner, and apply Theorem 3.2 to other powers of A, but then, the conditions obtained are not easy or practical to check. For example one has that*

$$\begin{aligned} A^3 &= \left[H(H^2+S^2)+S(HS+SH)\right]+\left[S(H^2+S^2)+H(HS+SH)\right] \\ &= \mathcal{H}(A^3)+\mathcal{S}(A^3). \end{aligned}$$

Theorem 3.1 says that positive definiteness may be obtained by means of a polynomial q of degree k such that $q(0)=0$, and we have shown that under suitable conditions this may be obtained for $q(\eta)=\eta^k$, $k=2,4$. Although they seem hard to find explicitly, more general polynomials of the same degree k may satisfy the definiteness condition. In the following example, we show that this is indeed the case.

Example 3.9. Let $q(\eta)=\eta^2+\alpha\eta$ with $\alpha>0$, and notice that

$$\eta^2+\alpha\eta-1<0 \quad \text{iff} \quad \eta \in \left(-\frac{\alpha}{2}-\sqrt{\frac{\alpha^2}{4}+1}, -\frac{\alpha}{2}+\sqrt{\frac{\alpha^2}{4}+1}\right) := (\ell_1,\ell_2),$$

with $\ell_1<0$ and $\ell_2>0$. Let $A=H+S$ with S skew-symmetric and orthogonal, and H symmetric with both positive and negative eigenvalues in (ℓ_1,ℓ_2). Using the eigenvalue decomposition $H=U\Lambda U^T$, $\Lambda=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ and the orthogonality of

S, for any $x \neq 0$ it follows

$$\begin{aligned} x^T q(A)x &= x^T(H^2 + S^2 + \alpha H)x \\ &= x^T H^2 x - x^T x + \alpha x^T H x \\ &= z^T \Lambda^2 z - z^T z + \alpha z^T \Lambda z = z^T(\Lambda^2 + \alpha\Lambda - I)z < 0. \end{aligned}$$

Here we used $z = U^T x$. The final inequality follows from the fact that the matrix in parenthesis is diagonal, and that $\lambda_i \in (\ell_1, \ell_2)$ for all i's, so that the matrix is negative definite. Note that for $|\lambda_1| \leq \ldots \leq 1 \leq \ldots \leq |\lambda_n|$, the quantity $x^T A^2 x = z^T(\Lambda^2 - I)z$ remains indefinite. As a numerical example, we take $\alpha = 10$ so that $\ell_1 \approx -10.099$ and $\ell_2 \approx 0.09902$. For

$$A = H + S, \quad H = \begin{bmatrix} -8 & 0 \\ 0 & 0.01 \end{bmatrix}, \quad S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

we obtain $\sigma(H) = \{-8, 0.01\}$ and $\sigma(\mathcal{H}(A^2)) = \{-0.9999, 63\}$, whereas $\sigma(\mathcal{H}(A^2 + \alpha A)) = \{-17, -0.8999\}$.

A completely analogous example showing that $x^T q(A)x > 0$ for all $x \neq 0$ may be obtained for $\sigma(H)$ contained in $\mathbb{R} \setminus [\ell_1, \ell_2]$. We have thus derived a class of matrices A for which $q(A)$ is definite, and A^2 may not be.

4. Discussion and additional examples. We begin by discussing the conditions of Theorem 3.2. Observe that having $\|H^{-1}S\| < 1$ implies that the matrix $H^{-1}A = I + H^{-1}S$ has its spectrum in the right half plane. This fact was used to consider H as a precondtioner; see [7], [37], and also [2].

The splitting $A = H - (-S)$ was used to generate convergent classical stationary iterative methods, often with some relaxation parameters or acceleration so that $\|H^{-1}S\| < 1$; see, [5], [9], [25].

We mention that in [21] a result is given for the case $H = 0$, i.e., for nonsingular skew-symmetric matrices, showing non-stagnation of GMRES(k) for $k \geq 2$. This result follows from the normality of S and also from Theorem 3.2 since in this case $\|HS^{-1}\| = 0 < 1$.

Example 3.7 has the 2×2 block structure typical of matrices stemming from saddle point problems. Indeed, other examples of this type may be constructed all taking the form

$$M = \begin{bmatrix} A & B \\ -B^T & 0 \end{bmatrix},$$

where A is $n \times n$ and symmetric, while B is a full rank $n \times m$ matrix; see, e.g., [4]. Note that both $H = \mathcal{H}(M)$ and $S = \mathcal{S}(M)$ are singular for $n > m$. For the case where $A = \mu I$, $\mu > 0$ as in [14], and assuming $\mu^2 I - B^T B$ is nonsingular, algebraic calculations show that

$$\|(HS + SH)(H^2 + S^2)^{-1}\| = \max_{\sigma_i} \left\{ \frac{\mu\sigma_i}{|\mu^2 - \sigma_i^2|}, \frac{\mu}{\sigma_i} \right\},$$

where σ_i, $i = 1, \ldots, m$ are the (nonzero) singular values of B. Therefore, for μ such that $\|(HS + SH)(H^2 + S^2)^{-1}\| < 1$, Theorem 3.6 applies. An explicit discussion of stagnation for this saddle point matrix when $A = \mu I$ can be found in [14].

