NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2004; **11**:in press (DOI: 10.1002/nla.393)

Interpolating functions of matrices on zeros of quasi-kernel polynomials

I. Moret[1,2] and P. Novati[1,2,*,†]

[1] *Universita' degli Studi di Trieste, Dipartimento di Scienze Matematiche, Via Valerio 12, 34100 Trieste, Italy*
[2] *Universita' degli Studi dell'Aquila, Dipartimento di Matematica Pura ed Applicata, Via Vetoio, Coppito, 67010 L'Aquila, Italy*

SUMMARY

The paper deals with Krylov methods for approximating functions of matrices via interpolation. In this frame residual smoothing techniques based on quasi-kernel polynomials are considered. Theoretical results as well as numerical experiments illustrate the effectiveness of our approach. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: Krylov subspace methods; functions of matrices; quasi-kernel polynomials

1. INTRODUCTION

The computation of functions of matrices is a classical topic in numerical linear algebra. In the recent years research in this area has received new impulse due to the introduction of Krylov subspace techniques for the treatment of functions of large and sparse matrices, in particular in the context of the solution of differential problems. Such techniques are projective in nature, since they resort to computing functions of matrices in lower dimensions, but, at the same time, they can also be viewed as polynomial interpolation methods. More precisely, for the computation of $y = f(A)v$, where f is a function analytic in a domain containing the spectrum of the square matrix A, and v is a given vector, these methods produce approximations of the type $y_m = p_{f,m-1}(A)v$, where the polynomial $p_{f,m-1}$, of degree $m-1$, interpolates f (in the Hermite sense) in suitably chosen points. Therefore a great attention must be payed to the choice of such points, with the aim of reflecting the eigenvalue distribution of the argument matrix and maintaining at the same time a limited cost for the arising algorithm. The Ritz values associated to Krylov subspace methods (see References [1–4]) are of course natural candidates to this task. If some information on the spectrum is available, it turns out also to be effective the use of zeros of Chebychev or Faber polynomials (see Reference [5]), or of

*Correspondence to: P. Novati, Universita' degli Studi dell'Aquila, Dipartimento di Matematica Pura ed Applicata, Via Vetoio, Coppito, 67010 L'Aquila, Italy.
†E-mail: novati@univaq.it

Received 17 March 2003
Revised 25 February 2004

Copyright © 2004 John Wiley & Sons, Ltd.

other types of special points employed in complex approximation, such as Fejer [6] or Lejia points [7].

Several error analyses for various specific methods can be found in the literature. Some of them are contained in the papers quoted above. Hence we see that the convergence of the approximations depends on the regularity of the function, on the spectral properties of the argument matrix and on the choice of the interpolation points. Roughly speaking we can say that the regularity of the function characterizes the asymptotic behaviour (cf. Reference [8]), while, when m is not too large, which is the case of practical interest, the others two factors take a more crucial role.

In this paper we are mainly interested to discuss the third factor, that is the choice of the interpolation points, in correspondence with some well-known procedures used for building up Krylov subspaces, such as the incomplete orthogonalization method (IOM) and the unsymmetric Lanczos algorithm, which are popular cost-reducing alternatives to the classical Arnoldi method. In fact, in the practical use of these procedures, dealing with functions with non-removable singularities and with non-normal matrices, we have often observed unstable behaviours with a very oscillating error-curve and in several cases the breakdown of the convergence. In the case of the inverse function, that is dealing with the solution of linear systems, these difficulties are usually overcome by resorting to minimization techniques, which lead to the so-called quasi-minimal residual (QMR) methods [9]. As shown by several numerical experiments, the introduction of such strategies in our more general context produces substantial improvements. In the next section, we present a theoretical analysis which can aid to explain this fact. In Section 3, we revisit the idea of the QMR methods and the underlying concept of quasi-kernel polynomials (cf. [10]). In Section 4, we give a brief description of the algorithms adopted. The numerical results are reported in Section 5.

2. KRYLOV METHODS AND ERROR ESTIMATES

At first let us introduce some notations. In what follows we denote by $\langle , \rangle$ the Euclidean scalar product. The norms of vectors and matrices are the 2-norms. We denote by Π_j the set of the algebraic polynomials of degree $\leqslant j$. The maximum-norm of a function f analytic in a set Ω is represented by $\|f\|_\Omega$. The spectrum of a matrix A is $\sigma(A)$. The mth Krylov subspace generated by A and v ($\|v\| = 1$) is denoted by $K_m(A, v)$. This is the space of all the vectors which can be expressed as $p(A)v$, where $p \in \Pi_{m-1}$. Accordingly, we associate to the vector v, the *grade* of v, with respect to A, that is the degree of the non-zero monic polynomial p of lowest degree such that $p(A)v = 0$. Let m^* ($\leqslant N$) be the *grade* of v. As well known $K_m(A, v)$ has dimension m if and only if $m^* \geqslant m$.

Let us consider the computation of $y = f(A)v$, where A is an $N \times N$ square matrix A and v a given N-dimensional vector. Here and in the sequel, we assume that $\|v\| = 1$. We start by considering a sequence of polynomials $\{q_0, q_1, \ldots, q_j, \ldots\}$, $q_0 = 1$, $q_j \in \Pi_j$ for each j, defined by

$$z q_{j-1}(z) = \sum_{i=1}^{j+1} h_{i,j} q_{i-1}(z) \quad \text{for } j \geqslant 1 \tag{1}$$

Copyright © 2004 John Wiley & Sons, Ltd.

where the $h_{i,j}$'s are scalars, with $h_{j+1,j} \neq 0$ for any j of interest. Furthermore, we consider the sequence of vectors $\{v_1, v_2, \ldots, v_j, \ldots\}$ defined by

$$v_j = q_{j-1}(A)v \in K_j(A, v)$$

Introducing the matrix $V_m = [v_1, v_2, \ldots, v_m]$, for $m \geqslant 1$, we easily obtain the fundamental relationship

