NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2000; **7**:117–131

Generalization of convergence conditions for a restarted GMRES

Jan Zítko*,†

Charles University of Prague, Department of Numerical Mathematics, Faculty of Mathematics and Physics, Sokolovská 83, 186 75 Praha 8, Czech Republic

SUMMARY

We consider the GMRES(s), i.e. the restarted GMRES with restart s for the solution of linear systems $Ax = b$ with complex coefficient matrices. It is well known that the GMRES(s) applied on a real system is convergent if the symmetric part of the matrix A is positive definite. This paper introduces sufficient conditions implying the convergence of a restarted GMRES for a more general class of non-Hermitian matrices. For real systems these conditions generalize the known result initiated as above. The discussion after the main theorem concentrates on the question of how to find an integer j such that the GMRES(s) converges for all $s \geq j$. Additional properties of GMRES obtained by derivation of the main theorem are presented in the last section. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: restarted GMRES; convergence

1. INTRODUCTION

The GMRES method proposed by Saad and Schultz [1] is a standard procedure for solving the non-Hermitian and non-singular system

$$Ax = b \tag{1.1}$$

where $A \in \mathbb{C}^{n\times n}$ and $x, b \in \mathbb{C}^n$. Let x_0 be an initial approximation and $r_0 = b - Ax_0$ the associated residual. In s steps of GMRES we compute the Krylov subspace $\mathcal{K}_s(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{s-1}r_0\}$ and we then take the new approximate solution x_s in $x_0 + \mathcal{K}_s(A, r_0)$ such that $\|b - Ax_s\|$ is minimal. It is well known that if exact arithmetic is used and $m = \min\{s \mid \dim \mathcal{K}_s(A, r_0) = \dim \mathcal{K}_{s+1}(A, r_0)\}$, then $x_m = A^{-1}b$. Therefore the GMRES will converge in n iterations at most. There is a class of matrices for which n iterations can be achieved. In References [2], [3] it is proved that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are complex numbers and $f_0 \geq f_1 \geq \cdots \geq f_{n-1} > 0$, then there exists a matrix A with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and a right-hand side b such that $\|b\| = f_0$ and $\|r_k\| = f_k$ for $k = 1, 2, \ldots, n-1$. In the worst case,

* Correspondence to: Jan Zítko, Department of Numerical Mathematics, Faculty of Mathematics and Physics, Charles University of Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic
† E-mail: zitko@karlin.mff.cuni.cz)

Contract/grant sponsor: Grant Agency of the Czech Republic; contract grant numbers: 201/98/0528, CEZ J13/98:113200007

Received 11 February 1999
Revised 7 October 1999

we could have

$$\|r_0\| = \|r_1\| = \ldots = \|r_{n-1}\| > 0 \tag{1.2}$$

and for a special choice of eigenvalues it is easy to present real linear systems, the residuals of which behave according to the relation (1.2).

To avoid these difficulties, the precondition strategy is usually considered; see References [4–10]. In brief, preconditioning denotes the replacement of Equation (1.1) by the system $M_1AM_2y = M_1b$, $x = M_2y$, where M_1 and M_2 are non-singular and specially constructed matrices. Usually M_2M_1 approximate in some sense A^{-1}. In what follows we will suppose that the preconditioned system is written in the form (1.1) and we do not introduce any other notation. Another technique for improving the convergence that has been suggested recently is to deflate the matrix explicitly from eigenspaces that hamper convergence; see Reference [11].

Considering the GMRES from a practical viewpoint, the Arnoldi-modified Gram–Schmidt or Householder algorithm is usually used to build an orthonormal basis in the Krylov subspace; see Reference [1]. When the number of iterations k increases, the computational cost and the number of vectors requiring storage increases as $O(k^2n)$ and $O(kn)$, respectively. To avoid these difficulties, we can truncate the orthogonalization, restart the process periodically or restart the truncated version. Now we come to the subject of this paper, i.e. to the restarted algorithm.

We denote by GMRES(s) the restarted version of the GMRES algorithm with restart s. The convergence is discussed in References [4], [8], [1]. The theoretical lower estimate for restart s is constructed (see Reference [1], Corollary 6, p. 867) for diagonalizable matrices and depends on scattering of eigenvalues and the condition number of the transformation matrices. In Elman's thesis [4] the restarted version of GCR, which is mathematically equivalent to GMRES, is studied. For the residual $r_s = b - Ax_s$ we have $r_s = p_s(A)r_0$, where

$$\|p_s(A)r_0\| = \min_{p \in \Pi_s, p(0)=1} \|p(A)r_0\| \tag{1.3}$$

Here Π_s denotes the set of polynomials of degree s. Elman [4] (pp. 40–42) performs the estimate

$$\|r_s\| \le \|p_1(A)^s\| \|r_0\| \le \|p_1(A)\|^s \|r_0\| \tag{1.4}$$

where $p_1 \in \Pi_1$ fulfils (1.3) for $s = 1$. Let us remark that he investigates real linear systems. For the norm $\|p_1(A)\|$ he obtains the estimate

$$\|p_1(A)\| \le \left(1 - (\alpha/\beta)^2\right)^{1/2} \tag{1.5}$$

where $\alpha = \lambda_{\min}\left((A + A^{\mathrm{T}})/2\right)$ and $\beta = \|A\|$ supposing that the symmetric part of A is positive definite. Substituting estimate (1.5) into estimate (1.4) he obtains the estimate

$$\|r_s\| \le \left(1 - (\alpha/\beta)^2\right)^{s/2} \|r_0\| \tag{1.6}$$

This formula is introduced in Reference [1] (p. 866) as well. If the symmetric part of matrix A is positive definite, then $0 < \alpha \le \beta$ (see Reference [4], p. 28) and therefore

$$0 \le \left(1 - (\alpha/\beta)^2\right)^{s/2} < 1 \tag{1.7}$$

for every positive integer s. If $k = js$ and r_k denotes the residual vector after the jth restart, then (1.6) yields $\|r_k\| \leq [(1-(\alpha/\beta)^2)^{s/2}]^j \|r_0\|$ and from Inequality (1.7) it follows that $r_k \to 0$ for $k \to \infty$ and therefore GMRES(s) converges for every positive integer s.