We conclude with some examples with elliptic operators.

EXAMPLE 4.1. We consider the class of matrices stemming from the centered finite difference discretization with mesh size 1/41 of the differential operator

$$L(u) = (\alpha u_x)_x + (\beta u_y)_y + \gamma u_x + \delta u_y + \eta u \tag{4.1}$$

on the unit square, with Dirichlet boundary conditions, giving rise to a matrix of dimension $n = 1600$. In Table 4.1 we report the minimum eigenvalue of H, and also $\|SH^{-1}\|$, for some choices of the coefficients. In all cases, H is indefinite but the condition (3.2) of Theorem 3.2 holds.

α	β	γ	δ	η	$\lambda_{\min}(H)$	$\|SH^{-1}\|$
$-\exp(-xy)$	$-\exp(xy)$	1	1	-100	-0.04719	0.6194
-1	-1	$1/(.1x + 100y)$	0	-100	-0.04775	0.1577
-1	-1	$-1/10(x - y)$	0	-100	-0.04772	0.1838
-1	-1	$-1/10(x + y)$	0	-100	-0.04772	0.5819
-1	-1	-0.2	0	-100	-0.04781	0.5811

TABLE 4.1
Coefficients and corresponding values of $\lambda_{\min}(H)$ and $\|SH^{-1}\|$ for the matrix associated with the operator in (4.1).

Example 4.1 shows that for the condition (3.2) to hold it is sufficient that the symmetric part of the operator "dominates" the skew-symmetric one, as discussed in the previous section. Therefore, further test matrices may be obtained by appropriately choosing the coefficients γ and δ in (4.1).

As a final remark, we note that the new results may be used in the solution of (preconditioned) linear systems stemming from discretized partial differential equations. If for those problems it could be shown that $\hat{c}$ and $\hat{C}$ in Theorem 3.1 are independent of the underlying mesh size then the worst-case convergence rate bound of GMRES applied to these problems would be valid for all possible mesh refinements.

Acknowledgement. Work on this paper was supported in part by the U.S. Department of Energy under grant DE-FG02-05ER25672.

REFERENCES

[1] Paola F. Antonietti and Blanca Ayuso. Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. Technical Report PV-20, IMATI-CNR, Pavia, 2005. *M2AN Mathematical Modelling and Numercal Analysis*, to appear.

[2] Owe Axelsson. A generalized Conjugate Gradient, least squares method. *Numerische Mathematik*, 51:209–227, 1987.

[3] Bernhard Beckermann, Sergei A. Goreinov, and Eugene E. Tyrtyshnikov. Some remarks on the Elman estimate for GMRES. *SIAM Journal on Matrix Analysis and Applications*, 27:772–778, 2006.

[4] Michele Benzi, Gene H. Golub, and Jörg Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[5] Muddun Bhuruth. A note on Hermitian splitting induced relaxation methods for convection-diffusion equations. *Numerical Methods for Partial Differential Equations*, 14:582–591, 1998.

[6] Xiao-Chuan Cai and Jun Zou. Some observations on the l^2 convergence of the additive Schwarz preconditioned GMRES method. *Numerical Linear Algebra with Applications*, 9:379–397, 2002.

[7] Paul Concus and Gene H. Golub. A generalized Conjugate Gradient method for nonsymmetric linear equations. In Roland Glowinski and Jacques-Louis Lions, editors, *Computing Methods in Applied Science and Engineering*, volume 134 of *Lecture Notes in Economics and Mathematical Systems*, pages 56–65. Springer, New York, 1976.

[8] Michael Eiermann and Oliver G. Ernst. Geometric aspects in the theory of Krylov subspace methods. *Acta Numerica*, 10:251–312, 2001.

[9] Michael Eiermann, Wilhelm Niethammer, and Richard S. Varga. Acceleration of relaxation methods for non-Hermitian linear systems. *SIAM Journal on Matrix Analysis and Applications*, 13:979–991, 1992.

[10] Stanley C. Eisenstat, Howard C. Elman, and Martin H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis*, 20:345–357, 1983.

[11] Howard C. Elman. *Iterative methods for large sparse nonsymmetric systems of linear equations.* PhD thesis, Department of Computer Science, Yale University, New Haven, CT, 1982. Research Report #229.

