

Pseudospectra of rectangular matrices

THOMAS G. WRIGHT AND LLOYD N. TREFETHEN

Oxford University Computing Laboratory, Parks Road, Oxford OX1 3QD, UK

[Received on 30 July 2001; revised on 15 January 2002]

Pseudospectra of rectangular matrices vary continuously with the matrix entries, a feature that eigenvalues of these matrices do not have. Some properties of eigenvalues and pseudospectra of rectangular matrices are explored, and an efficient algorithm for the computation of pseudospectra is proposed. Applications are given in (square) eigenvalue computation (Lanczos iteration), square pseudospectra approximation (Arnoldi iteration), control theory (nearest uncontrollable system) and game theory.

Keywords: eigenvalues; pseudospectra; rectangular matrices; control theory.

1. Introduction

Let A be an $m \times n$ matrix with $m \geq n$. An eigenvalue of A might be defined by the condition

$$(A - \lambda \tilde{I})v = 0 \quad (1.1)$$

for some nonzero n -vector v , where \tilde{I} denotes the $m \times n$ ‘identity’ with 1 on the main diagonal and 0 elsewhere. This definition (sometimes in the generalized case with $\tilde{I} = B$) has been used by Boley (1990), Thompson & Weil (1972) and Stewart (1994), for example, but could not be said to have become a mainstream concept of linear algebra. Applications are limited, and (1.1) has the awkward feature that most matrices have no eigenvalues at all, whilst for those that do, an infinitesimal perturbation will in general remove them. This can be seen by noting that if (λ, v) satisfies (1.1), then we have

$$\begin{pmatrix} A_1 - \lambda I_n \\ A_2 \end{pmatrix} v = 0 \quad (1.2)$$

(with A_1 denoting the upper $n \times n$ section of A), so not only must (λ, v) be an eigenpair of A_1 , but v must also be in the nullspace of A_2 .

We have found that the pseudospectra of a rectangular matrix may be more useful. We say that λ is an ϵ -pseudoeigenvalue of A if for some n -vector v with $\|v\| = 1$ (a corresponding pseudoeigenvector),

$$\|(A - \lambda \tilde{I})v\| \leq \epsilon.$$

For the moment we take $\|\cdot\|$ to represent the matrix norm induced by the Euclidean norms on the domain and range spaces \mathbb{C}^n and \mathbb{C}^m ; generalizations are discussed in Section 6. The ϵ -pseudospectrum of A , $\Lambda_\epsilon(A)$, is the set of ϵ -pseudoeigenvalues, a closed subset

of \mathbb{C} . As we will note, $\Lambda_\epsilon(A)$ depends continuously on A (for $\epsilon > 0$) and may be empty for non-zero ϵ (unlike the pseudospectra of square matrices).

So far as we are aware, pseudospectra of rectangular matrices have been considered previously only by Toh & Trefethen (1996), Wright & Trefethen (2001) and Higham & Tisseur (2002). The aim of the present paper is to explore this notion more systematically, to propose algorithms, and to consider possible applications. We begin in Section 2 by setting forth fundamental properties, and elaborate on the basic ideas in Section 3 by presenting simple examples. Our set of proposed algorithms is described in Section 4; these are incorporated in the first author's Pseudospectra GUI (Wright, 2001), a graphical Matlab code. Section 5 considers various applications, including the analysis of Lanczos iterations for symmetric matrices, even though symmetric matrices are a field where one might not expect pseudospectra to have much use. Section 6 considers generalizations such as the generalized rectangular eigenvalue problem

$$(A - \lambda B)v = 0$$

and non-Euclidean norms, and finally Section 7 summarizes some of our observations.

Though pseudospectra of rectangular matrices have rarely been considered explicitly in the past, our definitions and algorithms are closely related to certain developments in the study of controllability, and distance to uncontrollability, of linear control systems. Some of these connections are mentioned in Section 5.

2. Basic properties

Throughout this paper, we assume that A is an $m \times n$ matrix (real or complex) with $m \geq n$. Probably none of the results in this section are fundamentally new, although they may not have been put together in this form before.

The definition of ϵ -pseudospectra of rectangular matrices just given,

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C} : \|(A - z\tilde{I})v\| \leq \epsilon \text{ for some } v \in \mathbb{C}^n, \|v\| = 1\}, \quad (2.1)$$

suggests an equivalent definition in terms of exact eigenvalues of perturbed matrices (where $\Lambda(A)$ denotes the set of eigenvalues of A),

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C} : z \in \Lambda(A + E) \text{ for some } E \text{ with } \|E\| \leq \epsilon\}. \quad (2.2)$$

To see that these two definitions are equivalent, suppose first that for $z \in \mathbb{C}$ there exists some unit vector $v \in \mathbb{C}^n$ such that $\|(A - z\tilde{I})v\| \leq \epsilon$. Define $y = (A - z\tilde{I})v$, so that $\|y\| \leq \epsilon$, and define $E = -yv^*$. Then

$$(A + E)v = Av - yv^*v = z\tilde{I}v,$$

so z is an eigenvalue of $A + E$ with $\|E\| \leq \epsilon$. On the other hand, if $(A + E)v = z\tilde{I}v$ for some E with $\|E\| \leq \epsilon$ and $\|v\| = 1$, then $\|(A - z\tilde{I})v\| = \|Ev\| \leq \epsilon$. (This result is proved in the more general rectangular matrix polynomial case by Higham & Tisseur (2002).)

For both analytical and computational reasons, it is often convenient to characterize $\Lambda_\epsilon(A)$ in terms of the pseudoinverse, which can be defined in the following manner. Let

the singular value decomposition (SVD) of $z\tilde{I} - A$ be given by

$$z\tilde{I} - A = U\Sigma V^* = \sum_{j=1}^p \sigma_j u_j v_j^*, \tag{2.3}$$

where $p = \min(m, n)$, $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary (with columns u_j and v_j respectively), and $\Sigma \in \mathbb{C}^{m \times n}$ is diagonal with main diagonal entries σ_j , ordered in non-increasing order. The pseudoinverse $(z\tilde{I} - A)^+$ can be written as

$$(z\tilde{I} - A)^+ = \sum_{j=1}^p \frac{1}{\sigma_j} v_j u_j^*, \tag{2.4}$$

see e.g. Golub & Van Loan (1996); Stewart & Sun (1990). (Note that our definition of the pseudoinverse differs from that in the above references which take $p = \text{rank}(z\tilde{I} - A)$, since for rank-deficient matrices our pseudoinverse is undefined.)