For the computation of a solution of Equation (1.1) Professor Axelsson considers minimizing the quadratic functional $f(x) = \frac{1}{2}(Ax-b)^{\mathrm{T}} M_0 (Ax-b)$ in his paper [12], where M_0 is a symmetric and positive definite matrix. It is assumed that the symmetric part of matrix $M_0 A$ is positive definite. He presents *a generalized truncated conjugate gradient, least square* (GCG-LS) *algorithm* for the computation of a minimizer. He denotes $r_k = Ax_k - b$ and for two following residual vectors proves the inequality

$$\|r_{k+1}\|_{M_0} \leq q\, \|r_k\|_{M_0} \tag{1.8}$$

where $q^2 = \big(1 - 1/\|[(\tilde{A}+\tilde{A}^{\mathrm{T}})/2]^{-1}\|\, \|[(\tilde{A}^{-1}+\tilde{A}^{-T})/2]^{-1}\|\big)$, $\|x\|_{M_0} = (x^{\mathrm{T}} M_0 x)^{1/2}\ \forall x \in \mathbb{R}^n$ and $\tilde{A} = M_0^{-\frac{1}{2}} A M_0$ in this formula. Taking into account that the new approximation x_{k+1} is taken in the linear variety $x_k + \mathrm{span}\{d^1, d^2, \ldots, d^t\}$ where $d^1, d^2, \ldots, d^t$ are called *search directions* in this paper, we see that the above estimate could be transformed and used for restarted GMRES. According to the method described in Reference [12] we easily obtain the relation between $\|r_s\|_{M_0}$ and $\|r_0\|_{M_0}$ in the form $\|r_s\|_{M_0} \leq q^s \|r_0\|_{M_0}$, which is an inequality analogous to (1.6). Therefore GMRES(s) converges if the symmetric part of matrix $M_0 A$ is positive definite. If M_0 is an identity matrix then we obtain

$$\|[(A+A^{\mathrm{T}})/2]^{-1}\| = \frac{1}{\lambda_{\min}\big((A+A^{\mathrm{T}})/2\big)}$$

$$\|[(A^{-1}+A^{-\mathrm{T}})/2]^{-1}\| \leq \frac{\|A\|^2}{\lambda_{\min}\big((A+A^{\mathrm{T}})/2\big)}$$

Hence $q^2 \leq \big(1-(\alpha/\beta)^2\big)$ and therefore Axelsson's estimate is sharper than the estimate (1.6). Moreover, the estimate (1.8) can be used for the proof of convergence of a truncated version of GMRES.

Now the question is to estimate the norm $\|p_s(A)\|$ immediately without any factorization. Such an estimate gives more information about convergence for a wider class of matrices.

In Section 2, the more general estimate of the norm $\|r_s\|$ including Inequality (1.6) is presented together with some necessary basic properties and the convergence theorem. The discussion following the main theorem concentrates on the question of how to find an integer j such that GMRES(s) converges for all $s \geq j$. The additional properties of GMRES obtained by derivation of the main theorem are presented in the last section, Section 4, where various formulas for the coefficients of the polynomial p_s and for $\sin \angle(r_0, r_s)$ are introduced.

Now we present certain notational conventions. The set of all complex $n \times k$ matrices is a complex vector space denoted by $\mathbb{C}^{n\times k}$. The superscript H will be used for conjugating and transposing.

e_i^n	the ith column of the identity matrix $I_n \in \mathbb{C}^{n\times n}$,
e^n	$\sum_{i=1}^n e_i^n$,
S_n	the unit sphere in $\mathbb{C}^n$,
$F(A)$	$\{z^H A z \mid \forall z \in S_n\}$; (the field of values of A),
$\mathbb{C}_+$	the set of all $\lambda \in \mathbb{C}$ such that $\mathrm{Re}\,\lambda > 0$,
$\mathbf{i}$	$\sqrt{-1}$.

We use the Euclidean matrix and vector norm in the whole paper. We suppose that exact arithmetic is used and that all Krylov subspaces have the maximal dimension for all considered restarts.

2. NEW ESTIMATE OF THE NORM $\|r_s\|$

At the beginning of this section we firstly present the GMRES(s) algorithm.

Algorithm 2.1. ***GMRES(s)***
$max \geq 1$ is the maximal number of restarts;
ε is the tolerance for the residual norm;
s: the GMRES algorithm is restarted every s steps;
choose x_0 ;
  convergence := false;
  cycle := 0;
  until *convergence or cycle $\geq max$* **do**
    compute $r_0 = b - Ax_0$;
      if $\|r_0\| < \varepsilon$ *convergence:=true;*
      else
      compute x_s by using GMRES algorithm;
      $x_0 := x_s$; $r_0 := b - Ax_s$;
      cycle := cycle+1;
      endif
  end do
$k := cycle \times s$; $x_k := x_0$; $r_k := r_0$;
output k, x_k, r_k, *convergence, cycle;*
end of **Algorithm 2.1.**

Firstly we investigate one loop in the cycle. We compute the Krylov subspaces $\mathcal{K}_s(A, r_0)$ and $\mathcal{K}_s(A, Ar_0)$ and then take $x_s = x_0 + u_s$, where $u_s \in K_s(A, r_0)$ such that

$$\|r_0 - Au_s\| = \min_{u \in \mathcal{K}_s(A,r_0)} \|r_0 - Au\| \tag{2.1}$$

This means that Au_s is the best approximation in $K_s(A, Ar_0)$ to r_0. Introducing the matrix

$$Q_{r_0,s} = (Ar_0, A^2r_0, \ldots, A^s r_0) \in \mathbb{C}^{n\times s} \tag{2.2}$$

we can rewrite Equation (2.1) in the equivalent form

$$Au_s = Q_{r_0,s}z_s, \qquad \text{where } \|r_0 - Q_{r_0,s}z_s\| = \min_{z\in\mathbb{C}^s} \|r_0 - Q_{r_0,s}z\| \tag{2.3}$$

