Numer. Math. (2008) 109:477–487
DOI 10.1007/s00211-008-0145-y

Numerische Mathematik

New conditions for non-stagnation of minimal residual methods

Valeria Simoncini · Daniel B. Szyld

Received: 17 April 2007 / Revised: 14 September 2007 / Published online: 15 February 2008

Abstract In 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 worst-case bound for the convergence rate of the iterative method, due to Elman. This bound has been extensively used, e.g., when the linear system comes from discretized partial differential equations, to show that the convergence of GMRES is independent of the underlying mesh size. In this paper we introduce more general non-stagnation conditions, 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.

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

V. Simoncini
Dipartimento di Matematica, Università di Bologna,
Piazza di Porta S. Donato,
5, I-40127 Bologna, Italy

V. Simoncini
CIRSA, Ravenna, Italy
e-mail: valeria@dm.unibo.it

D. B. Szyld (✉)
Department of Mathematics, Temple University (038-16),
1805 N. Broad Street, Philadelphia,
Pennsylvania 19122-6094, USA
e-mail: szyld@temple.edu

1 Introduction

Minimal residual Krylov subspace methods, and in particular in the implementation given in GMRES [20], are routinely employed for the solution of large real linear systems of the form $Ax = b$, and especially of those systems arising in the discretization of partial differential equations; see, e.g., [15,19,23]. 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 for one or more iterations, that is $\|r_m\| = \|r_{m-1}\| = \cdots = \|r_{m-k}\|$ holds for some m, and some $k < m$; see, e.g., [7,16,27,28] for examples and discussion of this issue.

Elman [10] 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 [9]. 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 u,

$$c = \min_{u \neq 0} \frac{(u, Au)}{(u, u)} = \min_{u \neq 0} \frac{(u, Hu)}{(u, u)} > 0, \tag{1.1}$$

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 $C = \|A\| = \max_{u \neq 0} \frac{\|Au\|}{\|u\|}$. 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. From (1.1) 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.

The convergence of GMRES is in most cases superlinear (see, e.g., the discussion in [23]), 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 provides important information and is in fact 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 in (1.2), can be bounded independently of the mesh size of the discretization; see, e.g., [25, Sects. 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], [25, Sect. 3.6]. In [5] 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 one can formulate some other conditions for non-stagnation that are applicable to matrices whose symmetric part is not positive definite. This longstanding question is of theoretical and practical importance since, as already mentioned, practitioners have strived to design preconditioners precisely so that (1.2) holds, but have found it difficult to do so for many problems. In addition, more general non-stagnation conditions have been sought after to increase the understanding of GMRES, which usually does not stagnate even when (1.2) fails. Several attempts to look at this question during the last quarter century are known to have been undertaken. Few results have been obtained (see, e.g., [11,28]) but these cannot be easily checked or derived from the properties of the underlying problem.

In this paper, we answer the above 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 some 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 $\mathcal{H}(A)$ be positive definite. In Sect. 3, we present our new conditions together with a few elementary examples of their applicability, while in Sect. 4 we discuss the new results and present additional illustrative examples. We end with a short concluding section.

In the preceeding expressions, as well as in the rest of this paper, the inner product is the Euclidean one $(v, w) = v^T w$, with its associated vector 2-norm. 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. [7,22,24]. Our analysis focuses on real matrices, however the reported results can be appropriately adapted to handle complex matrices.

2 Preliminary and related results

As already mentioned, the bound in (1.2) may be used to ensure convergence in a *restarted* process. Indeed, for non-symmetric matrices, optimal minimal residual methods are usually characterized by large computational costs, which grow superlinearly with the number of iterations, and large memory requirements.

For these reasons, methods such as GMRES are often stopped after a fixed number of iterations, called the restart parameter, 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., [19, Sect. 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 [27].

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., [9], [15, Sect. 3.2], [19, Sect. 6.11.14], [23]. See also [18] for improvements of these bounds in certain cases, and [13] for analogous bounds for non-diagonalizable matrices.