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^H \tag{2}$$

where here and below, $e_j \in \mathbb{C}^m$ is the jth unit vector and H_m is the $m \times m$ upper Hessenberg matrix with entries $h_{i,j}$, for $i, j = 1, 2, \ldots, m$ ($h_{i,j} = 0$, for $i > j + 1$). The eigenvalues of H_m, which is *non-derogatory*, are the zeros of the polynomial $q_m(z)$. More precisely,

$$q_m^{(0)}(z) = (\Pi_{j=1}^m h_{j+1,j}) q_m(z)$$

where

$$q_m^{(0)}(z) = \det(zI - H_m)$$

Let $G \subset \mathbb{C}$ be a Cauchy domain which contains both $\sigma(A)$ and $\sigma(H_m)$ and let the function f be analytic in G and continuous on the boundary δG. Let $p_{f,m-1} \in \Pi_{m-1}$ be the polynomial that interpolates f (in the Hermite sense) in the zeros of $q_m^{(0)}(z)$. Then we approximate $y = f(A)v$ by $y_m = p_{f,m-1}(A)v$. By (2) one realizes that

$$y_m = V_m f(H_m) e_1 \tag{3}$$

This means that the y_m can be obtained, without knowing explicitly the interpolation points, by performing a matrix-function evaluation on a lower dimensional space. This reflects the underlying projective nature of the approach.

As mentioned before, we are particularly interested to functions having a critical point, possibly close to the spectrum of A. For simplicity we put this point at zero and accordingly, from now on, we suppose that both A and H_m are non-singular, and that $0 \notin G \cup \delta G$.

In order to estimate the error,

$$E_m = (f(A) - p_{f,m-1}(A))v$$

various approaches can be adopted. As well known, for normal matrices one can work directly on $\sigma(A)$. For non-normal ones, the so-called *departure from normality* (cf. Reference [11, Chaper 11]) takes a crucial role. In these cases the study must consider sets larger than $\sigma(A)$, like the *numerical range* or *ε-pseudospectra*. The starting point is the integral representation,

$$E_m = (f(A) - p_{f,m-1}(A))v = \frac{1}{2\pi \mathrm{i}} \int_\Gamma (f(\lambda) - p_{f,m-1}(\lambda))(\lambda I - A)^{-1} v \, \mathrm{d}\lambda \tag{4}$$

where Γ ($\subset G$) is a suitable contour, chosen in such a way that the resolvent $(\lambda I - A)^{-1}$, as well as $(f(\lambda) - p_{f,m-1}(\lambda))$, for $\lambda \in \Gamma$, can be estimated.

We recall that the *ε-pseudospectrum* of A, here denoted by $\sigma_\varepsilon(A)$, is the set of all $\lambda \in \mathbb{C}$ such that $\|(\lambda - A)^{-1}\| \geqslant \varepsilon^{-1}$, with $\|(\lambda - A)^{-1}\| = \infty$ iff $\lambda \in \sigma(A)$. In other words any eigenvalue of $A + E$ with $\|E\| \leqslant \varepsilon$ belongs to $\sigma_\varepsilon(A)$. For a general discussion and for references on this

Copyright © 2004 John Wiley & Sons, Ltd.

topic see Reference [12]. Accordingly, taking Γ as the contour of $\sigma_\varepsilon(A) \subset G$ and denoting by l_ε the length of Γ, from (4) we have

$$\|E_m\| = \frac{l_\varepsilon}{2\pi\varepsilon} \|f - p_{f,m-1}\|_{\sigma_\varepsilon(A)} \tag{5}$$

The practical evaluation of ε-*pseudospectra* is not, in general, simple to carry out. Various numerical methods have been considered and discussed in Reference [13]. Here, we notice that, as shown in References [10, 14], good approximations to pseudospectra can be obtained from lemniscates associated to the quasi-kernel polynomials later introduced.

Concerning the position of the spectrum of H_m with respect to the pseudospectra of A, we observe what follows. Let λ be an eigenvalue of H_m and assume that $U_m = [u_1, u_2, \ldots, u_m]$ is a matrix such that $U_m^H V_m = I$. Then, either $\lambda \in \sigma(A)$ or, since by (2) it is

$$(A - h_{m+1,m} v_{m+1} u_m^H) V_m = V_m H_m$$

λ is an eigenvalue of $(A - h_{m+1,m} v_{m+1} u_m^H)$ and thus

$$\|(\lambda - A)^{-1}\| \geqslant \|h_{m+1,m} u_m\|^{-1}$$

This generalizes results given in Reference [15] (Proposition 5.2), [16, 17] (Theorem 16).

Estimates of the resolvent norm in (4) can be obtained also referring to the *numerical range* of A, denoted by $W(A)$. In practice, this set can be estimated by the Bendixon–Hirsch theorem (see Reference [18]). Methods for Toeplitz matrices are illustrated in Reference [19]. Among the various interesting properties of $W(A)$, the following well-known inequality relates it to the growth of the resolvent

$$\|(\lambda - A)^{-1}\| \leqslant \frac{1}{\operatorname{dist}(\lambda, W(A))} \tag{6}$$

for $\lambda \notin W(A)$. Thus, if Γ strictly surrounds $W(A)$, then (6) can be used for getting norm estimates for E_m as done in Reference [5]. Connections between $W(A)$ and ε-*pseudospectra* can be found in Reference [17].

Other integral representations can be employed for getting estimates of E_m. Following Reference [3], in light of (3), we express the error as

$$E_m = \frac{1}{2\pi \mathrm{i}} \int_\Gamma f(\lambda)[(\lambda I - A)^{-1} v - V_m(\lambda I - H_m)^{-1} e_1] \, \mathrm{d}\lambda \tag{7}$$

Moreover, since the approximation turns out to be exact when f is a polynomial of degree $\leqslant m - 1$, we have

$$E_m = (f(A) - p_{m-1}(A))v - V_m(f(H_m) - p_{m-1}(H_m))e_1$$

for every polynomial $p_{m-1} \in \Pi_{m-1}$. Thus we write (7) as

$$E_m = \frac{1}{2\pi \mathrm{i}} \int_\Gamma (f(\lambda) - p_{m-1}(\lambda))[(\lambda I - A)^{-1} v - V_m(\lambda I - H_m)^{-1} e_1] \, \mathrm{d}\lambda \tag{8}$$

This expression is particularly meaningful. It shows that the error depends on how well $f(\lambda)$ can be approximated by polynomials and on how well the interpolatory method works on the resolvent $(\lambda I - A)^{-1}$, for $\lambda \in \Gamma$. Here below, this latter aspect will be discussed more in detail.

