Properties and Computations of Matrix Pseudospectra

Yuming Shen[1,2], *Jinxi Zhao*[1] and *Hongjun Fan*[1,3]
1.State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, P.R.China
2.Department of Mathematics, Guangxi Normal University, Guilin 541004, P.R.China
3.Department of Mathematics, Nanjing University, Nanjing 210008, P.R.China

Abstract

Pseudospectra were introduced as early as 1975 and became popular tool during the 1990s.In this paper, we give a new definition of pseudospectra by using QR decomposition. Some properties of pseduospectra are explored and an algorithm for the computation of pseduospectra is given.
Key words and phrases: eigenvalues, pseudospectra,QR decomposition

1 Introduction

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

$$(A - \lambda \tilde{I})\nu = 0, \tag{1.1}$$

for some nonzero n-vector ν, where $\tilde{I}$ denotes the $m \times n$ 'identity' with 1 on the main diagonal and 0 elsewhere. If (λ, ν) satisfies (1.1), then we have

$$\begin{pmatrix} A_1 - \lambda I_n \\ A_2 \end{pmatrix} \nu = 0,$$

where A_1 denotes the $n \times n$ upper part of A. Hence not only (λ, ν) must be an eigenpair of A_1, but ν must also be in the nullspace of A_2. Obviously, if A is a square matrix, then we get the canonical definition of eigenvalue.

Four equivalent definitions of pseudospectra of square matrix were introduced by [1, 7, 8, 9]. Pseudospectra of rectangular matrix has been considered by Toh,Wright and Trefethen[6, 11, 12] , Higham and Tisseur[2]. Here we present these equivalent definitions of pseudospectra as follows[12],

Definition 1.1 *Let $A \in \mathcal{C}^{m\times n}$ and $\epsilon \geq 0$ be arbitrary. The $\epsilon-$pseudospectrum $\Lambda_\epsilon(A)$ of A is the set of $z \in \mathcal{C}$ such that*

$$\|(z\tilde{I} - A)^\dagger\| \geq \epsilon^{-1}, \tag{1.2}$$

where '$\dagger$' denotes the pseudoinverse and $\tilde{I}$ denotes the $m\times n$ identity with 1 on the main diagonal and 0 elsewhere, $\mathcal{C}$ denotes the complex plane.

Definition 1.2 $\Lambda_\epsilon(A)$ *is the set of* $z \in \mathcal{C}$ *such that* $z \in \Lambda(A+E)$ *for some* $E \in \mathcal{C}^{m\times n}$ *with* $\|E\| \leq \epsilon$.

Definition 1.3 $\Lambda_\epsilon(A)$ *is the set of* $z \in \mathcal{C}$ *such that* $\|(z\tilde{I} - A)\nu\| \leq \epsilon$ *for some* $\nu \in \mathcal{C}^n$ *with* $\|\nu\| = 1$.

Definition 1.4 (assuming that the norm is $\|.\|_2$) $\Lambda_\epsilon(A)$ *is the set of* $z \in \mathcal{C}$ *such that*

$$\sigma_{\min}(z\tilde{I} - A) \leq \epsilon, \tag{1.3}$$

where $\sigma_{\min}$ denotes the smallest singular value.

In section 2 we give a new definition of pseudospectra. In section 3 we consider some fundamental properties of this new definition.In section4 we present some numerical examples to examine our conclusions. For simplicity, our norm $\|.\|$ will always be the vector 2-norm.

2 A new definition of matrices pseudospectra

Let $B = z\tilde{I} - A = [b_1, b_2, ..., b_n]$. It is shown that a system of vectors $\{b_1, b_2, ..., b_n\}$ is dependence if and only if $G[b_1, b_2, ..., b_n] = 0$, where $G[b_1, b_2, ..., b_n]$ is Gram determinant, i.e., $G[b_1, b_2, ..., b_n] \equiv det(B^*B)$. We can see that if $z \in \Lambda(A)$ is an eigenvalue of A then we must have $det^{\frac{1}{2}}(B^*B) = 0$. Based on this consideration we give another definition of pseudospectra.