This leads to the following definition of pseudospectra:

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C} : \|(z\tilde{I} - A)^+\| \geq \epsilon^{-1}\} \tag{2.5}$$

(where the norm is taken to be infinite if the pseudoinverse is undefined). From (2.4) and the unitary invariance of the 2-norm, it is clear that (2.5) is equivalent to

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C} : \sigma_{\min}(z\tilde{I} - A) \leq \epsilon\}. \tag{2.6}$$

To show that these definitions are also equivalent to (2.1) and (2.2), we first prove that (2.1) implies (2.5), and then that (2.6) implies (2.1). Suppose that for $z \in \mathbb{C}$ there exists a unit vector $v \in \mathbb{C}^n$ such that $(A - z\tilde{I})v = -y$ with $\|y\| \leq \epsilon$. Pre-multiplying this equation by the pseudoinverse yields $v = (z\tilde{I} - A)^+ y$, and since $1 = \|v\| \leq \|(z\tilde{I} - A)^+\| \|y\|$, we have $\|(z\tilde{I} - A)^+\| \geq \epsilon^{-1}$. The final part of our equivalence follows from noting that the SVD (2.3) implies that $\|(z\tilde{I} - A)v_p\| = \|\sigma_p u_p v_p^* v_p\| = \sigma_p = \sigma_{\min}(z\tilde{I} - A)$.

We have established the following theorem.

THEOREM 2.1 The definitions (2.1), (2.2), (2.5), and (2.6) are equivalent.

One of the reasons for studying pseudospectra of rectangular matrices is that $\Lambda_\epsilon(A)$ depends continuously on A for $\epsilon > 0$, even though the eigenvalues of A can vanish under infinitesimal perturbations. From definition (2.6), for example, it can be seen that $\Lambda_\epsilon(A)$ changes continuously with $\epsilon > 0$ since the singular values are continuous functions of the matrix entries.

Another motivation for this work is the following monotonicity result (expressed in ‘Matlab notation’).

THEOREM 2.2 Let A be an $m \times n$ matrix. Then

- (i) $\Lambda_\epsilon(A(:, 1:k)) \subseteq \Lambda_\epsilon(A(:, 1:k+1))$, $1 \leq k < n$.
- (ii) $\Lambda_\epsilon(A(1:k+1, :)) \subseteq \Lambda_\epsilon(A(1:k, :))$, $1 \leq k < m$.

Proof. Both results follow from the definition of the minimum singular value of a matrix B ,

$$\sigma_{\min}(B) = \min_{\|x\|_2=1} \|Bx\|_2,$$

since (i) removing columns of B can only increase this minimum, whereas (ii) removing rows of B can only decrease it. \square

An alternative explanation for part (ii) of Theorem 2.2 (for $k \geq n$) can be obtained as follows: for (λ, v) to be a ϵ -pseudoeigenpair of A , not only must it be a ϵ -pseudoeigenpair of the upper $n \times n$ section A_1 of A , but v must also be ‘nearly’ in the null-space of A_2 , the lower $(m - n) \times n$ section of A , since

$$\epsilon^2 \geq \|(\lambda \tilde{I} - A)v\|_2^2 = \|(\lambda I_n - A_1)v\|_2^2 + \|A_2v\|_2^2.$$

Removing rows of A_2 is equivalent to reducing $\|A_2v\|_2$, or in other words, allowing (λ, v) to be an ϵ -pseudoeigenpair of A_1 for a larger value of ϵ .

Many theorems for pseudospectra of square matrices also apply to rectangular matrices, or at least have analogues. For example, Embree & Trefethen (2001) present generalizations of 16 eigenvalue theorems to pseudospectra theorems, and about half of their pseudospectra theorems also apply in the rectangular case. Here are two of them.

THEOREM 2.3 Let A be an $m \times n$ matrix. Then

- (i) $\lambda \in \Lambda(A) \implies \lambda \in \Lambda_\epsilon(A)$ for all $\epsilon \geq 0$;
- (ii) $\Lambda_{\epsilon|\beta|}(\alpha \tilde{I} + \beta A) = \alpha + \beta \Lambda_\epsilon(A)$ for $\alpha, \beta \in \mathbb{C}$.

Proof.

- (i) This follows immediately from the definitions of $\Lambda(A)$ and $\Lambda_\epsilon(A)$.
- (ii) This is trivial for $\beta = 0$. For $\beta \neq 0$, we have

$$|\beta| \| (z \tilde{I} - (\alpha \tilde{I} + \beta A))^+ \| = \| (\beta^{-1}(z - \alpha) \tilde{I} - A)^+ \|. \quad \square$$

This theorem and the preceding discussion may give the impression that the mathematics of pseudospectra of rectangular matrices is much like that for square matrices. However, the opposite is true: the rectangular case is a very different problem. For example, for a square matrix A the resolvent $(zI - A)^{-1}$ is an analytic function of z for $z \notin \Lambda(A)$, and this implies that $\log \|(zI - A)^{-1}\|$ is subharmonic, that $\|(zI - A)^{-1}\|$ satisfies a maximum principle, and that projections and other matrix functions can be computed by Cauchy integrals. All these properties vanish in the rectangular case, beginning with the fact that because the pseudoinverse of a matrix B is defined by a formula $B^+ = (B^*B)^{-1}B^*$, which contains a complex conjugate, $(zI - A)^+$ does not depend analytically on z when A is rectangular. A rectangular matrix cannot be construed as mapping a vector space to itself, there is no determinant or Jordan canonical form, and its properties cannot be obtained by reduction to eigenvalues. We must rely on other techniques.

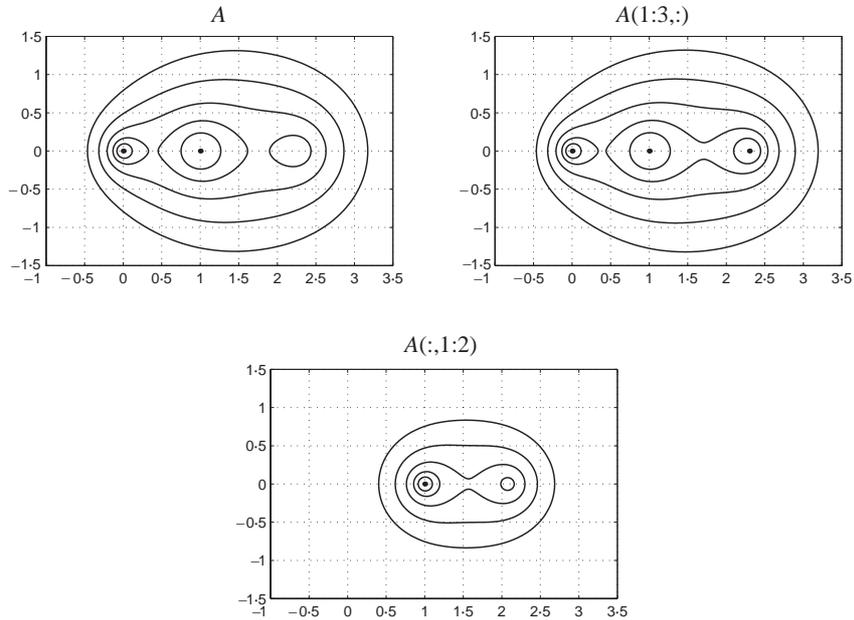


FIG. 1. ϵ -pseudospectra ($\epsilon = 10^{-1}, 10^{-1.25}, \dots, 10^{-2}$) for different portions of the 4×3 matrix (3.1). The sets are nested as predicted by Theorem 2.2.