We assume that the matrix $Q_{r_0,s}$ has maximal rank. The least squares solution gives the vector $z_s = (Q_{r_0,s}^H Q_{r_0,s})^{-1} Q_{r_0,s}^H r_0$ and the new residual is now $r_s = r_0 - P_s r_0$, where $P_s = Q_{r_0,s}(Q_{r_0,s}^H Q_{r_0,s})^{-1} Q_{r_0,s}^H$ is the projector for the space $\mathcal{K}_s(A, Ar_0)$. Therefore $r_s \perp \mathcal{K}_s(A, Ar_0)$ and especially $r_s \perp P_s r_0$. Moreover, we have

$$\|P_s r_0\| = \|r_0\| \cos\angle(r_0, P_s r_0), \qquad \|r_s\| = \|r_0\| \sin\angle(r_0, P_s r_0) \tag{2.4}$$

$$\|r_0\|^2 = \|r_s\|^2 + \|P_s r_0\|^2 = \|r_s\|^2 + d^H_{r_0,s}(Q^H_{r_0,s} Q_{r_0,s})^{-1} d_{r_0,s} \tag{2.5}$$

where

$$d_{r_0,s} = (r_0^H A r_0, r_0^H A^2 r_0, \ldots, r_0^H A^s r_0)^H \tag{2.6}$$

It is easy to see from Equation (2.5) that if vectors $Ar_0, A^2r_0, \ldots, A^s r_0$ are linearly independent then $r_s = r_0$ if and only if $r_0 \perp A^i r_0 \,\forall i = 1, 2, \ldots, s$. Therefore GMRES makes some progress in the first s iterations if $d_{r_0,s} \neq 0$. But the residual vector changes after every s iteration and only condition $d_{r_0,s} \neq 0$ does not guarantee the convergence. Let us turn our attention back to formula (2.5). If we put $y = r_0/\|r_0\|$ and analogously to (2.2) and (2.6) we denote $Q_{y,s} = (Ay, A^2y, \ldots, A^s y)$ and $d_{y,s} = (y^H Ay, y^H A^2 y, \ldots, y^H A^s y)^H$, then

$$\|r_s\|^2/\|r_0\|^2 = 1 - d^H_{y,s}\left(Q^H_{y,s} Q_{y,s}\right)^{-1} d_{y,s} \stackrel{\text{def}}{=} q_{y,s} \tag{2.7}$$

Denoting $M_{y,s} = \left(Q^H_{y,s} Q_{y,s}\right)$ we obtain by easy manipulations the following estimate for $q_{y,s} = \|r_s\|^2/\|r_0\|^2$:

$$\begin{aligned}
q_{y,s} &= 1 - d^H_{y,s} M^{-1}_{y,s} d_{y,s} \leq 1 - \|d_{y,s}\|^2 \lambda_{\min}\left(M^{-1}_{y,s}\right) \\
&= 1 - (|y^H Ay|^2 + |y^H A^2 y|^2 + \ldots + |y^H A^s y|^2)/\lambda_{max}\left(M_{y,s}\right) \\
&\leq 1 - \left(\sum_{j=1}^{s} |y^H A^j y|^2\right) \Big/ \left(y^H A^H A y + \ldots + y^H (A^s)^H A^s y\right) \\
&\leq 1 - \min_{z \in S_n} \left(\sum_{j=1}^{s} |z^H A^j z|^2\right) \Big/ \left(\lambda_{max}(A^H A) + \ldots + \lambda_{max}((A^s)^H A^s)\right) \\
&= 1 - \min_{z \in S_n} \left(\sum_{j=1}^{s} |z^H A^j z|^2\right) \Big/ \sum_{j=1}^{s} \|A^j\|^2
\end{aligned} \tag{2.8}$$

From Inequality (2.8) we have

$$q_{y,s} \leq 1 - \min_{z \in S_n} \left(\sum_{j=1}^{s} |z^H A^j z|^2\right) \Big/ \sum_{j=1}^{s} \|A^j\|^2 \stackrel{\text{def}}{=} q_s \tag{2.9}$$

Remark 1

Defining analogously the quotient $q_{y,m} = \|r_m\|^2/\|r_0\|^2$ for all positive integers $m \leq s$ then it is easy to see that the estimate (2.9) is valid if we substitute m instead of s in this formula.

Remark 2

If the matrix A, right-hand side b and the initial approximation are real, it suffices to consider the minimum only in the set $S_n \cap \mathbb{R}^n$.

If i denotes the number of restarts and if we put $k = i \times s$, then according to Algorithm 2.1 we obtain

$$\|r_k\| \leq q_s^{i/2} \|r_0\| \tag{2.10}$$

and therefore GMRES(s) is convergent if $q_s < 1$. The inequalities (2.9) and (2.10) imply the following propositions.

Theorem 2.1. *Let $s \in \{1, \ldots, n-1\}$. Let at least one of these conditions be valid.*

(i) *There exists $j \in \{1, \ldots, s\}$ such that $\min\limits_{z \in S_n} |z^H A^j z| > 0$.*

(ii) *There exists $j \in \{1, \ldots, s\}$ such that for every $z \in S_n$ for which $z^H A^j z = 0$, an integer $l = l(z) \in \{1, \ldots, s\}$ exists such that $|z^H A^l z| > 0$.*

(iii) *The system of equations*

$$z^H A^j z = 0 \qquad \text{for } j = 1, \ldots, s \tag{2.11}$$

does not have any solution in S_n.

Then GMRES(s) *is convergent. In case (i), moreover,* GMRES(m) *is convergent for all $m \geq j$.*

Proof

According to Equation (2.7) and Inequality (2.9), the number $q_s \in [0, 1]$. It suffices to prove that conditions (i) or (ii) eliminate the case $q_s = 1$ because it is easy to see that the conditions (ii) and (iii) are equivalent. The set S_n is compact and therefore there exists a point $z_0 \in S_n$ such that the continuous function $\sum_{j=1}^{s} |z^H A^j z|^2$ has its minimum in z_0, i.e. $z_0 = \arg_{z \in S_n} \min \left(\sum_{j=1}^{s} |z^H A^j z|^2 \right)$. Since $|z_0^H A^j z_0| > 0$ in the case (i), we have then $q_s < 1$. Moreover $q_m < 1 \; \forall m \geq j$. In case (ii), if $|z_0^H A^j z_0| = 0$ then an $l \in \{1, \ldots, s\}$ exists such that $|z_0^H A^l z_0| > 0$ and therefore $q_s < 1$. ■

Remark 3

If matrix A, right-hand side b and the initial approximation x_0 are real then (iii) can be formulated in the form

(iii′) The system of equations (2.11) does not have any solution in $S_n \cap \mathbb{R}^n$.