Other convergence bounds using the field of values $\mathcal{F}(A)$ were developed, where for a given $n \times n$ matrix A and denoting with v^* the conjugate transpose of v,

$$\mathcal{F}(A) = \{\omega \in \mathbb{C} \mid \omega = \frac{(v^* A v)}{(v^* v)}, \quad v \in \mathbb{C}^n, v \neq 0\}.$$

All these bounds assume that $0 \notin \mathcal{F}(A)$; see, e.g., [7, Corollary 6.2], [15, Sect. 3.2], [24]. It is precisely for those cases with $0 \in \mathcal{F}(A)$ that we look for new non-stagnation conditions.

We say that a matrix A is positive (negative) definite, if $x^T A x > 0$ for all non-zero real vectors x (if $-A$ is positive definite). In this case, $\mathcal{F}(A) \subset \mathbb{C}^+$ ($\mathcal{F}(A) \subset \mathbb{C}^-$) 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 $x_m - x_0 \in \mathcal{K}_m = \text{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{m-1} r_0\}$, 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 $\|p_m(A) r_0\| < \|r_0\|$. Thus, if for some $q_m \in \mathcal{P}_m$, $q_m(A)$ is positive (or negative) definite, using Elman's result on it, we have $||p_m(A) r_0|| \leq ||q_m(A) r_0|| < ||r_0||$. A more formal statement of this idea was given by Grcar [14], and it is reproduced in the next section.

If A is normal, i.e., if $AA^T = A^T A$, one can characterize $\sigma(A)$ for which $\sigma(A^k) \subset \mathbb{C}^+$, and thus its field of values is also contained in $\mathbb{C}^+$. It can be shown that if

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

then $\sigma(A^k) \subset \mathbb{C}^+$, see, e.g., [28]. For the special case of $k = 2$, this is [15, Exercise 2.8] where it is shown that for A normal, A^2 is positive definite if $|\text{Re}(\lambda)| > |\text{Im}(\lambda)|$ 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 [2,3,17] and to [21], for convergence results that also require $\mathcal{H}(A)$ to be positive definite.

3 The new conditions

We begin by stating the mentioned result of Grcar [14].

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 $x_0 + \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$ contains a vector $\hat{x}$ for which $\|b - A\hat{x}\| \leq \rho\|r_0\|$, 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 $\hat{C} = \|q(A)\|$.

For $q(A) = A$, the hypothesis is that $\mathcal{H}(A)$ is definite, and one recovers (1.2). We mention that this result has not been much used, in part, because the construction of such polynomials q, so that $q(A)$ is definite does not seem to be a simple matter. We do not construct such polynomials. Instead, we give conditions so that $\mathcal{H}(q(A))$ is either positive or negative definite in two specific cases: $q(\eta) = \eta^2$ and $q(\eta) = \eta^4$.

Thus, using Theorem 3.1 we then conclude that, in the case of $q(\eta) = \eta^2$, 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 non-singular, then $\mathcal{H}(A^2)$ is positive definite if and only if*

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

3. *If S is non-singular, 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

$$H^{-T}S^TSH^{-1} = Q\Lambda Q^T, \text{ with} Q^TQ = I, \quad \text{and} \quad \Lambda \geq 0 \text{ diagonal;} \tag{3.4}$$

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

$$\begin{aligned} x^TA^2x &= x^TH^THx - x^TS^TSx = y^Ty - y^TH^{-T}S^TSH^{-1}y \\ &= z^Tz - z^T\Lambda z = z^T(I - \Lambda)z. \end{aligned} \tag{3.5}$$

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 non-zero vectors x, $x^TA^2x = z^T(I - \Lambda)z > 0$ 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, since for any symmetric matrix H and any skew-symmetric matrix S, it holds that $\|HS\| = \|(HS)^T\| = \|S^TH^T\| = \|-SH\| = \|SH\|$.

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 non-singular, 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 non-singular 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* $\min_i |\lambda_i(H)| > \max_j |\lambda_j(S)|$, *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 $\max_i |\lambda_i(H)| < \min_j |\lambda_j(S)|$. Thus, if S is non-singular 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, and GMRES(4) does not stagnate.