3. Examples

The 4×3 matrix

$$A = \begin{pmatrix} 1 & 10 & 10 \\ 0 & 2.1 & 4.2 \\ 0 & 0.1 & 0.2 \\ 0 & 0.1 & 0.2 \end{pmatrix} \tag{3.1}$$

has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$, as can be readily verified. (Corresponding eigenvectors are $(10, -2, 1)^T$ and $(1, 0, 0)^T$.)

The top left plot of Fig. 1 shows the pseudospectra of A , with the eigenvalues drawn as dots. Note that although there is a region of the complex plane near $z = 2.3$ in the $10^{-1.75}$ -pseudospectrum, there are no points near $z = 2.3$ in the 10^{-2} -pseudospectrum: this could not happen if A were square, since for a square matrix each connected component of the pseudospectrum contains an eigenvalue. In the top right, we see the pseudospectra of the upper 3×3 square of A , and sure enough, there is now an eigenvalue near $z = 2.3$. Finally, the bottom plot shows the pseudospectra of the first two columns of A , and now there is just the single eigenvalue $\lambda = 1$.

The inclusion properties of Theorem 2.2 can be clearly seen: the pseudospectra of $A(1:3, :)$ are bigger than those of A , whilst the pseudospectra of $A(:, 1:2)$ are smaller.

Figure 2 shows the pseudospectra of A in another manner. We see that the resolvent

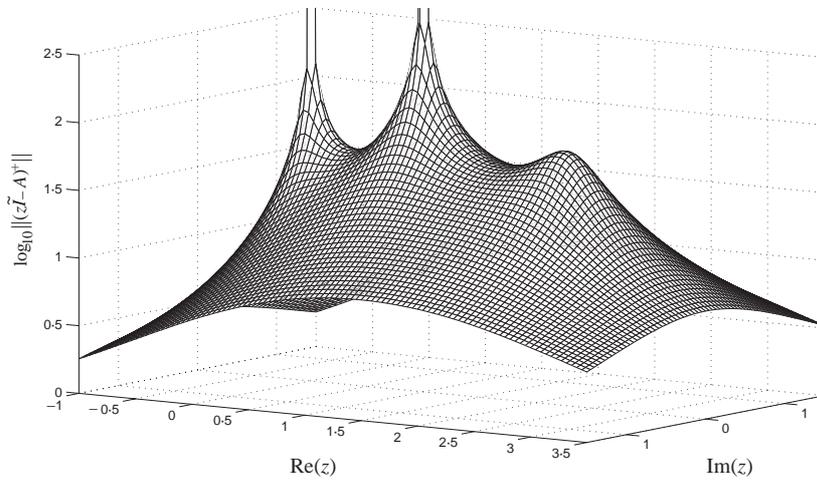


FIG. 2. Surface plot of $\log_{10} \|(z\tilde{I} - A)^+\|$ for the matrix A of (3.1). Whilst poles appear at the eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$, the resolvent norm is bounded near $z = 2.3$.

norm has poles at the eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$, but a local maximum at a pseudoeigenvalue near $z = 2.3$.

The matrix (3.1) is not typical of rectangular matrices, since in general they will have no eigenvalues, only pseudoeigenvalues. Perhaps a more typical example is

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0.1 & 0.2 & 0.2 \end{pmatrix}, \quad (3.2)$$

(shown in Fig. 3) for which any number z near $\pm i$ or 1 is an ϵ -pseudoeigenvalue for a small value of ϵ ; there are no eigenvalues.

4. Algorithms

Having decided that pseudospectra of rectangular matrices may provide useful information, we face the question of how to compute them. For dense, square matrices, methods for the computation of pseudospectra fall broadly into two categories: grid-based methods (see e.g. Trefethen, 1999), and path-following methods (see e.g. Brühl, 1996). Here we will consider only the former.

In recent years, grid-based algorithms for the dense, square case have improved greatly from the basic idea of computing a full SVD of $zI - A$ at each point z in a grid. The current state-of-the-art algorithm for a general dense matrix has two phases (Lui, 1997; Trefethen, 1999). The first is to compute a Schur decomposition of the matrix, $A = UTU^*$, where T is triangular and U unitary, and the second is to compute just the smallest singular value at

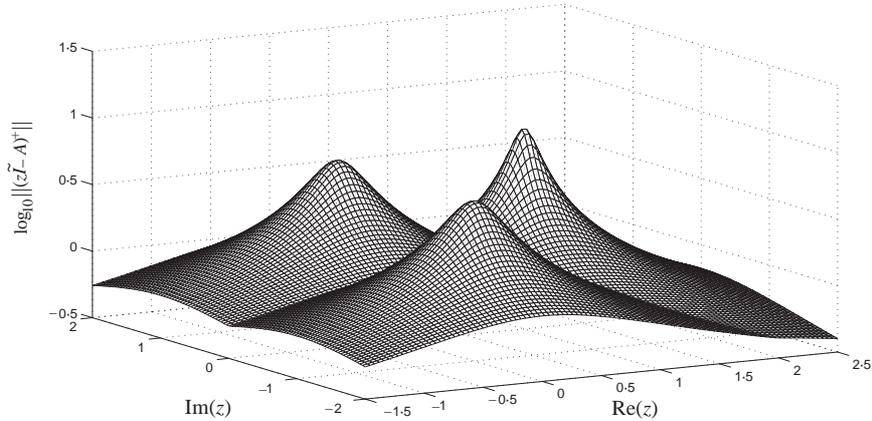


FIG. 3. Like Fig. 2, but for the matrix C of (3.2). $z = 1$ is an ϵ -pseudoeigenvalue for a slightly smaller value of ϵ than $z = \pm i$, but this matrix has no eigenvalues.

each grid point,

$$\sigma_{\min}(zI - A) = \sigma_{\min}(U(zI - T)U^*) = \sigma_{\min}(zI - T). \tag{4.1}$$

This can be done efficiently by computing the smallest eigenvalue of the matrix $(zI - T)^*(zI - T)$ using an inverse Lanczos iteration (i.e. a Lanczos iteration for the largest eigenvalue of $(zI - T)^{-1}(zI - T)^{-*}$), since the systems that have to be solved are triangular.

Ideally, we would like to follow a similar procedure in the case of rectangular matrices, but unfortunately we have not found a generalization of the Schur decomposition that preserves shift-invariance (4.1). However, we can follow this general pattern, performing some computation once and for all at the start, affording us cheaper computations at the grid points than we would otherwise have.

Our current algorithm distinguishes two classes of matrices, those with $m \geq 2n$ and those with $m < 2n$. The overall aim is the same for each: to reduce the problem to the computation of the smallest singular value of a triangular matrix, which, as just stated, can be relatively cheap.