Copyright © 2004 John Wiley & Sons, Ltd.

We start by rewriting (8) as

$$E_m = \frac{1}{2\pi \mathrm{i}} \int_\Gamma [f(\lambda) - p_{m-1}(\lambda)](\lambda I - A)^{-1} r_m(\lambda)\, \mathrm{d}\lambda \tag{9}$$

for any $p_{m-1} \in \Pi_{m-1}$, where

$$r_m(\lambda) = [v - (\lambda I - A)V_m(\lambda I - H_m)^{-1} e_1]$$

By (2), one can see that

$$r_m(\lambda) = h_{m+1,m}(e_m^H(\lambda I - H_m)^{-1} e_1) v_{m+1} \tag{10}$$

from which, by the properties of H_m, one gets (cf. Reference [5]),

$$r_m(\lambda) = \frac{q_m^{(0)}(A)v}{q_m^{(0)}(\lambda)} \tag{11}$$

Then, let us consider the polynomial

$$\rho_m(\lambda) = \frac{q_m^{(0)}(\lambda)}{q_m^{(0)}(0)}$$

This is the canonical residual-polynomial associated to (2), in the sense that $\rho(A)v = r_m(0)$. Noticing that $r_m(\lambda) = \rho_m(\lambda)^{-1} r_m(0)$, from (9), by simple algebra one obtains the following result, which relates E_m to $r_m(0)$.

Proposition 1
Let Γ be the contour of an ε-*pseudospectrum* $\sigma_\varepsilon(A) \subset G$, and let l_ε be the length of Γ. Assume that, for some $0 < \gamma_m < 1$, the set

$$\Omega_{\gamma_m} = \{\lambda : |\, \rho_m(\lambda)| < \gamma_m\} \tag{12}$$

is contained in $\sigma_\varepsilon(A)$, then, for any $p_{m-1} \in \Pi_{m-1}$,

$$\|E_m\| \leqslant \frac{l_\varepsilon}{2\pi\varepsilon\gamma_m} \|f - p_{m-1}\|_\Gamma \|r_m(0)\| \tag{13}$$

An analogous result can be obtained, with the obvious changes, when (6) can be employed. For a discussion on sets enclosed by lemniscates see Reference [20].

A norm-relationship between E_m and the error $A^{-1}r_m(0)$ can be obtained from (8) by a trivial use of the triangular inequality.

Corollary 2
Under the assumptions of the proposition above, then, for any $p_{m-1} \in \Pi_{m-1}$,

$$\|E_m - A^{-1} r_m(0)\| \leqslant \frac{l_\varepsilon}{2\pi\varepsilon\gamma_m} \|\Delta - p_{m-1}\|_{\sigma_\varepsilon(A)} \|r_m(0)\|$$

where $\Delta(z) = f(z) - z^{-1}$.

Copyright © 2004 John Wiley & Sons, Ltd.

For monitoring the approximations, some authors (see Reference [21]) have proposed the so-called *generalized residual*, defined as

$$R_m = \frac{1}{2\pi \mathrm{i}} \int_\Gamma f(\lambda) r_m(\lambda)\, \mathrm{d}\lambda$$

which is obtained from (7) by replacing the error $[(\lambda I - A)^{-1}v - V_m(\lambda I - H_m)^{-1}e_1]$ with the corresponding residual $r_m(\lambda)$. By (10) it is

$$R_m = h_{m+1,m}(e_m^H f(H_m)e_1)v_{m+1} \tag{14}$$

so that stopping criteria based on R_m can be easy introduced. Numerical experiments show the effectiveness of such criteria.

Remark 3
We notice that, because of the projective nature of the approximations, in various interesting cases R_m is in fact the residual of an equation we are actually solving. For instance suppose that $f(A)v$ represents a solution, at $t>0$, of a differential system of the type

$$Ly + Ay = 0$$

where $L = \mathrm{d}/\mathrm{d}t$, or $L = \mathrm{d}^2/\mathrm{d}t^2$. Then what we actually do is to consider the reduced problem

$$Lu_m + H_m u_m = 0$$

and to approximate y by $V_m u_m$ where u_m is a solution to this. So, by (2), the residual is given by

$$V_m L u_m + A V_m u_m = h_{m+1,m} v_{m+1} e_m^H u_m$$

Observe that, since H_m is upper Hessenberg, we also have, for $m>1$

$$R_m = h_{m+1,m}(e_m^H[(f(H_m) - p_{m-2}(H_m)]e_1)v_{m+1} \tag{15}$$

for every $p_{m-2} \in \Pi_{m-2}$.

Analyses based on the study of R_m have the advantage of taking into account only of H_m moreover it is easy to compare R_m with the residual $r_m(0)$, as we do here below, giving the counterparts of the results stated above for E_m.

Proposition 4
Let $\Gamma \subset G$ be a contour enclosing the set Ω_{γ_m} defined in (12). Then for every $p_{m-2} \in \Pi_{m-2}$ it is

$$\frac{\|R_m\|}{\|r_m(0)\|} \leqslant \frac{l_\Gamma}{2\pi\gamma_m} \|f - p_{m-2}\|_\Gamma \tag{16}$$

where l_Γ is the length of Γ.

Copyright © 2004 John Wiley & Sons, Ltd.