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

1. *If $H^2 + S^2$ is non-singular, 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 non-singular, 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 by 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 non-singular, 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

$$x^T q(A)x = x^T(H^2 + S^2 + \alpha H)x = x^T H^2 x - x^T x + \alpha x^T H x$$
$$= z^T \Lambda^2 z - z^T z + \alpha z^T \Lambda z = z^T(\Lambda^2 + \alpha\Lambda - I)z < 0.$$

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 $|\lambda_1| \leq \cdots \leq 1 \leq \cdots \leq |\lambda_n|$, 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 preconditioner; see, e.g., [6,26]. 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, e.g., [8].

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 [12], and assuming that $\mu^2 I - B^T B$ is non-singular, algebraic calculations show that

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

Table 1 Coefficients and corresponding values of $\lambda_{\min}(H)$ and $\|SH^{-1}\|$ for the matrix associated with the operator $L(u)$ for $\eta = 100$

α	β	γ	δ	$\lambda_{\min}(H)$	$\|SH^{-1}\|$
$\exp(-x_1x_2)$	$\exp(x_1x_2)$	-1	-1	-0.04719	0.6194
1	1	$-1/(0.1x_1 + 100x_2)$	0	-0.04775	0.1577
1	1	$1/10(x_1 - x_2)$	0	-0.04772	0.1838
1	1	$1/10(x_1 + x_2)$	0	-0.04772	0.5819
1	1	0.2	0	-0.04781	0.5811

where $\sigma_i, i = 1, \ldots, m$ are the (non-zero) 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 [12].

We conclude with some examples with elliptic operators.

Example 4.1 We consider the class of matrices stemming from the five point centered finite difference discretization of the differential operator

$$L(u) = -(\alpha u_{x_1})_{x_1} - (\beta u_{x_2})_{x_2} + \gamma u_{x_1} + \delta u_{x_2} - \eta u \tag{4.1}$$

on the unit square, with Dirichlet boundary conditions. We use a mesh size $1/41$ giving rise to a matrix of dimension $n = 1600$. In Table 1 we report the minimum eigenvalue of H, and also $\|SH^{-1}\|$, for the case $\eta = 100$, and some choices of the coefficients. In all cases, H is indefinite but the condition (3.2) of Theorem 3.2 holds.

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).

5 Conclusion

We have presented new conditions which guarantee the non-stagnation of minimal residual methods, such as GMRES. Up to now, the only such condition which was relatively easy to check, was the one due to Elman, and which states that $\mathcal{H}(A)$ should be positive definite. We have shown cases where $\mathcal{H}(A)$ is indefinite, but with the new conditions, non-stagnation is assured.

We note that, in the same manner that Elman's results have been applied, 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. Constructing such preconditioners may be a challenge, though.

Acknowledgments We would like to thank Michele Benzi, Anne Greenbaum, Marcus Sarkis, and the three referees, for their comments.