4.1 Matrices with $m \geq 2n$

The first thing to note for matrices $A \in \mathbb{C}^{m \times n}$, $m \geq 2n$ is that although we need to compute $\sigma_{\min}(z\tilde{I} - A)$ at each grid point, only the upper $n \times n$ portion of this matrix changes from point to point:

$$z\tilde{I} - A = \begin{pmatrix} zI_n \\ 0 \end{pmatrix} - \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}.$$

Since singular values are invariant under unitary transformations, we can replace A_2 by QA_2 for any unitary matrix Q . In particular, we can perform a QR factorization $A_2 = QR$, and then

$$\sigma_{\min}(z\tilde{I} - A) = \sigma_{\min} \begin{pmatrix} zI_n - A_1 \\ -A_2 \end{pmatrix} = \sigma_{\min} \begin{pmatrix} zI_n - A_1 \\ -R \end{pmatrix}.$$

If R has any rows of zeros (as it certainly will if $m > 2n$), these will not affect the singular values and can be removed, leaving a matrix S of dimension $2n \times n$ with the following structure:

$$S = \begin{pmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{pmatrix}.$$

The structure of S could be called *trapezoidal*, which is the term we will use for the remainder of this paper.

Even after performing this reduction, we still cannot use S efficiently within an inverse Lanczos iteration, so we now have to do a further reduction *after* shifting for each grid point. Fortunately, this can be done reasonably quickly with a further QR factorization (again dropping rows of zeros, leaving an $n \times n$ triangular matrix), taking advantage of the zeros we have already introduced. We call this a banded QR factorization, since the trapezoidal form of S means that it is banded below the diagonal (with bandwidth n).

The cost of this banded QR factorization is $\mathcal{O}(n^3)$ for each grid point z (in addition to the $\mathcal{O}(n^2)$ inverse Lanczos iteration). This figure is higher than the $\mathcal{O}(n^2)$ one gets for square matrices, but this appears to be the price to be paid for working with rectangular matrices.

4.2 Matrices with $m < 2n$

If the matrix is not of the previous form, the focus of the computation shifts: at the cost of more work in the first phase, we can further reduce the work at each grid point in the second phase. First we note that we no longer have a such a large block of the matrix $z\tilde{I} - A$ that is unchanged at each grid point. In fact, since A_2 now has fewer rows than columns, we cannot transform A to trapezoidal form using the same technique used above. What we can do, however, is split the matrix pencil $z\tilde{I} - A$ differently as follows:

$$z\tilde{I} - A = z \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} - \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},$$

where I_1 and A_1 are $(m - n) \times n$ and I_2 and A_2 are $n \times n$. If we now perform a QZ factorization (see e.g. Golub & Van Loan (1996)) $Q(zI_2 - A_2)Z = zT_2 - S_2$ and update the pencil $zI_1 - A_1$ correspondingly (by post-multiplying by Z), we are left with a pencil of the form

$$zT - S = z \begin{pmatrix} I_1 Z \\ Q I_2 Z \end{pmatrix} - \begin{pmatrix} A_1 Z \\ Q A_2 Z \end{pmatrix} = z \begin{pmatrix} I_1 Z \\ T_2 \end{pmatrix} - \begin{pmatrix} A_1 Z \\ S_2 \end{pmatrix},$$

where S and T are in trapezoidal form of dimension $m \times n$, and Q and Z are unitary and of dimension $n \times n$. This means that $zT - S$ has the same singular values as $z\tilde{I} - A$ for all values

```

procedure RECTPSA( $A$ , grid)
  if  $m \geq 2n$  then
     $[Q, R] = \text{QR}(A(n+1:m, :))$ 
     $S(1:n, :) = A(1:n, :)$ 
     $S(n+1:2n, :) = R(1:n, :)$ 
     $T = \tilde{I}$ 
  else
     $[S_2, T_2, Z] = \text{QZ}(A(m-n+1:m, :), \tilde{I}(m-n+1:m, :))$ 
     $S(1:m-n, :) = A(1:m-n, :)Z$ 
     $S(m-n+1:m, :) = S_2$ 
     $T(1:m-n, :) = Z(1:m-n, :)$ 
     $T(m-n+1:m, :) = T_2$ 
  end if
  for each  $z \in \text{grid}$ ,
     $[Q, R] = \text{BAND\_QR}(zT - S)$ 
     $R = R(1:n, :)$ 
     $\sigma_{\min}(R) = \text{INV\_LANCZOS}(R^*R)$ 
  end for
end procedure

```

FIG. 4. Algorithm for pseudospectra of rectangular matrices.

of z . (Note that updating I_1 and A_1 can be done more efficiently if the transformations used to construct Z are applied as they are created during the QZ algorithm, rather than by explicitly forming and post-multiplying by Z .)

We can now use the banded QR factorization of the trapezoidal matrix $zT - S$ at each grid point z to generate a square, triangular matrix that can then be used with the inverse Lanczos iteration exactly as before.

Although the QZ factorization is much more expensive than the QR factorization used when $m \geq 2n$ (by a factor of approximately 20), we only have to do this once, and now the bandwidth of the pencil $zT - S$ is only $m - n$, meaning that the QR factorization at each grid point involves $\mathcal{O}((m - n)n^2)$ operations instead of $\mathcal{O}(n^3)$. In other words, the closer to square the matrix A is, the cheaper the computation is at each grid point.

If A is square, the QZ factorization reduces to the Schur decomposition, and we recover the standard algorithm for square matrices. (In this case the lower bandwidth is 0, so we do not need any banded QR step at each grid point.)

Our complete algorithm for rectangular matrices is given in Fig. 4.

5. Applications

5.1 Arnoldi iteration

Rectangular matrix pseudospectra were first introduced to handle the rectangular Hessenberg matrices generated by the Arnoldi iteration applied to a non-Hermitian matrix A (Arnoldi, 1951; Saad, 1980). Interest in these matrices arose from their inclusion

properties (Theorem 2.2). Toh & Trefethen (1996) showed that the pseudospectra of the $(k+1) \times k$ matrix \tilde{H}_k from the k th step of the Arnoldi process,

$$AV_k = V_k H_k + f_k e_k^* = V_{k+1} \tilde{H}_k,$$

are contained within those of A , and often provide good approximations. This is very different to the familiar situation with the Ritz values of A (the eigenvalues of the $k \times k$ square matrix H_k), which in general can lie far from the true eigenvalues of A .

Although the result of Toh & Trefethen was of theoretical interest, in practice large subspace dimensions k are often needed before good approximations to the pseudospectra are obtained. Wright & Trefethen (2001) showed that when the matrices \tilde{H}_k are computed using the implicitly restarted Arnoldi iteration (Sorensen, 1992; Lehoucq *et al.*, 1998) however, accurate approximations to the pseudospectra can be obtained for much smaller subspace dimensions. In particular, pseudospectra approximations can be obtained for a very low relative cost once the eigenvalue computation is complete. The algorithm described in that paper uses a Hessenberg QR factorization (see e.g. Golub & Van Loan, 1996) at each gridpoint and is a special case of the algorithm described here in Section 4; no initial reduction to trapezoidal form is necessary since \tilde{H}_k already has the necessary structure.