Let $A^j = H_j + \mathbf{i} K_j$, where H_j and $\mathbf{i} K_j$ are the Hermitian and skew-Hermitian parts of A^j, respectively. According to the decomposition of the matrix A^j, we have for all z

$$|z^H A^j z|^2 = (z^H H_j z)^2 + (z^H K_j z)^2 \tag{2.12}$$

From Equation (2.12) we can see directly that $\min_{z \in S_n} |z^H A^j z| > 0$ if and only if the system

$$z^H H_j z = 0, \qquad z^H K_j z = 0 \tag{2.13}$$

does not have any solution in the unit sphere S_n. For a special case we can formulate the following theorem, which generalizes the known result [4], [1] for real linear systems.

Theorem 2.2. *Let $j \in \{1, \ldots, s\}$ exist such that at least one of matrices H_j or K_j is positive definite. Then* GMRES(m) *is convergent for all $m \geq j$.*

Proof
If matrix H_j (or K_j) is positive definite then the system (2.13) does not have any solution in the unit sphere S_n and therefore $\min_{z \in S_n} |z^H A^j z| > 0$. The rest is obvious. ■

In the next section we find for special cases the smallest restart $s \in \{1, \ldots, n-1\}$ for which GMRES(s) is convergent.

3. DISCUSSION OF CONVERGENCE CONDITIONS

For real linear systems the statement in Theorem 2.2 for $j = 1$ gives the well-known proposition described already in Elman's thesis [4].

Manteuffel [13] investigates the five-point discretization of the partial differential equation

$$\begin{aligned} -\varepsilon \triangle \mathbf{x}(r,t) + \sigma_1 \mathbf{x}_r(r,t) + \sigma_2 \mathbf{x}_t(r,t) &= \mathbf{c}(r,t) \qquad \text{on} \quad \Omega = (0,1) \times (0,1) \\ \mathbf{x}(r,t) &= \mathbf{g}(r,t) \qquad \text{on} \quad \partial\Omega \end{aligned} \tag{3.1}$$

where σ_1, σ_2 and $\varepsilon > 0$ are real numbers. If the first derivatives are approximated by centred differences (see References [14], [7]) then the system of linear algebraic equations is obtained with matrix $A = M + \mathbf{i}N$, where $M = (A + A^{\mathrm{T}})/2$ is positive definite and $N = (A - A^{\mathrm{T}})/2\mathbf{i}$ is Hermitian. Hence the field of values $F(A)$ lies in the half-plane $\mathbb{C}_+$ and GMRES(s) is convergent for every s according to Theorem 2.1.

Now we investigate normal matrices. If A is normal, then a unitary matrix U and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ exist such that $A = U^H \Lambda U$. Apparently A^k is normal and the numbers $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ are eigenvalues of A^k for every positive integer k. Let us define the sets $\mathcal{P}_j$ for $j \in \{1, 2, \ldots\}$:

$$\mathcal{P}_j = \left\{ \lambda \in \mathbb{C} \mid \lambda = |\lambda| \mathrm{e}^{\mathbf{i}\varphi_\lambda} \quad \text{where } \varphi_\lambda \in \bigcup_{k=0}^{j-1} [(-\pi + 4\pi k)/2j, (\pi + 4\pi k)/2j] \right\}$$

If $z \in \mathcal{P}_j$ then $z^j \in \mathbb{C}_+$. Therefore if all eigenvalues of A lie in $\mathcal{P}_j$ then the eigenvalues of A^j lie in $\mathbb{C}_+$. Hence $F(A^j)$ does not contain zero and GMRES(s) converges if $s \geq j$. The following more general theorem generalizes this result.

Theorem 3.1. *Let A be a normal matrix, $\omega \in [0, 2\pi)$ and $s \in \{1, 2, \ldots, n-1\}$. Let at least one of the following conditions be valid.*

(i) *There exists $j \in \{1, \ldots, s\}$ such that all eigenvalues of matrix A lie in the set $\mathrm{e}^{\mathbf{i}\omega}\mathcal{P}_j$.*
(ii) *The rectangular system*

$$\left.\begin{aligned} \alpha_1 \quad + \alpha_2 \quad + \cdots + \alpha_n \quad &= 1 \\ \alpha_1\lambda_1 + \alpha_2\lambda_2 + \cdots + \alpha_n\lambda_n &= 0 \\ \alpha_1\lambda_1^2 + \alpha_2\lambda_2^2 + \cdots + \alpha_n\lambda_n^2 &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ \alpha_1\lambda_1^s + \alpha_2\lambda_2^s + \cdots + \alpha_n\lambda_n^s &= 0 \\ \alpha_i \geq 0 \;\; \forall i & \end{aligned}\right\} \tag{3.2}$$

does not have any solution.

Then GMRES(s) *is convergent. In case (i), moreover,* GMRES(m) *is convergent for all* $m \geq j$.

Proof

(i) If the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of matrix A lie in the set $e^{\mathbf{i}\omega}\mathcal{P}_j$ then the eigenvalues of A^j lie in $e^{\mathbf{i}\omega j}\mathcal{P}_1 = e^{\mathbf{i}\nu}\mathbb{C}_+$ where we have put $\nu = \omega j$. But A^j is normal and therefore $F(A^j)$ coincides with the convex hull of the eigenvalues of A^j (see Reference [15], p. 81 or Reference [13], p. 190). But the convex compact set $F(A^j)$ lies in the half-plane $e^{\mathbf{i}\nu}\mathbb{C}_+$ and therefore the distance $F(A)$ from zero is positive, i.e. $\min_{z \in S_n} |z^H A^j z| > 0$ and according to Theorem 2.1(i) GMRES(m) is convergent for all $m \geq j$.

(ii) Let u_i denote the ith column of the transformation matrix U. If $z = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n \in S_n$ then $\sum_{i=1}^n |\beta_i|^2 = 1$, and if we put $\alpha_i = |\beta_i|^2 \geq 0\ \forall i$ then $z^H A^j z = \alpha_1 \lambda_1^j + \alpha_2 \lambda_2^j + \cdots + \alpha_n \lambda_n^j$ and we see that system (3.2) is a rewritten system (2.11) for normal matrices and we can use Theorem 2.1 (iii). ∎

Let us give an example. In paper [16] conjugate gradient type methods are presented for the solution of linear systems with complex coefficient matrices of type $A = T + \mathbf{i}\sigma I$ where T is Hermitian and σ a real scalar. The eigenvalues of A lie in the half-plane $e^{\mathbf{i}\pi/2}\mathbb{C}_+$ if $\sigma > 0$ and in the half-plane $e^{-\mathbf{i}\pi/2}\mathbb{C}_+$ if $\sigma < 0$. Therefore GMRES(s) is convergent for all positive integers s according to (i) of the last theorem.