Proof
Write $f(H_m) - p_{m-2}(H_m)$ in the integral form, obtaining, by (15) and (10),

$$R_m = \frac{1}{2\pi \mathrm{i}} \int_\Gamma [f(\lambda) - p_{m-2}(\lambda)] r_m(\lambda)\, \mathrm{d}\lambda$$

Finally use the identity $r_m(\lambda) = \rho_m(\lambda)^{-1} r_m(0)$. □

Corollary 5
Under the assumptions of proposition (4), for any $p_{m-2} \in \Pi_{m-1}$, we have

$$\frac{\|R_m - r_m(0)\|}{\|r_m(0)\|} \leqslant \frac{l_\Gamma}{2\pi\gamma_m} \|\Delta - p_{m-2}\|_\Gamma$$

where $\Delta(\lambda) = f(\lambda) - \lambda^{-1}$.

3. QUASI-KERNEL POLYNOMIALS

The relationships stated in the previous section stress the fact, confirmed by several numerical applications, that when a 'good' decay of the residuals $r_m(0)$ occurs it is reasonable to expect a 'good' behaviour of the approximations. This suggest to choose the interpolation points in accordance to residual-smoothing strategies adopted in the context of the solution of linear systems. In order to do this we employ the basic ideas of the well-known quasi-minimal residual methods (cf. Reference [9]). Here below we revisit that approach.

At first, let us consider the general process for transforming bases of $K_m(A, v)$. Let $\{w_1, w_2, \ldots, w_j, \ldots\}$ be a sequence of vectors, such that $w_1 = v$, and $\|w_j\| = 1$ for every j, to which, for $m \geqslant 1$, we now associate the matrix

$$W_m = [w_1, w_2, \ldots, w_m]$$

Assume that $S_m = [s_{i,j}]$ (with $s_{1,1} = 1$) is an $m \times m$ non-singular upper triangular matrix such that,

$$W_m = V_m S_m$$

Let us represent S_{m+1} in the partitioned form,

$$\begin{matrix} S_m & d_m \\ 0 & s_{m+1,m+1} \end{matrix} \tag{17}$$

Proposition 6
Under the above assumptions and notation we have,

$$AW_m = W_m S_m^{-1} H_m^{(1)} S_m + k_{m+1,m} w_{m+1} e_m^H \tag{18}$$

where

$$H_m^{(1)} = H_m - h_{m+1,m} b_m e_m^H$$

Copyright © 2004 John Wiley & Sons, Ltd.

with

$$b_m = \frac{d_m}{s_{m+1,m+1}}$$

$$s_{m+1,m+1} = \|V_m b_m + v_{m+1}\|^{-1}$$

and

$$k_{m+1,m} = \frac{h_{m+1,m} s_{m,m}}{s_{m+1,m+1}} \tag{19}$$

Proof
By our assumptions we get

$$AW_m = W_m S_m^{-1} H_m S_m + h_{m+1,m} v_{m+1} e_m^H S_m$$

Since

$$w_{m+1} = V_{m+1} S_{m+1} e_{m+1} = V_m d_m + s_{m+1,m+1} v_{m+1} \tag{20}$$

the results follows by simple computation. □

It is easy to see that both the matrices $H_m^{(1)}$ and $S_m^{-1} H_m^{(1)} S_m$ are upper Hessenberg. The Sherman–Morrison formula leads to the following result.

Proposition 7
If H_m is non-singular, then $H_m^{(1)}$ is non-singular provided that

$$1 + h_{m+1,m} e_m^H H_m^{-1} b_m \neq 0$$

Now, the mth approximation to $f(A)v$, associated to the new basis defined by (18), i.e. that obtained by interpolating in the eigenvalues of $H_m^{(1)}$ (assuming that f is analytic there) is

$$W_m f(S_m^{-1} H_m^{(1)} S_m) e_1 = V_m f(H_m^{(1)}) e_1 \tag{21}$$

Thus, the explicit knowledge of W_m is not required. In other words, now, instead of (2), we use

$$AV_m = V_m H_m^{(1)} + \frac{h_{m+1,m}}{s_{m+1,m+1}} w_{m+1} e_m^H \tag{22}$$

which is a reformulation of (18).

Assume that H_m is non-singular and that b_m is chosen in such a way that $H_m^{(1)}$ is non-singular too. Then consider the residual,

$$\bar{r}_m^{(1)}(0) = v - AV_m H_m^{(1)-1} e_1 \tag{23}$$

By the usual argument (cf. (11)), referring now to the polynomial

$$\bar{q}_m^{(0)}(z) = \det(zI - S_m^{-1} H_m^{(1)} S_m) = \det(zI - H_m^{(1)})$$

Copyright © 2004 John Wiley & Sons, Ltd.

we have

$$\bar{r}_m^{(1)}(0) = \frac{\bar{q}_m^{(0)}(A)}{\bar{q}_m^{(0)}(0)} v$$

Since,

$$\bar{q}_m^{(0)}(A)v = \left(\prod_{j=1}^{m} k_{j+1,j} \right) w_{m+1}$$

we have

$$\|\bar{q}_m^{(0)}(A)v\| = \left| \prod_{j=1}^{m} h_{j+1,j} \right| \|V_m b_m + v_{m+1}\|$$

and thus

$$\|\bar{r}_m(0)\| = \frac{|\prod_{j=1}^{m} h_{j+1,j}| \|V_m b_m + v_{m+1}\|}{|\det(H_m)(1 - h_{m+1,m} e_m^H H_m{}^{-1} b_m)|} \tag{24}$$

$$= \|r_m(0)\| \frac{\|V_m b_m + v_{m+1}\|}{|(1 - h_{m+1,m} e_m^H H_m{}^{-1} b_m)|} \tag{25}$$

Clearly the ideal choice of b_m comes from minimizing

$$\Psi(b_m) = \frac{\|V_m b_m + v_{m+1}\|}{|(1 - h_{m+1,m} e_m^H H_m{}^{-1} b_m)|} \tag{26}$$

In the case of orthonormal basis, this minimization problem is solvable without any substantial extra cost and we get the popular GMRES method. In the general case one usually resorts to the so-called quasi-minimization. In our framework this consists in minimizing the functional,

$$\bar{\Psi}(b_m) = \frac{\|\bar{b}_m\|}{|(1 - h_{m+1,m} e_m^H H_m{}^{-1} b_m)|}$$

where

$$\bar{b}_m = \begin{matrix} b_m \\ 1 \end{matrix}$$

instead of Ψ. More in general one could take $\bar{b}_m = D\bar{b}_m$, with D a suitable weight diagonal matrix (cf. Reference [9]).