5.2 Block Arnoldi iteration

In practice, the computer memory hierarchy sometimes makes it advantageous to work on matrices rather than just vectors, and a block variant of the Arnoldi iteration can be used for this purpose (Saad, 1992). The idea is similar to the usual Arnoldi iteration, except that instead of starting with a single vector, v_0 , one starts with a block of j orthonormal starting vectors, $V_0 = [v_0, v_1, \dots, v_{j-1}]$. After the k th step of the iteration, one has a block-Hessenberg matrix

$$\tilde{H}_k = \begin{pmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,k} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,k} \\ & \ddots & \ddots & \vdots \\ & & H_{k,k-1} & H_{k,k} \\ & & & H_{k+1,k} \end{pmatrix},$$

of dimension $j(k+1) \times jk$, where $H_{i,l}$ is dense for $i \leq l \leq k$ and $H_{i,i-1}$ is upper triangular for $2 \leq i \leq k+1$. The upshot of this is that \tilde{H}_k is already in trapezoidal form, so we can compute its pseudospectra without any phase one pre-processing. What is more, since \tilde{H}_k is a rectangular section of the final banded Hessenberg matrix H , which is similar to A , the nesting property holds as in the case of the standard Arnoldi iteration:

$$\Lambda_\epsilon(\tilde{H}_k) \subseteq \Lambda_\epsilon(H) = \Lambda_\epsilon(A).$$

Restarting techniques (including implicit restarting (Sorensen, 1992)) that result in matrices which are left in a block-Arnoldi factorization will also yield matrices that have this property.

Small block sizes are typically used in practice (e.g. $j = 2$ to $j = 6$), which means that the bandwidth of \tilde{H}_k is small, and the pseudospectra can be computed efficiently.

5.3 Bounds for Lanczos iterations

Whilst one would not usually consider using pseudospectra to say something about the behaviour of normal matrices, in this section we show that they can lead to useful eigenvalue bounds. The Lanczos factorization for a symmetric matrix A can be written as

$$AQ_k = Q_k T_k + \beta_k q_{k+1} e_k^* = Q_{k+1} \tilde{T}_k,$$

where Q_k is a matrix with orthonormal columns and T_k is tridiagonal. This is a special case of the Arnoldi factorization, and so, as before, $\Lambda_\epsilon(\tilde{T}_k) \subseteq \Lambda_\epsilon(A)$.

The eigenvalues θ_i of T_k (the Ritz values) are conventionally taken as approximations to the eigenvalues of A , and the following bound is often used as a test of the quality of the approximation (see e.g. Bai *et al.*, 2000, Golub & Van Loan, 1996):

$$\min_{\lambda \in \Lambda(A)} |\theta_i - \lambda| \leq \beta_k |(s_i)_k| = \|Ay_i - \theta_i y_i\|_2, \quad i = 1, 2, \dots, k. \quad (5.1)$$

Here (θ_i, s_i) is an eigenpair of T_k and $y_i = Q_k s_i$ is the corresponding Ritz vector of A .

Recall that the ϵ -pseudospectra of A consist simply of balls of radius ϵ centred at the eigenvalues, or equivalently, if $z \in \Lambda_\epsilon(A)$, there must be an eigenvalue within a distance ϵ of z ,

$$\min_{\lambda \in \Lambda(A)} |z - \lambda| \leq \epsilon. \quad (5.2)$$

Since $\Lambda_\epsilon(\tilde{T}_k) \subseteq \Lambda_\epsilon(A)$, there must also be an eigenvalue of A within a distance ϵ of any point $z \in \Lambda_\epsilon(\tilde{T}_k)$. Another way to put this is that the distance of the Ritz values θ_i to the true eigenvalues of A can be bounded in terms of the norm of the resolvent:

$$\min_{\lambda \in \Lambda(A)} |\theta_i - \lambda| \leq \|(\theta_i \tilde{I} - \tilde{T}_k)^+\|_2^{-1}. \quad (5.3)$$

What is more, we have the following theorem.

THEOREM 5.1 The pseudospectra bound (5.3) is always at least as good as the bound (5.1):

$$\min_{\lambda \in \Lambda(A)} |\theta_i - \lambda| \leq \|(\theta_i \tilde{I} - \tilde{T}_k)^+\|_2^{-1} \leq \|Ay_i - \theta_i y_i\|_2.$$

Proof. We have just proved the first inequality. The second can be proved as follows:

$$\begin{aligned} \|(\theta_i \tilde{I} - \tilde{T}_k)^+\|_2^{-1} &= \sigma_{\min}(\theta_i \tilde{I} - \tilde{T}_k) \\ &\leq \|(\theta_i \tilde{I} - \tilde{T}_k) s_i\|_2 \\ &= \|(\theta_i Q_{k+1}^* Q_k - Q_{k+1}^* A Q_k) s_i\|_2 \\ &= \|Q_{k+1}^* (\theta_i I_n - A) y_i\|_2 \\ &\leq \|Q_{k+1}^*\|_2 \|(\theta_i I_n - A) y_i\|_2 \\ &= \|\theta_i y_i - A y_i\|_2. \end{aligned}$$

□

In fact, this bound is not new. Lehmann (1949) and Lehmann (1966) (see also Parlett (1998, Section 10.5)) describe more general bounds for the eigenvalues of A in terms of not only the smallest singular value, but also the second smallest and so on. What is new here is the formulation in terms of pseudospectra, and as a result an efficient algorithm for computing the bounds. Beattie & Ipsen (2001) extend Lehmann's bounds to non-Hermitian matrices, but since these new bounds depend on the condition number of the eigenvector matrix, it is difficult to use them in practice.

Another closely related topic is that of *refined Ritz vectors* described by Jia (1997) for non-Hermitian matrices (see also Jia & Stewart, 2001). The idea here is that whilst Ritz vectors may not converge to eigenvectors of A or may converge slowly, refined Ritz vectors corresponding to $Q_k u_n$, where u_n is the right singular vector corresponding to the smallest singular value of $zI - \tilde{T}_k$, can do better. The refined Ritz vector is a pseudoeigenvector of A for a smaller value of ϵ than the Ritz vector, since the refined Ritz vector is optimal for the particular Ritz value.

5.4 Pseudospectra lower bounds for some matrices

For certain matrices, it is possible to approximate the pseudospectra by the pseudospectra of the first few columns of the matrix, which we know from Theorem 2.2 are lower bounds. This technique may work for example when the eigenvalues of a leading submatrix A_{11} of dimension $k \ll n$ are well separated and the $(n - k) \times k$ block A_{21} below this has small norm. In such a case, the eigenvalues of A_{11} will be ϵ -pseudoeigenvalues of the first k columns of A for all $\epsilon \geq \|A_{21}\|$, and one might hope that for $\epsilon \gg \|A_{21}\|$ one will get good approximations to the pseudospectra of A near the eigenvalues of A_{11} .

Figure 5 shows the results of this technique applied to a random matrix A of dimension 500 with the k th row and column scaled by $1/(1 + 5(k - 1))$ for each k : that is, a_{ij} is a random sample from the normal distribution of mean 0 and standard deviation $[(5i - 4)(5j - 4)]^{-1/2}$. The plot on the left shows the true pseudospectra near some of the eigenvalues, and took nearly 12 minutes to compute on a Sun Ultra 5 workstation, whilst the right plot shows the lower bound obtained by taking the first 30 columns, and took less than 20 seconds to compute (on the same grid). The lower block has norm $\|A_{21}\| \approx 10^{-1.5}$, so we hope to see accurate epsilon contours down to approximately that value. In fact, we do better than that.