Let us remark that we would obtain the same results if we applied Theorem 2.2. We have $K_1 = \sigma I$. This matrix is positive (negative) definite if $\sigma > 0$ ($\sigma < 0$).

Now we investigate diagonalizable matrices. Let A be diagonalizable, i.e. $A = X\Lambda X^{-1}$ where X is non-singular and

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{3.3}$$

is a diagonal matrix. If $s \in \{1, 2, \ldots, n-1\}$ and a complex number $\gamma \neq 0$ exists such that $\Lambda^s = \gamma \times I_n$, then for every $z \in S_n$ it is $z^H A^s z = z^H X \Lambda^s X^{-1} z = \gamma$ and therefore $0 \notin F(A^s) = \gamma$ which proves that GMRES(s) is convergent according to Theorem 2.1.We expect that the convergence of restarted GMRES will be preserved if there exists an eigenvalue of matrix A(without any loss of generality let λ_1 be this eigenvalue) such that λ_k^s for $k = 2, \ldots, n$ will be situated in some small neighbourhood of λ_1^s.

Theorem 3.2. *Let A be diagonalizable, $A = X\Lambda X^{-1}$, where X is regular and Λ is a diagonal matrix given by* (3.3). *Denote $\kappa(X) = \|X\|\|X^{-1}\|$. If $s \in \{1, 2, \ldots, n-1\}$ exists such that*

$$|\lambda_1|^s > \kappa(X) \max_{k \in \{2,\ldots,n\}} |\lambda_k^s - \lambda_1^s| \tag{3.4}$$

then GMRES(s) *is convergent. Moreover,* GMRES(m) *is convergent for all* $m > s$.

Proof

Let us split $\Lambda^s = \lambda_1^s I_n + M$ where $M = \mathrm{diag}(0, \lambda_2^s - \lambda_1^s, \ldots, \lambda_n^s - \lambda_1^s)$. Let $z \in S_n$. Then $z^H A^s z = \lambda_1^s + z^H X M X^{-1} z$. Since

$$|z^H X M X^{-1} z| \leq \kappa(X) \max_{k \in \{2,\ldots,n\}} |\lambda_k^s - \lambda_1^s|$$

it follows from Inequality (3.4) that $F(A^s)$ lies in a closed circle centred at $|\lambda_1|^s$ and having radius $\kappa(X)\max_{k\in\{2,\dots,n\}}|\lambda_k^s-\lambda_1^s|$. But this circle does not contain zero, i.e., $0 \notin F(A^s)$. The rest is obvious. ∎

It is easy to see that Theorem 3.2 can be generalized for non-diagonalizable matrices.

Corollary 3.1. *Let* $s \in \{1,\dots,n-1\}$ *exist such that* $A^s = X\Omega X^{-1}$ *where* X *is nonsingular,* $\Omega = \gamma I_n + B$ *and* $|\gamma| > \kappa(X)\|B\|$. *Then* GMRES($s$) *is convergent. Moreover,* GMRES(m) *is convergent for all* $m > s$.

The proof is an easy modification of the previous one.

Now we return to Theorem 3.2 and try to construct the class of matrices fulfilling the condition (3.4). We shortly denote $\kappa = \kappa(X)$. Let $\lambda_1 > 0$ and let the following conditions for eigenvalues of a diagonalizable matrix A be valid:

- λ_1 is an eigenvalue of A,
- if λ is an eigenvalue of A then

$$\lambda = (\lambda_1+\delta)\exp\{[(2\pi k/s)+\vartheta]\mathbf{i}\} \quad \text{for some} \quad k\in\{0,1,\dots,s-1\} \tag{3.5}$$

Our task is to find intervals for the real numbers δ and ϑ such that the inequality

$$|\lambda_1|^s > \kappa|\lambda^s - \lambda_1^s| \tag{3.6}$$

holds. We conclude from (3.6) that it is reasonable to assume that

$$|\delta/\lambda_1| < 1, \quad \text{and} \quad |s\vartheta| < \pi/2 \tag{3.7}$$

Now we discuss Inequality (3.6). By direct verification, it is found that

$$|\lambda^s - \lambda_1^s|^2 = \lambda_1^{2s}[1 + (1+\delta/\lambda_1)^{2s} - 2(1+\delta/\lambda_1)^s\cos(s\vartheta)] \tag{3.8}$$

Substituting (3.8) into (3.6) we obtain that

$$\frac{1}{\kappa^2} > 1 + (1+\delta/\lambda_1)^{2s} - 2(1+\delta/\lambda_1)^s\cos(s\vartheta) \tag{3.9}$$

Inequality (3.9) is valid for $\delta = \vartheta = 0$. However, the right-hand side of (3.9) is a continuous function of the variables δ and ϑ in $\mathbb{R}^2$. Hence there exists a neighbourhood $\mathcal{U}_s = (-\delta_s,\delta_s)\times(-\vartheta_s,\vartheta_s) \subset \mathbb{R}^2$of zero such that Inequality (3.9) is valid for all $(\delta,\vartheta)\in\mathcal{U}_s$. It is $\delta_s > 0$ and $\vartheta_s > 0$. According to (3.7) it is

$$1 + (1+\delta/\lambda_1)^{2s} - 2(1+\delta/\lambda_1)^s\cos(s\vartheta) \ge (1-(1+\delta/\lambda_1)^s)^2 \tag{3.10}$$

Hence using Inequalities (3.9) and (3.10) we obtain an upper bound for δ_s from the inequality

$$\frac{1}{\kappa} > |(1-(1+\delta/\lambda_1)^s)| \tag{3.11}$$

If we approximate $(1 \pm 1/\kappa)^{1/s} \sim 1 \pm 1/\kappa s$ than we obtain the estimate

$$\delta_s < \lambda_1/\kappa s \tag{3.12}$$

From (3.9) we have

$$\cos(s\vartheta) > \frac{1 - 1/\kappa^2}{2(1+\delta/\lambda_1)^s} + \frac{1}{2}(1+\delta/\lambda_1)^s \tag{3.13}$$