On the other hand , let A be an $m \times n$ matrix with $m \geq n$, we write A as follows,

$$A = [a_1, a_2, ..., a_n]. \tag{2.1}$$

A system of vectors $\{a_1, a_2, ..., a_k\}, 1 \leq k \leq n$ is $\epsilon-$linear dependence, if $G^{\frac{1}{2}}[a_1, a_2, ..., a_k] \leq \epsilon$ for any given $\epsilon \geq 0$[3]. Obviously, if a system of vectors$\{a_1, a_2, ..., a_k\}$is $\epsilon-$linear dependence then a system of vectors$\{a_1, a_2, ..., a_r\}$ with $r > k$ is also $\epsilon-$linear dependence. And we can have the following result[4].

Suppose $\{b_1, b_2, ..., b_k\}$ is an orthogonal system and $\|b_i\| = \|a_i\|, i = 1, 2, ..., k$ then

$$G[a_1, a_2, ..., a_r] \leq G[b_1, b_2, ..., b_k] \tag{2.2}$$

. The equality is satisfied if and only if $\{a_1, a_2, ..., a_k\}$ is also an orthogonal system.

Based on this consideration we give a new definition of pseudospectra.

Definition 2.1 *Let* $A \in \mathcal{C}^{m\times n}$ *and* $\epsilon \geq 0$ *be arbitrary.The* $\epsilon-$*pseudospectrum* $\overline{\Lambda_\epsilon}(A)$ *of* A *is the set of* $z \in \mathcal{C}$ *such that*

$$\overline{\Lambda_\epsilon}(A) = \{z \in \mathcal{C} : G^{\frac{1}{2}}(z\tilde{I} - A) = G^{\frac{1}{2}}[b_1, b_2, ..., b_n] \leq \epsilon\} \tag{2.3}$$

As we will show, $\overline{\Lambda_\epsilon}(A)$ depends continuously on A(for $\epsilon > 0$) and is nonempty for sufficiently large ϵ.

3 Basic properties

Theorem 3.1 *Let A be an $m \times n$ matrix,$B = z\tilde{I} - A = QR$.Then*
(i) $\Lambda(A) \subseteq \overline{\Lambda_\epsilon}(A)$ *,where $\Lambda(A)$ denotes the set of eigenvalues of A.*
(ii) $\overline{\Lambda}_{\epsilon\alpha^{-1}}(A) = \alpha\overline{\Lambda_\epsilon}(A)$ *for any $\alpha > 0$.*
(iii) $\overline{\Lambda}_\epsilon(A(:, 1:k)) \subseteq \overline{\Lambda}_{\epsilon|\rho_{k+1,k+1}|}(A(:, 1:k+1)), 1 \leq k < n$,
where the monotonicity result is expressed in 'MATLAB notation',QR denotes QR decomposition and ρ_{kk} is the main diagonal elements of matrix R.

Proof. (i) For any $z \in \Lambda(A)$, we obtain that a system of vectors $\{b_1, ..., b_n\}$ is linear dependence, i.e.,$G[b_1, b_2..., b_n] = 0$. which yields, $z \in \overline{\Lambda}_\epsilon(A)$.

(ii) This result follows immediately from the definition of $\overline{\Lambda}_\epsilon(A)$.

(iii)The idea is to factor the matrix B as $B = z\tilde{I} - A = QR$, where

$$R = \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}, \qquad \tilde{R} = \begin{pmatrix} \rho_{11} & & & \\ & \rho_{22} & * & \\ & & \ddots & \\ & & & \rho_{nn} \end{pmatrix},$$

with $|\rho_{11}| \geq |\rho_{22}| \geq \cdots \geq |\rho_{nn}|$and Q is an $m \times n$ unitary matrix.
This is trivial for $z \in \Lambda(A)$, since $z\tilde{I} - A$ is singular. If $z \notin \Lambda(A)$, i.e., $rank(B) = n$ then $|\rho_{nn}| \neq 0$. Consider that $G[b_1, b_2, ..., b_k] = det(B_k^* B_k) = det(\tilde{R}_k^* \tilde{R}_k) = \rho_{11}^2.\rho_{22}^2...\rho_{kk}^2$, where $\tilde{R}_k$ is a $k \times k$ upper triangular matrix of $\tilde{R}, B_k = [b_1, b_2, ..., b_k]$.