Of course, this technique relies on a small portion of the top left of the matrix approximating the eigenvalues and pseudospectra we are interested in, which will probably not be the case in general. More often, use of the Arnoldi iteration provides a more reliable way to approximate pseudospectra whilst still maintaining inclusion.

5.5 Control theory

An important problem in control theory is that of estimating the distance from a linear control system

$$\dot{x} = Ax + Bu$$

(see e.g. Kailath (1980)) to the nearest uncontrollable system (a system which cannot be controlled from any given starting state to any given finishing state). In the notation of this

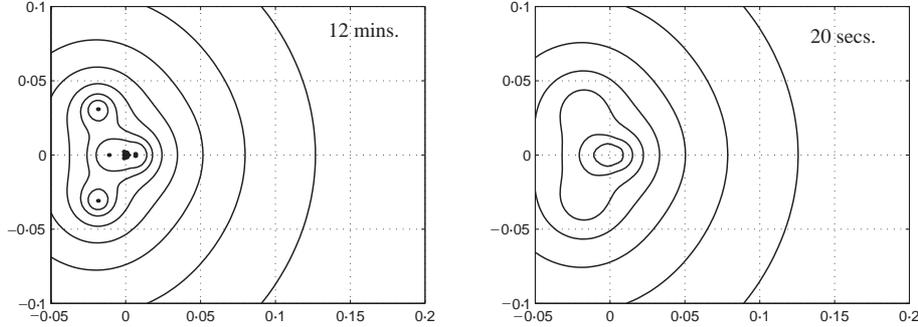


FIG. 5. Exact ϵ -pseudospectra ($\epsilon = 10^{-1}, 10^{-1.25}, \dots, 10^{-2.5}$) for a random matrix (left), and the lower bound obtained from the first 30 columns (right).

paper (which differs from that in control theory), A is $n \times n$, B is $n \times (m - n)$, x is an n -vector of states and u an $(m - n)$ -vector of control inputs. Eising (1984) shows that this distance is equal to the smallest singular value (minimized over all values of λ) of the $m \times n$ matrix

$$\begin{pmatrix} \lambda I - A^T \\ B^T \end{pmatrix},$$

which as pointed out in Higham & Tisseur (2002), is equal to

$$\min_{\epsilon} \left\{ \mathcal{A}_{\epsilon} \left(\begin{pmatrix} \lambda I - A^T \\ B^T \end{pmatrix} \right) \neq \emptyset \right\}. \tag{5.4}$$

In other words, the maximum value of the resolvent norm of this matrix is the inverse of the minimum distance to an uncontrollable system. Thus in our terms the problem of determining distance to uncontrollability becomes the problem of finding the maximum of the resolvent norm surface.

Various algorithms for computing this distance have appeared in the literature such as an early contribution by Boley (1987), and a notable recent contribution by Gu (2000). The algorithms Gu presents are iterative, and he observes that if the system is close to being uncontrollable, the number of iterations can increase. This increase in computation time is also seen by Byers (1990) in his ‘BFI’ method, which is similar to ours. He computes $h(\lambda) = \sigma_{\min}([A - \lambda I_n, B])$ for many different values of λ by first reducing A to upper Hessenberg form $H = Q^* A Q$, and then for each λ reducing $[H - \lambda I_n, Q^* B]$ to upper trapezoidal form, $[R, 0]$ where R is triangular, again using unitary transformations. He then estimates $\sigma_{\min}(R)$ using a condition number estimator (see e.g. Cline *et al.*, 1979), doing this for different values of λ defined in a grid. This approach proves too costly for systems which are close to being uncontrollable, and he also describes a second algorithm which reduces the necessary number of evaluations of $h(\lambda)$.

To give an indication of how our approach could be used in practice, consider the matrix C of (3.2). The distance to uncontrollability is the inverse of the maximum value of the resolvent norm, so if C represents a control system (where A is 3×3 and B is 3×1),

from a rotation of Fig. 3 it appears that the distance to uncontrollability is about $10^{-1} = 0.1$ near $z = 1$. In fact, refining the grid close to this point indicates that the true distance is approximately $10^{-1.009} = 0.0979$. Perhaps a hybrid method which gets an initial guess from the pseudospectra grid data and then uses one of the iterative methods to refine this estimate might prove useful.

5.6 Game theory

Game theory also leads to rectangular eigenvalue problems, as a two-player game depending on a given parameter λ can be represented by an $m \times n$ pencil $A - \lambda B = G_\lambda$, whose rows represent options for Player 1 and whose columns represent options for Player 2. The (i, j) entry of the pencil denotes the amount of money Player 1 receives from Player 2 if Player 1 chooses the i th option and Player 2 chooses the j th option (if this entry is negative, Player 2 receives the money from Player 1). The value of a game, $\phi(G_\lambda) = \min_x \max_y (y^T G_\lambda x) = \max_y \min_x (y^T G_\lambda x)$ (where x and y are real, non-negative 1-norm normalized vectors), is defined to be the maximum amount Player 1 can expect to win (unless Player 1 knows extra information about the strategy of Player 2), and this is equal to the minimum amount that Player 2 can expect to lose (see e.g. Karlin (1992)).

Thompson & Weil (1972) show that if the pencil has an eigenvalue λ corresponding to non-negative left and right eigenvectors, the value of the game for this value of λ is zero. In other words, by judicious play, both players can expect to come out at least even, no matter what the other player does; the game is fair. Theorem 4.2 of Thompson & Weil can be modified for the case of ϵ -pseudoeigenvalues instead of exact eigenvalues as follows. The pseudospectra here are defined by viewing $A - zB$ as a map from $\mathbb{C}_{1\text{-norm}}^n$ to $\mathbb{C}_{\infty\text{-norm}}^m$,

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C} : \|(A - zB)v\|_\infty \leq \epsilon \text{ for some } v \in \mathbb{C}^n, \|v\|_1 = 1\},$$

rather than the 2-norm pseudospectra discussed so far. However, since we have $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ for a vector x , the result can still be used for 2-norm pseudospectra, although it may be less sharp.

THEOREM 5.2 If λ is an ϵ -pseudoeigenvalue (in the above sense) of the $m \times n$ pencil $A - \lambda B = G_\lambda$ with a corresponding right pseudoeigenvector x_* that is real and non-negative, then the value of the game G_λ satisfies $\phi(G_\lambda) \leq \epsilon$.