A standard minimization technique shows that $\bar{\Psi}$ is minimized by

$$b_m = -h_{m+1,m} H_m{}^{-H} e_m \tag{27}$$

We can now consider the new matrix,

$$H_m^{(1)} = H_m + h_{m+1,m}^2 H_m^{-H} e_m e_m^H = H_m^{-H} (\tilde{H}_m^H \tilde{H}_m) \tag{28}$$

Copyright © 2004 John Wiley & Sons, Ltd.

and the corresponding approximation (21), which we denote by the acronymus *QK*, in accordance with the term *Quasi-Kernel polynomial* used in Reference [10] for denoting the corresponding residual polynomial

$$\bar{\rho}_m(\lambda) = \frac{\bar{q}_m^0(z)}{\bar{q}_m^0(0)}$$

As previously mentioned, it has been observed that in several cases the corresponding sets Ω_{γ_m}, defined by (12), that is sets enclosed by lemniscates of $\bar{\rho}_m$, represent good approximations to pseudospectra of A. Concerning the location of eigenvalues of matrix (28), the following result generalizes that given in Reference [22] for orthonormal bases.

Proposition 8
Let μ be any eigenvalue of the matrix $H_m^{(1)}$ defined by (28). Assume that U_{m+1} is a matrix such that $U_{m+1}^H V_{m+1} = I$. Then $\mu^{-1} \in W(U_{m+1}^H A^{-1} V_{m+1})$.

Proof
Let $H_m^{(1)} x = \mu x,\ x \neq 0$. Recall that b_m minimizes $\bar{\Psi}(b_m)$ that is $\|U_{m+1}^H \bar{r}_m^{(1)}(0)\|$ (cf. (23)). This means that w_{m+1} must be orthogonal to $U_{m+1}^H A V_m x$ (which is $\neq 0$).

From (22) we get,

$$U_{m+1}^H A V_m = U_{m+1}^H V_m H_m^{(1)} + k_{m+1,m} U_{m+1}^H w_{m+1} e_m^{\mathrm{T}}$$

Hence we get,

$$\langle U_{m+1}^H A V_{m+1} x, U_{m+1}^H A V_{m+1} x \rangle = \mu \langle U_{m+1}^H A V_{m+1} x, U_{m+1}^H V_m x \rangle$$

Then, setting $y = U_{m+1}^H A V_{m+1} x$, since $A V_m = V_{m+1} U_{m+1}^H A V_m$, we obtain

$$\langle y, y \rangle = \mu \langle y, U_{m+1}^H A^{-1} V_{m+1} y \rangle$$ □

4. THE BASIC METHODS

The commonly adopted procedures for building up Krylov subspaces, in absence of particular informations on the spectrum, are based on Arnoldi or Lanczos algorithms.

For our purposes, we have implicitly supposed that the algorithms fulfill exactly the fundamental formula (2). Actually instabilities may occur and may influence important properties, such as orthogonality or biorthogonality. Fortunately, this possible drawbacks can be often substantially reduced by resorting to clever implementations. Analyses which take into account of the errors due to computer arithmetic can be found in References [1, 4, 23].

The (full) Arnoldi method (cf. Reference [24]) constitutes our reference algorithm. It generates a sequence of orthonormal vectors $\{v_1, v_2, \ldots, v_j, \ldots\}$ such that, for $m \leqslant m^*$, $K_m(A, v) = \text{span}\{v_1, v_2, \ldots, v_m\}$. The entries of the matrices H_m are

$$h_{i,j} = \langle v_i, A v_j \rangle$$

For a collection of various implementations of the Arnoldi method, as well as for a discussion about, we refer to Reference [25]. A stable implementation ensures that $W(H_m) \subseteq W(A)$.

Copyright © 2004 John Wiley & Sons, Ltd.

The eigenvalues of H_m are usually called the Arnoldi–Ritz values. The QK-Arnoldi method applied to $f(z)=z^{-1}$ corresponds to GMRES. If we denote by μ any eigenvalue of the matrix $H_m^{(1)}$ defined by (28), Proposition 8 says that $\mu^{-1}\in W(A^{-1})$. For the functions we have considered, we observed that the two interpolatory methods (Arnoldi and QK-Arnoldi) give substantially very similar numerical results. They often perform well in terms of speed of convergence, but due to the growth of the work required, they become less effective, in terms of computational costs, with respect to other procedures.

As first alternative to Arnoldi method we have considered the *incomplete orthogonalization method* (IOM). We refer to References [26, 27], for a detailed discussion. This method consists in performing the orthogonalization of the vector v_{j+1} only against the previous p vectors $v_{j-p}, v_{j-p+1}, \ldots, v_j$. Usually it is denoted by IOM(p). This technique is suggested by the (heuristic) observation that, for fixed i, the elements $h_{i,j}$, produced by the Arnoldi algorithm, are decreasing as the column index j increases. In the IOM(p) the matrix H_m has a band structure, with possible non-zero elements only for $i-1\leqslant j\leqslant i+p$. Dealing with various functions, we performed several numerical experiments using IOM(p), with different reasonable values of p. In most of the cases we observed an instable behaviour of the approximations (as illustrated by the examples below). Yet, we were able to eliminate such instabilities by resorting to the corresponding QK variant (QK-IOM(p)), getting so a very effective algorithm.