References

1. Antonietti, P.F., Ayuso, B.: Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. ESAIM: Math. Model. Numer. Anal. **41**, 21–54 (2007)
2. Axelsson, O.: A generalized Conjugate Gradient, least squares method. Numer. Math. **51**, 209–227 (1987)
3. Beckermann, B., Goreinov, S.A., Tyrtyshnikov, E.E.: Some remarks on the Elman estimate for GMRES. SIAM J. Matrix Anal. Appl. **27**, 772–778 (2006)
4. Benzi, M., Golub, G.H., Liesen, J.: Numerical solution of saddle point problems. Acta Numer. **14**, 1–137 (2005)
5. Cai, X.-C., Zou, J.: Some observations on the l^2 convergence of the additive Schwarz preconditioned GMRES method. Numer. Linear Algebra Appl. **9**, 379–397 (2002)
6. Concus, P., Golub, G.H.: A generalized Conjugate Gradient method for non-symmetric linear equations. In: Glowinski, R., Lions, J.-L. (eds.) Computing Methods in Applied Science and Engineering, vol. 134 of Lecture Notes in Economics and Mathematical Systems, pp. 56–65. Springer, New York (1976)
7. Eiermann, M., Ernst, O.G.: Geometric aspects in the theory of Krylov subspace methods. Acta Numer. **10**, 251–312 (2001)
8. Eiermann, M., Niethammer, W., Varga, R.S.: Acceleration of relaxation methods for non-Hermitian linear systems. SIAM J. Matrix Anal. Appl. **13**, 979–991 (1992)
9. Eisenstat, S.C., Elman, H.C., Schultz, M.H.: Variational iterative methods for non-symmetric systems of linear equations. SIAM J. Numer. Anal. **20**, 345–357 (1983)
10. Elman, H.C.: Iterative methods for large sparse non-symmetric systems of linear equations. Ph.D thesis, Department of Computer Science, Yale University, New Haven, CT, 1982. Research Report #229
11. Faber, V., Greenbaum, A., Marshal, D.E.: The polynomial numerical hulls of Jordan blocks and related matrices. Linear Algebra Appl. **374**, 231–246 (2003)
12. Fischer, B., Ramage, A., Silvester, D.J., Wathen, A.J.: Minimum residual methods for augmented systems. BIT **38**, 527–543 (1998)
13. Freund, R.W.: Quasi-kernel polynomials and convergence results for quasi-minimal residual iterations. In: Braess, D., Schumaker, L.L. (eds.) Numerical Methods in Approximation Theory, vol. 9, pp. 77–95. Birkhäuser, Basel (1992)
14. Grcar, J.F.: Operator coefficient methods for linear equations. Technical Report SAND89-8691, Sandia National Laboratories (1989). Available at http://seesar.lbl.gov/ccse/Publications/sepp/ocm/SAND89-8691.pdf
15. Greenbaum, A.: Iterative Methods for Solving Linear Systems. SIAM, Philadelphia (1997)
16. Greenbaum, A., Ptàk, V., Strakoš, Z.: Any non-increasing convergence curve is possible for GMRES. SIAM J. Matrix Anal. Appl. **17**, 95–118 (1996)
17. Joubert, W.D.: On the convergence behavior of the restarted GMRES algorithm for solving non-symmetric linear systems. Numer. Linear Algebra Appl. **1**, 427–447 (1994)
18. Liesen, J.: Computable convergence bounds for GMRES. SIAM J. Matrix Anal. Appl. **21**, 882–903 (2000)
19. Saad, Y.: Iterative Methods for Sparse Linear Systems. The PWS Publishing Company, Boston, 1996. 2nd edn. SIAM, Philadelphia (2003)
20. Saad, Y., Schultz, M.H.: GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems. SIAM J. Sci. Stat. Comput. **7**, 856–869 (1986)
21. Sadok, H.: Analysis of the convergence of the minimal and the orthogonal residual methods. Numer. Algorithms **40**, 201–216 (2005)
22. Sarkis, M., Szyld, D.B.: Optimal left and right additive Schwarz preconditioning for minimal residual methods with Euclidean and energy norms. Comput. Methods Appl. Mech. Eng. **196**, 1612–1621 (2007)
23. Simoncini, V., Szyld, D.B.: Recent computational developments in Krylov subspace methods for linear systems. Numer. Linear Algebra Appl. **14**, 1–59 (2007)
24. Starke, G.: Field-of-values analysis of preconditioned iterative methods for non-symmetric elliptic problems. Numer. Math. **78**, 103–117 (1997)

25. Toselli, A., Widlund, O.B.: Domain Decomposition Methods—Algorithms and Theory. Springer, Berlin and Heidelberg (2005)
26. Widlund, O.B.: A Lanczos method for a class of non-symmetric systems of linear equations. SIAM J. Numer. Anal. **15**, 801–812 (1978)
27. Zavorin, I., Elman, H.C., O'Leary, D.P.: Complete stagnation of GMRES. Linear Algebra Appl. **367**, 165–183 (2003)
28. Zítko, J.: Generalization of convergence conditions for a restarted GMRES. Numer. Linear Algebra Appl. **7**, 117–131 (2000)