Proof. By Hölder's inequality, we have $y^T G_\lambda x \leq \|y\|_1 \|G_\lambda x\|_\infty = \|G_\lambda x\|_\infty$ for any normalised vector y , and since x_* is an ϵ -pseudoeigenvector, $\|G_\lambda x_*\|_\infty \leq \epsilon$, which gives

$$\begin{aligned} \phi(G_\lambda) &= \max_y \min_x (y^T G_\lambda x) \\ &\leq \max_y (y^T G_\lambda x_*) \\ &\leq \epsilon. \end{aligned}$$

□

As an example, consider the matrix A of (3.1) as a parametrized game $A - \lambda \tilde{T}$ (here we take $B = \tilde{T}$). Player 1 chooses a number i from 1 to 4, and simultaneously Player 2 chooses

a number j from 1 to 3. Now, as seen in Section 3, $z = 2.3$ is an ϵ -pseudoeigenvalue for A with $\epsilon = 0.0135$ in the 2-norm. Since the right singular vector of $A - \lambda \tilde{I}$ corresponding to the smallest singular value happens to be non-negative, we can apply Theorem 5.2 using this singular vector as our ϵ -pseudoeigenvector. This tells us that for $\lambda = 2.3$, the value of this game is less than 0.0135.

The 1-norm normalized ϵ -pseudoeigenvector gives probabilities for the numbers Player 2 should choose to ensure this. For example, in this game, Player 1 can never lose with $i = 4$ since all the entries of row 4 of the matrix are non-negative. However, as predicted by the theorem, if Player 2 uses probabilities given by the pseudoeigenvector, they lose less than $\epsilon = 0.0135$ in the long run.

6. Generalizations

6.1 Generalized rectangular eigenproblems

Some applications (such as game theory) involve rectangular matrix pencils,

$$A - \lambda B,$$

a generalization of the pencil $A - \lambda \tilde{I}$ we have mainly focused on up to now. For related work see Lavallée (1997). (Higham & Tisseur (2002) consider the still more general problem of rectangular polynomial eigenproblems and their application to control theory.) The analogue of (1.2) for this case is that for (λ, v) to be an eigenpair of the rectangular pencil, it must be an eigenpair of the square pencil $A_1 - \lambda B_1$, and v must be in the nullspace of $A_2 - \lambda B_2$.

The literature contains several different definitions of pseudospectra for square generalized eigenproblems: Riedel (1994) considers the case where B is self-adjoint and positive definite, van Dorsselaer (1997) only considers perturbations to A (not B), Frayssé *et al.* (1996) allow different sized perturbations to the matrices A and B , whilst Ruhe (1994) considers $\Lambda_\epsilon(B^{-1}A)$. Here, for simplicity, we generalize the definitions used by van Dorsselaer, since including perturbations to B in addition to those of A requires only minor changes[†]:

$$\begin{aligned} \Lambda_\epsilon(A, B) &= \{z \in \Lambda(A + E, B) : \|E\| \leq \epsilon\} \\ &= \{z \in \mathbb{C} : \|(zB - A)^+\| \geq \epsilon^{-1}\}. \end{aligned} \tag{6.1}$$

It turns out to be relatively straightforward to include generalized rectangular eigenproblems into our current algorithm, particularly in the case where $m < 2n$. Here, since the QZ factorization is designed for generalized eigenproblems, we do not need to make any changes at all. On the other hand, if $m \geq 2n$, we have to be more careful. We can no longer perform a QR factorization of the lower $(m - n) \times n$ block of A since the difference $zB - A$ will be dense in general. However, if we first perform a QR factorization $QR = B$ and update A to reflect this transformation $\hat{A} = Q^*A$ (thus keeping the singular values of the pencil unchanged), we now have matrices of the right form for the algorithm

[†]Suppose that the perturbations to A are bounded by $\alpha\epsilon$ and those to B by $\beta\epsilon$. Then $\Lambda_\epsilon(A) = \{z \in \mathbb{C} : \|(A - zB)v\| \leq \epsilon(\alpha + \beta|z|) \text{ for some } v \in \mathbb{C}^n, \|v\| = 1\}$ (cf. (2.1)), as proved by Higham & Tisseur (2002) in the more general case of rectangular matrix polynomials.

presented previously, since only the upper $n \times n$ block \widehat{A}_1 of \widehat{A} will be changed when subtracting from zB :

$$zQ^*B - Q^*A = z \begin{pmatrix} \times & \times & \times \\ & \times & \times \\ & & \times \\ \hline & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} - \begin{pmatrix} \widehat{A}_1 \\ \hline \widehat{A}_2 \end{pmatrix}.$$

A very similar algorithm is described by Spellucci & Hartmann (2000) for solving the least squares problem of minimising $\|(A + \lambda B)x - f\|_2$ for several values of λ . They consider $m \gg n$, and their algorithm reduces A and B to staircase triangular form (each column has two fewer zero entries than the previous one). They then use Givens rotations to reduce this pencil to triangular form (the step we do with banded QR, based on Householder reflectors), since this can take advantage of existing sparsity in their problem.

6.2 Computing pseudospectra in weighted norms

Another generalization we may consider is how to compute pseudospectra in norms other than the 2-norm; this arises, for example, when an ‘energy’ norm is the correct way to measure perturbations to the matrix. We currently have no way of including this directly into our algorithm for general rectangular matrices, but in the special case where the rectangular matrices are derived from an orthogonal projection (for example Arnoldi’s method or the JDQR algorithm (Sleijpen & Van der Vorst, 1996)) there is more we can say. The norms we consider here are of the form $\|x\|_G = \|Gx\|_2$ for nonsingular G , which induce the matrix norm $\|A\|_G = \|GAG^{-1}\|_2$ (for square matrices, A), and have the associated inner product $\langle x, y \rangle_G = \langle Gx, Gy \rangle_2 = x^*G^*Gy$.

The problem (using Arnoldi’s method for illustration) is that when we come to compute the pseudospectra of the projected matrix \widetilde{H}_k , all our algorithms use the 2-norm, not the energy norm appropriate for the problem. The solution is to compute the orthonormal basis for the Krylov subspace in Arnoldi’s method not in the 2-norm, but in the required energy norm. The reason for this is as follows: suppose you build up a full Arnoldi factorization $AV = VH$, where V has n orthonormal columns (in the weighted inner product, $\langle \cdot, \cdot \rangle_G$) or in other words

$$V^*G^*GV = I = GVV^*G^*. \tag{6.2}$$

Now, to compute the G -weighted pseudospectra of A , we can either evaluate $\sigma_{\min}(G(zI - A)G^{-1}) = \sigma_{\min}(zI - GAG^{-1})$, or, since $GAG^{-1} = GVVH^*G^*$,

$$\begin{aligned} \sigma_{\min}(zI - GAG^{-1}) &= \sigma_{\min}(zI - GVVH^*G^*) \\ &= \sigma_{\min}(zI - H). \end{aligned} \tag{6.3}$$

This comes from the fact that GV and V^*G^* have 2-norm orthonormal columns (6.2), and so do not change the singular values.

What (6.3) shows is that the G -weighted pseudospectra of A are the same as the 2-norm pseudospectra of H , if we have used the G -norm for the orthogonalization. In a practical Arnoldi iteration, we do not actually compute all of the matrix H (we only have the upper left $(k + 1) \times k$ section, \tilde{H}_k), but we know that the 2-norm pseudospectra of \tilde{H}_k are contained within those of H , and so they must be contained within those of A by (6.3).