The function

$$f(x) = \frac{1 - 1/\kappa^2}{2x} + \frac{1}{2}x$$

has its minimum on the interval $[0, \infty)$ in the point $\sqrt{1 - 1/\kappa^2}$ and this minimum is equal to $\sqrt{1 - 1/\kappa^2}$. Therefore, we obtain an upper bound for ϑ_s from the inequality

$$\cos(s\vartheta_s) > \sqrt{1 - 1/\kappa^2} \tag{3.14}$$

From (3.12) and (3.14) we can see that the upper bound for δ_s and ϑ_s decreases if κ or s increases. Moreover, both bounds converge to 0 provided that $\kappa \to \infty$ or $s \to \infty$. The spectrum of another matrix fulfilling (3.6) can be obtained by rotation of the complex plane. If we substitute λ from (3.5) instead of λ_k in (3.4) we obtain (3.6) from (3.4). We can see that for given s and κ the class of matrices satisfying (3.4) is small. The assumption $0 \notin F(A^s)$ on which the proof of Theorem 3.1 is based is strong. However, using the weaker assumption (*iii*) from Theorem 2.1 we get the convergence even if $0 \in F(A^j)$ for all $j \leq s$. The next example validates this hypothesis.

For a general matrix, the field of values can be much bigger than the convex hull of the eigenvalues and it is very questionable to require $0 \notin F(A)$. But Theorem 2.1 (*ii*) or (*iii*) can still be used here. We consider the following academic example. If U is any unitary matrix then $F(U^H A U) = F(A)$ and consequently there is no restriction in supposing A to be upper triangular. Let us consider the matrix

$$A = \begin{pmatrix} 2 & \alpha & 0 \\ 0 & -4 & \alpha \\ 0 & 0 & 8 \end{pmatrix}$$

for $\alpha \in \mathbb{R}$. It is easy to find that $H_1 = (A + A^H)/2$ and $K_1 = (A - A^H)/2\mathbf{i}$ are indefinite for all α. The following statements can be obtained after easy calculation. Matrix $H_2 = (A^2 + (A^2)^H)/2$ is indefinite for $|\alpha| > \sqrt{\left(-20 + \sqrt{6544}\right)/3} \stackrel{\text{def}}{=} \tilde{\alpha} \doteq 4.5054$, positive semidefinite for $|\alpha| = \tilde{\alpha}$, and positive definite for $|\alpha| < \tilde{\alpha}$. Matrix $K_2 = (A^2 - (A^2)^H)/2\mathbf{i}$ is indefinite for all α. Hence GMRES(2) is convergent for $|\alpha| < \tilde{\alpha}$ according to Theorem 2.2. Let us put $\alpha = 5 > \tilde{\alpha}$. For this matrix A the symmetric part and antisymmetric part multiplied by $\mathbf{i}$ are indefinite for both matrices A and A^2. Hence $0 \in F(A^j)$ for $j = 1, 2$, i.e. $z^1 \in S_3$ and $z^2 \in S_3$ exist such that $(z^j)^H A^j z^j = 0$. Therefore Theorem 2.2 and Theorem 2.1(i) cannot be used. But (iii') in Remark 3 gives the convergence of GMRES(2).

It is $z^H A z = 0$ for all vectors $z = (z_1, z_2, z_3)^{\mathrm{T}} \in S_3 \cap \mathbb{R}^3$ satisfying the relation

$$2z_1^2 - 4z_2^2 + 8z_3^2 + 5z_1 z_2 + 5z_2 z_3 = 0 \tag{3.15}$$

We have

$$z^H A^2 z = 4z_1^2 + 16z_2^2 + 64z_3^2 - 10z_1z_2 + 25z_1z_3 + 20z_2z_3$$

Substituting the relation (3.15) in the equation $z^H A^2 z = 0$we obtain the equation

$$8 + 72z_3^2 + 30z_2z_3 + 25z_1z_3 = 0$$

The function

$$f(z_1, z_2, z_3) = 8 + 72z_3^2 + 30z_2z_3 + 25z_1z_3$$

has its minimum on the compact set $S_3 \cap \mathbb{R}^3$ at the points

$$\begin{aligned} y_1 &= (-0.620522, -0.744626, 0.245937) \\ y_2 &= (0.620522, 0.744626, -0.245937) \end{aligned}$$

and $f(y_1) = f(y_2) = 3.045752$. Therefore the system

$$z^H Az = 0, \qquad z^H A^2 z = 0$$

does not have any solution in $S_3 \cap \mathbb{R}^3$. According to statement (iii'), GMRES(2) is convergent. This example shows that the restarted GMRES also converges in the case that the Hermitian part and skew-Hermitian part multiplied by **i** of all powers of a matrix A are indefinite.

For arbitrary matrices it is not known how to choose s, in general. Let us consider a discrete approximation of the partial differential equation

$$-(\mathbf{P}\mathbf{x}_r)_r - (\mathbf{Q}\mathbf{x}_t)_t + \mathbf{R}_1 x_s + (\mathbf{R}_2 x)_r + \mathbf{S}_1 x_t + (\mathbf{S}_2 x)_t + \mathbf{T}\mathbf{x} = \mathbf{F} \tag{3.16}$$

defined in an open, connected and bounded set Ω in the plane with the Dirichlet condition on Γ, the boundary of Ω. Here $\mathbf{x}$=$\mathbf{x}(r, t)$ and we assume that the given functions **P, Q,** $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{S}_1$, $\mathbf{S}_2$,**T**,**F** of two variables r, t, are continuous in $\overline{\Omega}$, and satisfy conditions $\mathbf{P}(r, t) > 0$, $\mathbf{Q}(r, t) > 0$ and $\mathbf{T}(r, t) > 0$. For the sake of simplicity we assume that region Ω is square $(0, 1) \times (0, 1)$. For the coefficients we have taken the functions from the examples in References [4, 7]. We discretize the partial differential equation on a uniform grid with the mesh size $h = 1/(l + 1)$. The linear system $Ax = b$ is of order $n = l^2$. The behaviour of $\log_{10} \|r_k\|/\|r_0\|$ for various values of the restart was studied and, moreover, the computer time needed for the $\|r_k\|/\|r_0\|$ to be less than 10^{-10} was indicated . The time-comparing shows that the best restart lies in the interval [20, 40] for systems having about ten to twenty thousand unknowns. For matrices obtained by discretization of Equation (3.16) we have verified that $z^{\mathrm{T}} A^j z \neq 0$ for $j = 1, 2, \ldots, 20$ and for a lot of randomly chosen vectors $z \in S_n$. Therefore we have put $s = 20$ and have hoped that the system

$$z^{\mathrm{T}} A^j z = 0 \qquad \text{for} \quad j = 1, 2, \ldots, s$$

has no solution on S_n for all systems obtained by discretization of Equation (3.16). Since the time for calculation rapidly grow as s increase we have, moreover, given an upper bound for s. The interval [20, 40] for s is recommended for systems without preconditioning or if a diagonal matrix is used as preconditioner. The preconditioning strategy involving incomplete Cholesky factorization will be discussed later.