We arrived to the same conclusion even dealing with the *non-symmetric Lanczos process*, which represents the classical alternative to Arnoldi type algorithms. As well known, given $v_1=v$ and $u_1=u$, $\langle u,v\rangle\neq 0$, this method produces two biorthogonal sequences of vectors, $\{v_1,v_2,\ldots,v_j,\ldots\}$ and $\{u_1,u_2,\ldots,u_j,\ldots\}$, such that, for each m, $K_m(A,v)=\mathrm{span}\{v_1,v_2,\ldots,v_m\}$ and $K_m(A^{\mathrm{T}},u)=\mathrm{span}\{u_1,u_2,\ldots,u_m\}$. The sequence $\{v_1,v_2,\ldots,v_j,\ldots\}$ fulfills (2) and the matrix H_m is now tridiagonal. It is known that the Lanczos process could have 'breakdowns', that is it may have a premature stop. This can be overcome by resorting to variants based on the so-called look-ahead technique. In this case H_m will have a block tridiagonal structure. For computational schemes, details and references we refer to Reference [28]. Some numerical experiments in the next section will be dedicated to illustrate the advantages of considering the QK counterpart of Lanczos method.

5. NUMERICAL EXPERIMENTS

In our experiments we consider two types of matrices which are often adopted in the numerical tests.

The matrix A_1 is the GRCAR matrix, a Toeplitz matrix with sensitive eigenvalues, defined as

$$A_1=\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & \cdots & \\ -1 & 1 & 1 & 1 & 1 & 0 & \cdots \\ 0 & -1 & 1 & 1 & 1 & 1 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Copyright © 2004 John Wiley & Sons, Ltd.

This test matrix is interesting because A_1 is non-normal and $\sigma(A_1)$ is located in a curve, symmetric with respect to the real axis, close to the origin. It was used in References [10, 14] for illustrating the dependence of Krylov methods on pseudospectra and for showing that pseudospectra can be well approximated by sets enclosed by lemniscates of quasi-kernel polynomials.

The second matrix, we denote by A_2, is obtained by discretizing on the unit square $(0,1)\times(0,1)$, with Dirichelet boundary conditions, the second-order differential operator,

$$-\Delta + c_1\frac{\partial}{\partial x} + c_2\frac{\partial}{\partial y}, \quad c_1, c_2 \in \mathbb{R}$$

where Δ denotes the Laplacian. We introduce a uniform meshgrid with meshsize $h = 1/(n+1)$ and we employ the central differences for discretizing the Laplacian. Since we will consider only cases where the grid-Péclet numbers are greater than 1, we use the upwind differences for the advection terms (cf. Reference [29]). So doing we get the $N\times N$ $(N = n^2)$ matrix

$$A_2 = \frac{1}{h^2}[(I_n \otimes B) + (C \otimes I_n)] \in \mathbb{R}^{n^2\times n^2}$$

where I_n is the identity matrix of order n, B and C are tridiagonal matrices defined as

$$B = \begin{bmatrix} 4+c_1h+c_2h & -1 & & \\ -(1+c_1h) & 4+c_1h+c_2h & -1 & \\ & -(1+c_1h) & \ddots & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & -1 & & \\ -(1+c_2h) & 0 & -1 & \\ & -(1+c_2h) & \ddots & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

For a discussion on spectra of such matrices we refer to Reference [7].

In all the experiments we have set $v = [1, 1, \ldots, 1]^{\mathrm{T}}/N^{1/2}$. The approximations are evaluated via formula (3) and the computation of $f(H_m)e_1$ is performed by a standard MatLab routine. In the figures we report the behaviour (as m increases) of the Euclidean norm of the error E_m versus the computational cost measured in terms of scalar products. Actually a similar behaviour has been observed for the generalized residuals R_m. We must point out that the cost of a matrix by vector product is considered in terms of a certain number of scalar products, that depends on the sparsity pattern of the matrix.

In all examples we consider a comparison between the IOM(p) or the Lanczos method and their QK versions. The numerical results (Figures 1–7) show the effectiveness of the QK

Copyright © 2004 John Wiley & Sons, Ltd.

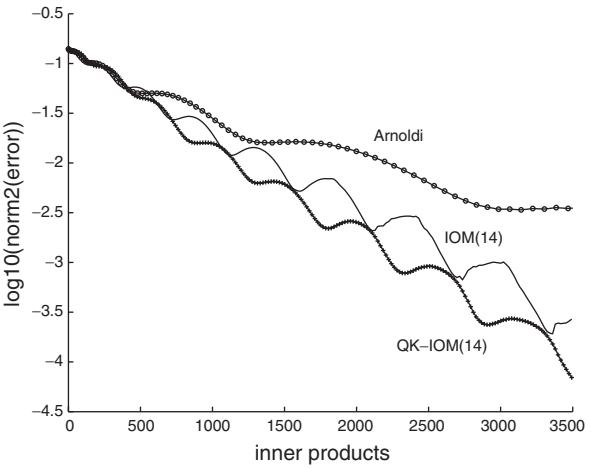

Figure 1.

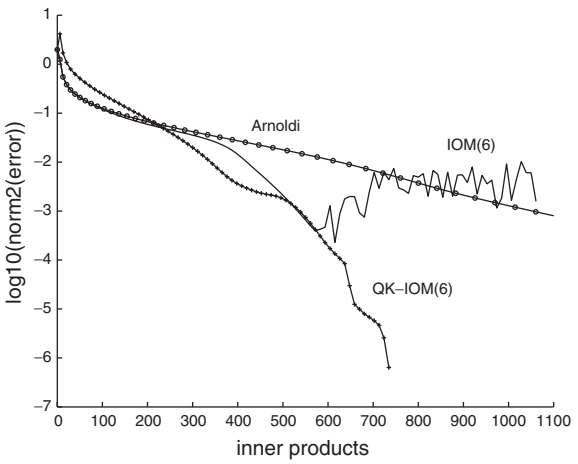

Figure 2.

approach especially in terms of stability. Moreover, regarding the IOM(p), it is possible to observe (see Example 2) that using small values of p the QK version converges even when the basic version does not. For comparison, in the pictures, the error curve of the Arnoldi method is also shown. As already mentioned, it represents approximately also the error curve of its QK version.