7. Summary

We have demonstrated that pseudospectra of rectangular matrices can provide useful information in a variety of fields. Using our algorithm, which is implemented in the Pseudospectra GUI (Wright, 2001), one can compute pseudospectra for $m \times n$ ($m \geq n$) matrices with n of the order of 500 to 1000 in a few minutes on a Sun Ultra 5 workstation. The cost of the algorithm is effectively independent of m , apart from the initial QR factorization needed when $m \geq 2n$.

Acknowledgements

We would like to thank Mark Embree, Nick Higham and Françoise Tisseur for their suggestions and insights which have greatly improved this paper. We are also grateful to Dan Boley and Paul Van Dooren for their advice about control theory.

REFERENCES

- ARNOLDI, W. E. (1951) The principle of minimized iteration in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, **9**, 17–29.
- BAI, Z., DEMMEL, J., DONGARRA, J., RUHE, A. & VAN DER VORST, H. (eds) (2000) *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. Philadelphia, PA: SIAM.
- BEATTIE, C. & IPSEN, I. C. (2001) Inclusion regions for matrix eigenvalues. *Linear Algebra Appl.*, (to appear).
- BOLEY, D. (1987) Computing rank-deficiency of rectangular matrix pencils. *Systems Control Lett.*, **9**, 207–214.
- BOLEY, D. (1990) Estimating the sensitivity of the algebraic structure of pencils with simple eigenvalue estimates. *SIAM J. Matrix Anal. Appl.*, **11**, 632–643.
- BRÜHL, M. (1996) A curve tracing algorithm for computing the pseudospectrum. *BIT*, **36**, 441–454.
- BYERS, R. (1990) Detecting nearly uncontrollable pairs. *Signal Processing, Scattering and Operator Theory, and Numerical Methods (Amsterdam, 1989)*. Boston, MA: Birkhäuser, pp. 447–457.
- CLINE, A. K., MOLER, C. B., STEWART, G. W. & WILKINSON, J. H. (1979) An estimate for the condition number of a matrix. *SIAM J. Numer. Anal.*, **16**, 368–375.
- EISING, R. (1984) Between controllable and uncontrollable. *Systems Control Lett.*, **4**, 263–264.
- EMBREE, M. & TREFETHEN, L. N. (2001) Generalizing eigenvalue theorems to pseudospectra theorems. *SIAM J. Sci. Comput.*, **23**, 583–590.
- FRAYSSÉ, V., GUEURY, M., NICOU, F. & TOUMAZOU, V. (1996) Spectral portraits for matrix pencils. Technical Report TR/PA/96/19, CERFACS.

- GOLUB, G. H. & VAN LOAN, C. F. (1996) *Matrix Computations*, 3rd edn. Baltimore, MD: Johns Hopkins University Press.
- GU, M. (2000) New methods for estimating the distance to uncontrollability. *SIAM J. Matrix Anal. Appl.*, **21**, 989–1003.
- HIGHAM, N. & TISSEUR, F. (2002) More on pseudospectra for polynomial eigenvalue problems and applications in control theory. *Linear Algebra Appl.*, (to appear).
- JIA, Z. (1997) Refined iterative algorithms based on Arnoldi's process for large unsymmetric eigenproblems. *Linear Algebra Appl.*, **259**, 1–23.
- JIA, Z. & STEWART, G. W. (2001) An analysis of the Rayleigh–Ritz method for approximating eigenspaces. *Math. Comp.*, **70**, 637–647.
- KAILATH, T. (1980) *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall.
- KARLIN, S. (1992) *Mathematical methods and theory in games, programming, and economics*. Vol. I: *Matrix Games, Programming, and Mathematical Economics*, Vol. II: *The Theory of Infinite Games*. New York: Dover Publications, Reprint of the 1959 original.
- LAVALLÉE, P.-F. (1997) *Nouvelles Approches de Calcul du ϵ -Spectre de Matrices et de Faisceaux de Matrices*, Ph.D. Thesis, L'Université de Rennes 1, Rennes, France.
- LEHMANN, N. J. (1949) Beiträge zur numerischen Lösung linearer Eigenwertprobleme. I. *Z. Angew. Math. Mech.*, **29**, 341–356.
- LEHMANN, N. J. (1966) Zur Verwendung optimaler Eigenwerteingrenzungen bei der Lösung symmetrischer Matrizenaufgaben. *Numer. Math.*, **8**, 42–55.
- LEHOUCQ, R. B., SORENSEN, D. C. & YANG, C. (1998) *ARPACK Users' Guide*. Philadelphia PA: SIAM.
- LUI, S. H. (1997) Computation of pseudospectra by continuation. *SIAM J. Sci. Comput.*, **18**, 565–573.
- PARLETT, B. N. (1998) *The Symmetric Eigenvalue Problem*. Philadelphia, PA: SIAM, Corrected reprint of the 1980 original.
- RIEDEL, K. S. & (1994) Generalized epsilon-pseudospectra. *SIAM J. Numer. Anal.*, **31**, 1219–1225.
- RUHE, A. (1994) The rational Krylov algorithm for large nonsymmetric eigen-values—mapping the resolvent norms (pseudospectrum). Unpublished note.
- SAAD, Y. (1980) Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra Appl.*, **34**, 269–295.
- SAAD, Y. (1992) *Numerical Methods for Large Eigenvalue Problems*. Manchester: Manchester University Press.
- SLEIJPEN, G. L. G. & VAN DER VORST, H. A. (1996) A Jacobi–Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, **17**, 401–425.
- SOERENSEN, D. C. (1992) Implicit application of polynomial filters in a k -step Arnoldi method. *SIAM J. Matrix Anal. Appl.*, **13**, 357–385.
- SPELLUCCI, P. & HARTMANN, W. M. (2000) A QR decomposition for matrix pencils. *BIT*, **40**, 183–189.
- STEWART, G. W. (1994) Perturbation theory for rectangular matrix pencils. *Linear Algebra Appl.*, **208/209**, 297–301.
- STEWART, G. W. & SUN, J.-G. (1990) *Matrix Perturbation Theory*. Boston, MA: Academic Press.
- THOMPSON, G. L. & WEIL, R. L. (1972) The roots of matrix pencils ($Ay = \lambda By$): Existence, calculations, and relations to game theory. *Linear Algebra Appl.*, **5**, 207–226.
- TOH, K.-C. & TREFETHEN, L. N. (1996) Calculation of pseudospectra by the Arnoldi iteration. *SIAM J. Sci. Comput.*, **17**, 1–15.

- TREFETHEN, L. N. (1999) Computation of pseudospectra. *Acta Numerica 1999*. Cambridge: Cambridge University Press, pp. 247–295.
- VAN DORSSELAER, J. L. M. (1997) Pseudospectra for matrix pencils and stability of equilibria. *BIT*, **37**, 833–845.
- WRIGHT, T. G. (2001) Matlab Pseudospectra GUI. <http://www.comlab.ox.ac.uk/pseudospectra/psagui/>.
- WRIGHT, T. G. & TREFETHEN, L. N. (2001) Computation of pseudospectra using ARPACK and eigs. *SIAM J. Sci. Comput.*, **23**, 591–605.