Finally, we consider the system obtained by discretization of Equation (3.1) with Dirichlet boundary condition, where $\sigma_1 = 2(1 + x^2)$, $\sigma_2 = 100$ and $c(r, t) = 0$. This system behaves differently from systems in References [4] or [7]. For $\varepsilon \in [10^{-2}, 10^2]$, there were no problems to obtain $\|r_k\|/\|r_k\| < 10^{-10}$ for $k < 1000$. To obtain the same precision for $k < 1000$ we have used $s = 60$ for $\varepsilon = 10^{-3}$and $s = 100$ for $\varepsilon = 10^{-4}$. The time of calculation grew astronomically, especially for small ε. But no problems occur if we have used the standard incomplete LU decomposition strategy for preconditioning. Let us return to our theory.

Let us suppose that the matrix of the preconditioned system has the form $A = I + C$, where $\|C\| < 1$. Evidently, the equation $z^{\mathrm{T}}Az = 0$ with the constrain $z^{\mathrm{T}}z = 1$ has no solution and GMRES(1) is convergent. Moreover the quotient $(\min_{z \in S_n} |z^{\mathrm{T}}Az|)/\|A\|$ can be close to 1 if $\|C\|$ is small and consequently for the quotient from (2.9) we have $q_1 \ll 1$. Generally we believe, on the basis of the above considerations, that by using an incomplete LU decomposition strategy for preconditioning, the system (2.11) does not have any solution for small s and that q_s is not close to 1. For $\varepsilon = 1$, GMRES(1) preconditioned by an incomplete LU decomposition strategy was about five times faster than GMRES(20) without preconditioning. Hence small s is recommended in preconditioned cases because, on the other hand, preconditioning means a work over. For example, the quotient $\|r_k\|/\|r_0\|$ converges in the preconditioned case incomparably faster to zero than the same one without preconditioning, but the time is only twice as good.

4. OBSERVATIONS AND REMARKS

This section extends the classical knowledge about moment matrices ([15]). The result of this section is concentrated in Theorem 4.1. The coefficients of the polynomial p_s, $\sin \angle(r_0, r_s)$ and $\|r_s\|$ are expressed by using the moment matrices defined by (4.2). Let $p_s(z) = \alpha_s z^s + \alpha_{s-1}z^{s-1} + \cdots + \alpha_0$ be the polynomial defined by (1.3). In this section we write e_i instead of e_i^{s+1} and e instead of e^{s+1} for brevity. We have shown that

$$r_s = p_s(A)r_0 = r_0 - Q_{r_0,s}(Q_{r_0,s}^H Q_{r_0,s})^{-1} d_{r_0,s} \tag{4.1}$$

where $Q_{r_0,s}$ is given by (2.2) and $d_{r_0,s}$ by (2.6). Let us introduce the moment matrices

$$\left.\begin{aligned} M_{Ar_0,s} &= Q_{r_0,s}^H Q_{r_0,s}, \qquad M_{Ay,s} = Q_{y,s}^H Q_{y,s} \\ M_{r_0,s+1} &= (r_0, Ar_0, \dots, A^s r_0)^H (r_0, Ar_0, \dots, A^s r_0) \end{aligned}\right\} \tag{4.2}$$

The matrix $M_{y,s+1}$ is defined analogously. The relation (4.1) implies

$$r_s = (r_0, Ar_0, A^2 r_0, \dots, A^s r_0) \begin{pmatrix} 1 \\ -M_{Ar_0,s}^{-1} d_{r_0,s} \end{pmatrix} \tag{4.3}$$

Evidently $\alpha_0 = 1$. Putting $a = (1, \alpha_1, \alpha_2, \dots, \alpha_s)^{\mathrm{T}}$ we obtain from (4.1) and (4.3) that

$$a = \begin{pmatrix} 1 \\ -M_{Ar_0,s}^{-1} d_{r_0,s} \end{pmatrix} \tag{4.4}$$

For $\sin\angle(r_0, r_s)$ we have

$$\sin\angle(r_0, r_s) = (1 - d_{r_0,s}^H M_{Ar_0,s}^{-1} d_{r_0,s}/\|r_0\|^2)^{1/2} = (1 - d_{y,s}^H M_{Ay,s}^{-1} d_{y,s})^{1/2} \tag{4.5}$$

where $y = r_0/\|r_0\|$.

For vector a and $\sin\angle(r_0, r_s)$ other formulas can be achieved immediately using the definition of r_s and (2.1).

Theorem 4.1. *Let the vectors $r_0, Ar_0, A^2r_0, \ldots, A^s r_0$ be linearly independent. Then the relations*

$$a = (e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} e_1)^{-1} M_{r_0,s+1}^{-1} e_1 \tag{4.6}$$

$$\sin\angle(r_0, r_s) = (e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} e_1)^{-1/2}/\|r_0\| \tag{4.7}$$

$$\|r_s\| = (e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} e_1)^{-1/2} \tag{4.8}$$

hold.