This formula yields

$$G[b_1, b_2, ..., b_{k+1}] = G[b_1, b_2, ..., b_k]\rho_{k+1,k+1}^2, \tag{3.1}$$

which implies

$$\overline{\Lambda}_\epsilon(A(:, 1:k)) \subseteq \overline{\Lambda}_{\epsilon\rho_{k+1,k+1}}(A(:, 1:k+1)), \tag{3.2}$$

If $|\rho_{k+1,k+1}| \leq 1$ then we get $\overline{\Lambda}_\epsilon(A(:, 1:k)) \subseteq \overline{\Lambda}_\epsilon(A(:, 1:k+1))$. □

Theorem 3.2 (Pseudospectra of Similarity Transformation) *Let $m = n$, S is an nonsingular matrix and $C = S^{-1}AS$. Then*

$$\overline{\Lambda}_\epsilon(A) = \overline{\Lambda}_\epsilon(C). \tag{3.3}$$

Proof. Let $C = S^{-1}AS = [c_1, c_2, ..., c_n]$ then we have $G[c_1, ..., c_n] = det(C^T C) = det(A^T A)$. which implies

$$G[b_1, b_2, ..., b_n] = G[c_1, c_2, ..., c_n], \tag{3.4}$$

i.e.,
$$\overline{\Lambda}_\epsilon(A) = \overline{\Lambda}_\epsilon(C)$$
□

The result demonstrates that pseudospectra are invariant under similarity transformation. Consider Definition2.1, we know that $\Lambda_\epsilon(A) \subseteq \Lambda_{\kappa(S)\epsilon}(C)$.($\kappa(S) = \|S\|\|S^{-1}\|$) The results follows from the inequality $\epsilon^{-1} \leq \|(z\tilde{I} - A)^{-1}\| \leq \|S\|\|S^{-1}\|\|(z\tilde{I} - C)^{-1}\|$. This means an ill-conditioned similarity transformation can alter pseudospectra.

Theorem 3.3 *Suppose A is a normal matrix then*
(i) $\overline{\Lambda}_\epsilon(A) = \overline{\Lambda}_\epsilon(\Lambda)$, *where Λ is a diagonal matrix with eigenvalues of A on the main diagonal.*

(ii) *for any $z \in \overline{\Lambda}_\epsilon(A)$, there exists $\lambda_s \in \lambda(A)$such that $|z - \lambda_s| \le \epsilon^{\frac{1}{n}}$ where λ_s is an eigenvalue of A that minimum $|z - \lambda_k|$ for $1 \le k \le n$.*

Proof. (i) This result follows from Theorem3.2.

(ii) Let

$$B = z\tilde{I} - \Lambda = \begin{pmatrix} z-\lambda_1 & & \\ & \ddots & \\ & & z-\lambda_n \end{pmatrix}, \tag{3.5}$$

then

$$det(B^T B) = |z-\lambda_1|^2 |z-\lambda_2|^2 ... |z-\lambda_n|^2.$$

Hence we get

$$G^{\frac{1}{2}}[b_1, b_2 ..., b_n] = |z-\lambda_1||z-\lambda_2|...|z-\lambda_n| \ge |z-\lambda_s|^n. \tag{3.6}$$

If $z \in \overline{\Lambda}_\epsilon(A)$ then we have $|z - \lambda_s| \le \epsilon^{\frac{1}{n}}$. □

Theorem 3.4 *Let A,B are square matrices then $\overline{\Lambda}_\epsilon(AB) = \overline{\Lambda}_\epsilon(BA)$.*

Proof. Notice that $\Lambda(AB) = \Lambda(BA)$. Then we have

$$det(\lambda I - BA) = det(\lambda I - AB).$$

Let $C_1 = \lambda I - BA = [c_1, ..., c_n]$, $\quad C_2 = \lambda I - AB = [c_1', ..., c_n']$ then

$$G[c_1, ..., c_n] = det(C_1^T C_1) = det(C_2^T C_2) = G[c_1', ..., c_n'], \tag{3.7}$$

which yields $$\overline{\Lambda}_\epsilon(AB) = \overline{\Lambda}_\epsilon(BA).$$ □

The same proof shows that if A is an $m \times n$ matrix and B is an $n \times m$ matrix, then AB and BA have the same psedoeigenvalues except that the product which is of higher order has $|m - n|$ extra zero eigenvalues.

The following theorem gives relationship between two definitions of pseudospectra.