[12] Howard C. Elman, David J. Silvester, and Andrew J. Wathen. *Finite Elements and Fast Iterative Solvers, with Applications in Incompressible Fluid Dynamics*, volume 21 of *Numerical Mathematics and Scientific Computation.* Oxford University Press, Oxford and New York, 2005.

[13] Mark Embree. *Convergence of Krylov subspace methods for non-normal matrices.* PhD thesis, Oxford University Computing Laboratory, Numerical Analysis Group, Michaelmas Term, 1999.

[14] Bernd Fischer, Alison Ramage, David J. Silvester, and Andrew J. Wathen. Minimum residual methods for augmented systems. *BIT Numerical Mathematics*, 38:527–543, 1998.

[15] Roland W. Freund. Quasi-kernel polynomials and convergence results for quasi-minimal residual iterations. In Dietrich Braess and Larry L. Schumaker, editors, *Numerical Methods in Approximation Theory, Vol. 9*, pages 77–95. Birkhäuser, Basel, 1992.

[16] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* The John Hopkins University Press, Baltimore, third edition, 1996.

[17] Joseph F. Grcar. Operator coefficient methods for linear equations. Technical Report SAND89-8691, Sandia National Laboratories, November 1989. Available at `http://seesar.lbl.gov/ccse/Publications/sepp/ocm/SAND89-8691.pdf`.

[18] Anne Greenbaum. *Iterative Methods for Solving Linear Systems*, volume 17 of *Frontiers in Applied Mathematics.* SIAM, Philadelphia, 1997.

[19] Anne Greenbaum, Vlastimil Ptàk, and Zdeněk Strakoš. Any nonincreasing convergence curve is possible for GMRES. *SIAM Journal on Matrix Analysis and Applications*, 17:95–118, 1996.

[20] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[21] Khalide Jbilou and Hassane Sadok. Analysis of some vector extrapolation methods for solving systems of linear equations. *Numerische Mathematik*, 70:73–89, 1995.

[22] Wayne D. Joubert. On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems. *Numerical Linear Algebra with Applications*, 1:427–447, 1994.

[23] Jörg Liesen. Computable convergence bounds for GMRES. *SIAM Journal on Matrix Analysis and Applications*, 21:882–903, 2000.

[24] Thomas A. Manteuffel. The Tchebychev iteration for nonsymmetric linear systems. *Numerische Mathematik*, 28:307–327, 1977.

[25] Wilhelm Niethammer and Richard S. Varga. Relaxation methods for non-Hermitian linear systems. *Results in Mathematics*, 16:308–320, 1989.

[26] Yousef Saad. *Iterative Methods for Sparse Linear Systems.* The PWS Publishing Company, Boston, 1996. Second edition, SIAM, Philadelphia, 2003.

[27] Yousef Saad and Martin H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7:856–869, 1986.

[28] Hassane Sadok. Analysis of the convergence of the minimal and the orthogonal residual methods. *Numerical Algorithms*, 40:201–216, 2005.

[29] Marcus Sarkis and Daniel B. Szyld. Optimal left and right additive Schwarz preconditioning for minimal residual methods with Euclidean and energy norms. *Computer Methods in Applied Mechanics and Engineering*, 196:1612–1621, 2007.

[30] Valeria Simoncini. A new variant of restarted GMRES. *Numerical Linear Algebra with Applications*, 6:61–77, 1999.

[31] Valeria Simoncini and Daniel B. Szyld. On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods. *SIAM Review*, 47:247–272, 2005.

[32] Valeria Simoncini and Daniel B. Szyld. Recent computational developments in Krylov subspace methods for linear systems. *Numerical Linear Algebra with Applications*, 14:1–59, 2007.

[33] Barry F. Smith, Petter E. Bjørstad, and William D. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, Cambridge, New York, and Melbourne, 1996.

[34] Gerhard Starke. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Numerische Mathematik*, 78:103–117, 1997.

[35] Andrea Toselli and Olof B. Widlund. *Domain Decomposition Methods - Algorithms and Theory*, volume 34 of *Springer Series in Computational Mathematics.* Springer, Berlin and Heidelberg, 2005.

[36] Henk A. van der Vorst and C. (Kees) Vuik. The superlinear convergence behaviour of GMRES. *Journal of Computational and Applied Mathematics*, 48:327–341, 1993.

[37] Olof B. Widlund. A Lanczos method for a class of nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis*, 15:801–812, 1978.

[38] Ilyia Zavorin, Howard C. Elman, and Dianne P. O'Leary. Complete stagnation of GMRES. *Linear Algebra and its Applications*, 367:165–183, 2003.

[39] Jan Zítko. Generalization of convergence conditions for a restarted GMRES. *Numerical Linear Algebra with Applications*, 7:117–131, 2000.