Example 1

In this first example we consider the function $f(z) = \exp(-tz^{-1})$, for $t > 0$, which has an essential singularity at $z = 0$. In Figure 1 we give a comparison between the IOM(14) and QK-IOM(14) when applied to the computation of $f(A_1)v$, with $t = 0.1$. The dimension of the problem is $N = 200$.

Copyright © 2004 John Wiley & Sons, Ltd.

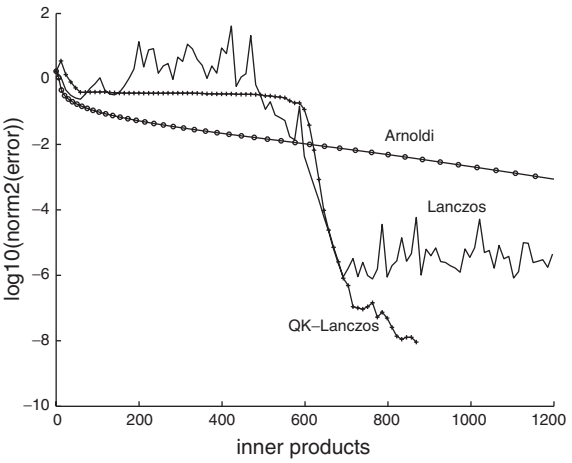

Figure 3.

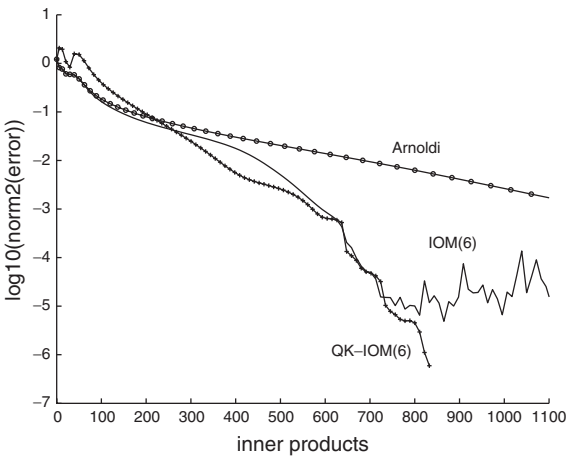

Figure 4.

Example 2

We consider the function $f(z) = z^{1/2}$. We recall that the function square root denotes the single-valued branch of the many-valued function $z^{1/2}$ that is defined in a domain not containing the origin (cf. References [30, p. 77], [31, vol. 1, p. 114]). We take the square root on the branch such that $1^{1/2} = 1$.

The problem examined in this example is the computation of $f(sA_2)v$, with $s > 0$. As well-known (see Reference [32]) $A_2^{1/2}$ is well defined.

Fixed $n = 30$, $s = 0.01$, $c_1 = 50$, $c_2 = 50$, in Figure 2 the error curves of the IOM(6) with respect to its QK version are shown, whereas in Figure 3, we test the Lanczos and the QK-Lanczos methods on the same problem.

Copyright © 2004 John Wiley & Sons, Ltd.

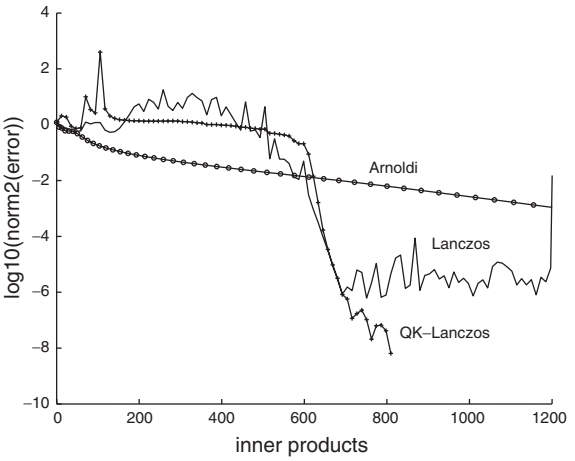

Figure 5.

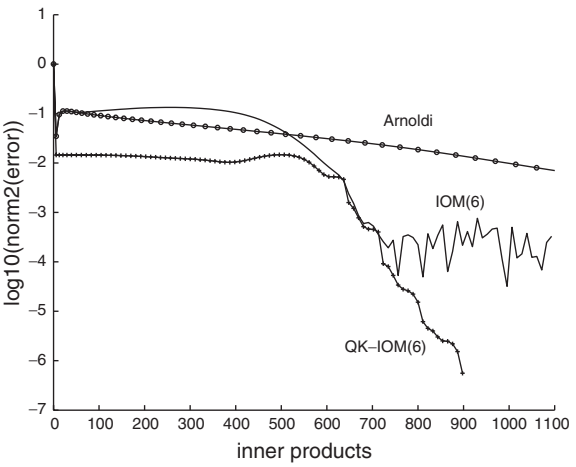

Figure 6.

Example 3

The function considered in this example is $f(z) = \sin(tz^{1/2})$. In Figures 4 and 5 we test our methods for the computation of $f(A_2)v$. In both the tests we have chosen $n = 30,\ t = 0.1,\ c_1 = 60,\ c_2 = 50$.

Example 4

Now consider $f(z) = (1 - \exp(-z))^{-1}$. This function has a pole of order 1 at $z = 0$. In fact its Laurent expansion is

$$f(z) = z^{-1} + \sum_{k=0}^{\infty} a_k z^k$$

Copyright © 2004 John Wiley & Sons, Ltd.