Theorem 3.5 *For any given $\epsilon \ge 0$,$\Lambda(A) \subseteq \overline{\Lambda}_\epsilon(A) \subseteq \Lambda_{\epsilon^{\frac{1}{n}}}(A)$.*

Proof. From Theorem3.1 we have that

$$G^{\frac{1}{2}}[b_1, b_2, ..., b_n] = |\rho_{11}||\rho_{22}|...|\rho_{nn}|,$$

where ρ_{nn} is the element of matrix $\tilde{R}$,with $|\rho_{11}| \ge |\rho_{22}| \ge \cdots \ge |\rho_{nn}|$.
From the definition of the minimum singular value of a matrix B,

$$\sigma_{min}(z\tilde{I} - A) = \sigma_{min}(B) = \min_{\|x\|_2=1} \|Bx\|_2.$$

Since $B = QR$ and the unitary invariance of the 2-norm, let $x = e_n$ we have that

$$\sigma_{min}(z\tilde{I} - A) \le |\rho_{nn}|.$$

This formula implies

$$\sigma_{min}^n(B) \leq |\rho_{11}||\rho_{22}|...|\rho_{nn}|,$$

i.e.,

$$\overline{\Lambda}_\epsilon(A) \subseteq \Lambda_{\epsilon^{\frac{1}{n}}}(A).$$ □

Remark. Since the singular values and ρ_{nn} are continuous functions of the matrix entries, hence if $\rho_{nn} \longrightarrow 0$ then $\sigma_{min}(B) \longrightarrow 0$. The converse is also true. It also can be seen that $\overline{\Lambda}_\epsilon(A)$ and $\Lambda_\epsilon(A)$ change continuously with $\epsilon > 0$.

4 Numerical experiments

Now let us calculate pseudospectra properly. The place to begin is with the column pivoting QR decomposition. In numerical experiments,we observe that, if $|\rho_{kk}| >> 1, k = 1, 2, ..., n$ then $G^{\frac{1}{2}}[b_1, b_2, ..., b_n] >> 1$. In order to avoid this situation, we modify our formula $G^{\frac{1}{2}}[b_1, b_2, ..., b_n] \leq \epsilon$ as $G^{\frac{1}{2}}[b_1, b_2, ..., b_n]/|\rho_{11}| \leq \epsilon/|\rho_{11}|$ Hence, the algorithm is to compute column pivoting QR decomposition of $z\tilde{I} - A$ for values of z on a grid in the plane and then generate a contour plot from this data. At last, we also notice that if $G^{\frac{1}{2}}[b_1, b_2, ..., b_n] \leq \epsilon^n$ then we get $\sigma_{min}(B) \leq \epsilon$, but the converse may not be true (see Lawson and Hanson[5]p31).

Algorithm5.1
(1)For each $z \in grid$ computing $B = z\tilde{I} - A = [b_1, ..., b_n]$;
(2) Computing the column pivoting QR decomposition of B;
(3) If $G^{\frac{1}{2}}[b_1, b_2, ..., b_n]/|\rho_{11}| \leq \epsilon/|\rho_{11}|$ then $z \in \overline{\Lambda}_\epsilon(A)$
else goto step (1).

Now we present some numerical examples to examine our conclusions.

Example1 We denote the matrix A=rand(5,5),

$$A = \begin{pmatrix} 0.1934 & 0.6979 & 0.4966 & 0.6602 & 0.7271 \\ 0.6822 & 0.3784 & 0.8998 & 0.3420 & 0.3093 \\ 0.3028 & 0.8600 & 0.8216 & 0.2897 & 0.8385 \\ 0.5417 & 0.8537 & 0.6449 & 0.3412 & 0.5681 \\ 0.1509 & 0.5936 & 0.8180 & 0.5341 & 0.3704 \end{pmatrix}.$$

Figure 1 shows the pseudospectra of A,which the eigenvalue drawn as dots. Note that the sets are nested as indicated in Theorem3.1.In figure2, we see the 5×4 matrix of A(:,1:4). The inclusion properties of Theorem3.1(iii) can be clearly seen that the pseudospectra of the square matrix A are bigger than those of A(:,1:4).

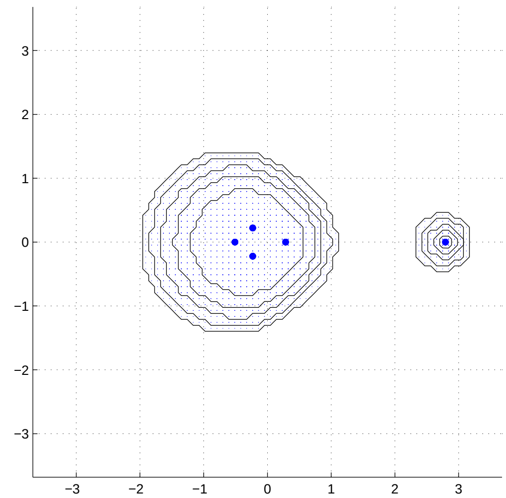

Figure 1:$\epsilon-$pseudospectra ($\epsilon = 0.5, 0.4, 0.3, 0.2, 0.1$) for the matrix A. The grid points are selected with v=80.