Proof
We have

$$\|r_s\|^2 = \|p_s(A)r_0\|^2 = r_0^H p_s(A)^H p_s(A) r_0 = a^H M_{r_0,s+1} a \stackrel{\text{def}}{=} F(a)$$

and $e_1^{\mathrm{T}} a = 1$. According to (1.3)

$$a = \arg \min_{\substack{v\in\mathbb{C}^{s+1}\\ e_1^{\mathrm{T}} v = 1}} v^H M_{r_0,s+1} v$$

Since $M_{r_0,s+1}$ is positive definite, a number $\beta \in \mathbb{R}$ exists such that the vector a is the only solution of the equations

$$\begin{aligned} F'(v) + \beta e_1^{\mathrm{T}} &= 0 \\ e_1^{\mathrm{T}} v &= 1 \end{aligned}$$

where $F'(v)$ denotes the Gâteaux derivative of the functional F at v. It is easy to see that the vector a is given by (4.6). From (4.1) we obtain

$$\|r_s\|^2 = \|p_s(A)r_0/\|r_0\|\|^2 \|r_0\|^2$$

Hence

$$\begin{aligned} \sin\angle(r_0, r_s) &= \left\| p_s(A) r_0/\|r_0\| \right\| = \frac{1}{\|r_0\|}(a^H M_{r_0,s+1} a)^{1/2} \\ &= \left(e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} (e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} e_1)^{-1} M_{r_0,s+1} (e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} e_1)^{-1} M_{r_0,s+1}^{-1} e_1 \right)^{1/2} / \|r_0\| \\ &= (e_1^{\mathrm{T}} M_{r_0,s+1}^{-1} e_1)^{-1/2}/\|r_0\| \end{aligned}$$

Relation (4.8) is evident. ■

Remark 1
If the moment matrix is constructed for the vector $y = r_0/\|r_0\|$, then

$$\sin\angle(r_0, r_s) = (e_1^{\mathrm{T}} M_{y,s+1}^{-1} e_1)^{-1/2}$$

Remark 2
If the moment matrix is constructed, moreover, for the matrix $I - A$, we obtain

$$\sin\angle(r_0, r_s) = (e^{\mathrm{T}} M_{y,s+1}^{-1} e)^{-1/2}$$

ACKNOWLEDGEMENTS

This paper was supported by the Grant Agency of the Czech Republic under Grant No. 201/98/0528 and under Grant CEZ J13/98:113200007. I thank the referees, whose reports lead to corrections and improvements in this paper.

REFERENCES

1. Saad Y, Schultz MH. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 1986; **7**:856–869.
2. Greenbaum A, Strakoš Z. Matrices that generate the same Krylov residual spaces. In *Advances in Iterative Methods*, Golub G, Greenbaum A, Luskin M (eds). Springer–Verlag: New York, 1994.
3. Greenbaum A, Pták V, Strakoš Z. Any nonincreasing convergence curve is possible for GMRES. *SIAM Journal on Matrix Analysis and Applications* 1996; **17**:465–469.
4. Elman HC. Iterative methods for large sparse nonsymmetric systems of linear equations. PhD thesis, Computer Science Department, Yale University, New Haven, CT, 1982.
5. Freund RW, Golub GH, Nachtigal NM. Iterative solution of linear systems. *Acta Numerica* 1991; 57–100.
6. Golub GH, Van Loan ChF. *Matrix Computation*. The John Hopkins University Press: Baltimore, MD, 1984.
7. Kincaid DR, Hayes LJ. Iterative methods for large linear systems. Papers from a conference held 1988 at the Center for Numerical Analysis of the University of Texas at Austin, Academic Press, 1989.
8. Saad Y. *Iterative Methods for Sparse Linear Systems*. PWS Publishing, 1996.
9. Weiss R. *Parameter–Free Iterative Linear Solver*. Academie Verlag: Berlin, 1996.
10. Zítko J. Improving the convergence of GMRES using preconditioning and pre-iterations. In *Proceedings of 'Prague Mathematical Conference 1996'*. 1996; 377–382.
11. Chapman A, Saad Y. Deflated and augmented Krylov subspace techniques. *Numerical Linear Algebra with Applications* 1997; **4**:43–66.
12. Axelsson O. A generalized conjugate gradient, least square method. *Numerische Mathematik* 1987; **51**:209–227.
13. Manteuffel TA. Adaptive procedure for estimating parameters for the nonsymmetric Tschebychev iteration. *Numerische Mathematik* 1978; **31**:183–208.
14. Elman HC, Golub GH. Iterative methods for cyclically reduced non-self-adjoint linear systems. *Mathematics of Computations* 1990; **54**:671–700.
15. Householder AS. *The Theory of Matrices in Numerical Analysis*. Blaisdell: New York, 1964.
16. Freund RW. On conjugate gradient type methods and polynomial preconditioners for class of complex non–Hermitian matrices. *Numerische Mathematik* 1990; **57**:285–312.
17. Fadeev DK, Fadeeva VN. *Computational Methods of Linear Algebra*. Freeman: San Francisco, 1963.
18. Fiedler M. *Special Matrices and their Applications in Numerical Mathematics*. Martinus Nijhoff: Dordrecht, The Netherlands, 1986.
19. Gutknecht MH. Lanczos-type solvers for nonsymmetric linear system of equations. Technical Report TR–97–04, 1997.

20. Hageman LA, Young DM. *Applied Iterative Method*. Academic Press: New York, 1981.
21. Hackbusch W. *Iterative Solution of Large Linear Systems of Equations*. Springer–Verlag: New York, 1994.
22. Saad Y. Krylov subspace methods for solving large unsymmetric linear systems. *Mathematics of Computations* 1981; **37**:105–126.
23. Stoer J. Solution of large linear systems of equations by conjugate gradient type methods. In *Mathematical Programming: the State of the Art*, Bachem A, Grötschel M, Korte B (eds). Springer–Verlag: Berlin 1983; 540–565.
24. Varga RL. *Matrix Iterative Analysis*. Prentice–Hall: Englewood Cliffs, NJ, 1962.
25. Van der Vorst HA, Vuik C. The superlinear convergence behaviour of GMRES. *Journal of Computational and Applied Mathematics* 1993; **48**:327–341.
26. Young DM. *Iterative Solution of Large Linear Systems*. Academic Press: New York, 1971.
27. Weiss R. A theoretical overview of Krylov subspace methods. In *Special Issue on Iterative Methods for Linear Systems*, Schönauer W, Weis R (eds). *Applied Numerical Methods* 1983; 540–565.
28. Zítko J. Improving the convergence of iterative methods. *Aplikace Matematiky* 1983; **28**:215–229.
29. Zítko J. Combining the GMRES and a matrix iterative method. In *Proceedings of ICIAM/GAMM 95*, vol. 76. 1996; 595–596.
30. Gantmacher FR. *The Theory of Matrices*. Chelsea: NY, 1959.
31. Zítko J. Using successive approximations for improving the convergence of GMRES method. *Applications of Mathematics* 1998; **43**:321-350.