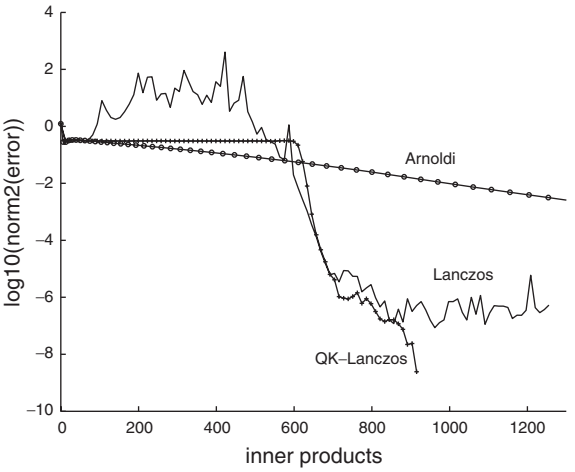

Figure 7.

for some constants $a_0,\ a_1,\ldots$. We are interested to compute $f(\xi A_2)v$, for $\xi>0$. It represents the vector $y(0)$, being $y(t)$ the solution of the differential problem

$$y' + A_2 y = 0$$

$$y(0) = y(\xi) + v$$

The following two pictures, Figures 6 and 7, show the results obtained taking $c_1 = c_2 = 100$ and $c_1 = c_2 = 50$, respectively. In both cases $n = 30$ and $\xi = 0.01$.

REFERENCES

1. Kniznherman L. Calculation of functions of unsymmetric matrices using Arnoldi's method. *Computational Mathematics and Mathematical Physics* 1991; **31**(1):1–9.
2. Saad Y. Analysis of some Krylov subspace approximations to the matrix exponential operator. *SIAM Journal on Numerical Analysis* 1992; **29**:209–228.
3. Hochbruck M, Lubich C. On Krylov subspace approximation to the matrix exponential operator. *SIAM Journal on Numerical Analysis* 1995; **34**(5):1911–1925.
4. Druskin V, Greenbaum A, Knizherman L. Using nonorthogonal Lanczos vectors in the computation of matrix functions. *SIAM Journal on Scientific Computing* 1998; **19**(1):38–54.
5. Moret I, Novati P. An interpolatory approximation of the matrix exponential based on Faber polynomials. *Journal of Computational and Applied Mathematics* 2001; **131**:361–380.
6. Novati P. A method based on Fejèr points for the computation of functions of nonsymmetric matrices. *Applied Numerical Mathematics* 2003; **44**:201–224.
7. Caliari M, Vianello M, Bergamaschi L. Interpolating discrete advection-diffusion propagators at spectral Leja sequences. Università di Padova, 2002, preprint.
8. Tal-Ezer H. High degree polynomial interpolation in Newton form. *SIAM Journal on Statistical and Scientific Computing* 1991; **12**:648–667.
9. Freund RW, Nachtigal NM. QMR: a quasi-minimal residual method for non-Hermitian linear systems. *Numerische Mathematik* 1991; **60**:315–339.
10. Freund RW. Quasi-kernel polynomials and their use in non-Hermitian matrix iterations. *Journal of Computational and Applied Mathematics* 1992; **43**:135–158.
11. Golub GH, Van Loan CF. *Matrix Computations.* North Oxford Academic: Oxford, 1983.
12. Trefethen LN. Pseudospectra of linear operators. *SIAM Review* 1997; **39**(3):383–406.
13. Trefeten LN. Computation of pseudospectra. *Acta Numerica* 1999; **8**:247–295.

Copyright © 2004 John Wiley & Sons, Ltd.

14. Nachtigal NM, Reichel L, Trefethen LN. A hybrid GMRES algorithm for nonsymmetric linear systems. *SIAM Journal on Matrix Analysis and Applications* 1992; **13**:796–825.
15. Nevanlinna O. Hessenberg matrices in Krylov subspaces and the computation of the spectrum. *Numerical Functional Analysis and Optimization* 1995; **16**(3–4):443–473.
16. Simoncini V, Gallopulos E. Transfer functions and resolvent norm approximation of large matrices. *Electronic Transactions on Numerical Analysis* 1998; **7**:190–201.
17. Embree M, Trefethen LN. Generalizing eigenvalue theorems to pseudospectra theorems. *SIAM Journal on Scientific Computing* 2001; **23**:583–590.
18. Halmos PR. *A Hilbert Space Problem Book.* D. Van Nostrand Company, Inc.: Princeton, 1967.
19. Eiermann M. Fields of values and iterative methods. *Linear Algebra and its Applications* 1993; **180**:167–197.
20. Walsh JL. *Interpolation and Approximation by Rational Functions in the Complex Domain*. AMS: New York, 1935.
21. Hochbruck M, Lubich C, Selhofer H. Exponential integrators for large systems of differential equations. *SIAM Journal on Scientific Computing* 1998; **19**:1552–1574.
22. Manteuffel TA, Starke G. On hybrid iterative methods for nonsymmetric systems of linear equations. *Numerische Mathematik* 1991; **73**:489–506.
23. Druskin V, Knizherman L. Krylov subspace approximations of eigenpairs and matrix functions in exact and computer arithmetic. *Numerical Linear Algebra with Applications* 1995; **2**:205–217.
24. Trefethen LN, Bau D. *Numerical Linear Algebra*. SIAM: Philadelphia, 1997.
25. Drkosova' J, Greenbaum A, Rozloznik M, Strakos Z. Numerical stability of GMRES. *BIT* 1995; **35**:309–330.
26. Saad Y. Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra and its Applications* 1980; **34**:269–295.
27. Saad Y. The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems. *SIAM Journal on Numerical Analysis* 1982; **19**:485–506.
28. Gutknecht MH. Lanczos-type solvers for unsymmetric linear systems of equations. *Acta Numerica* 1997; **6**: 271–397.
29. Quarteroni A, Valli A. *Numerical Approximation of Partial Differential Equations.* Springer Series in Computational Mathematics, vol. 23. Springer: Berlin, 1994.
30. Nehari Z. *Conformal Mapping.* Mc Graw-Hill: New York, 1952.
31. Gantmacher FR. *The Theory of Matrices*. AMS Chelsea Publ.: Providence, 1998.
32. Kato T. *Perturbation Theory for Linear Operators*. Springer: Berlin, 1976.

Copyright © 2004 John Wiley & Sons, Ltd.