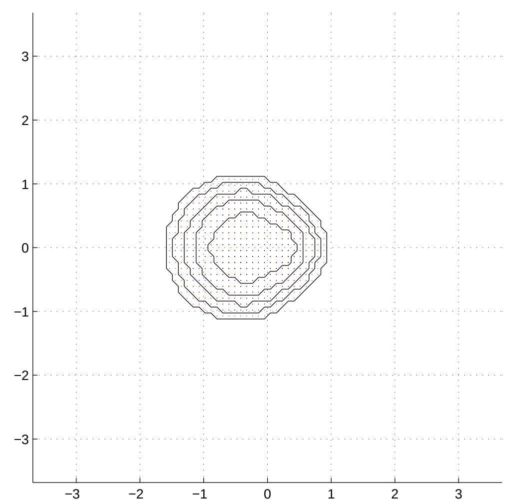

Figure 2:$\epsilon-$pseudospectra ($\epsilon = 0.5, 0.4, 0.3, 0.2, 0.1$) for the matrix A(:,1:4). The grid points are selected with v=80.

Example2 This test matrix is constructed by L.N.Trefethen (see[10]p255). Let $A = A_{N=20}$ and grid points are selected with v=80, the numerical result is shown in Figure 3.

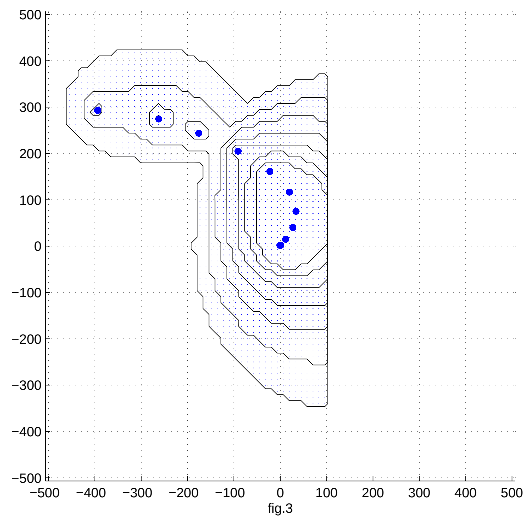

Figure 3 :$\epsilon-$pseudospectra ($\epsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}$) for the matrix A. The grid points are selected with v=80.

5 Summary

We present a new definition of pseudospectra of matrices by using QR decomposition. Based on this definition we get some basic properties. An algorithm for the computation of pseudospectra is given.

References

[1] H.J.Landau, *On Szegö's eigenvalue distribution theory and non-Hermitian kernels.*J.Analyse Math., 28(1975),pp.335-357

[2] N.J.Higham and F.Tisseur, *More on pseudospectra for polynomial eigenvalue problems and applications in control theory.*Tech. Reort 372(2001),Maths Dept.,University of Manchester

[3] X.He, *Theory of numerical linear dependence and its applications.*Numer.Math.J.Chinese Univ.,1(1979),pp.11-19

[4] X.He , *The fundamental properties of pseudoinverse and its applications.*Shanghai Scientific Technology Pulishing House,1985

[5] C.L.Lawson and R.J.Hanson, *Solving Least Squares Problems.*Prentice-Hall,Englewood Cliffs,NJ.Reprinted with a detailed "new developments'appendix in 1996 by SIAM Publications,Philadelphia,PA

[6] K.-C.Toh and L.N.Trefethen, *Calculation of pseudospectra by the Arnoldi iteraion* .SIAM J.Sci.Comput.,17(1996),pp.1-15

[7] L.N.Trefethen,*Pseudospectra of matrices,*in Numerical Analysis 1991,D.F.Griffiths and G.A.Watson,eds.,Longman Scientific and Technical,Harlow,UK,1992,pp.234-266.

[8] L.N.Trefethen,*Pseudospectra of linear operators,*SIAM Rev.,39(1997),pp.383-406.

[9] L.N.Trefethen,*Spectra and pseudospectra:The behavior of non-normal matrices and operators,*in The Graduate Student's Guide to Numerical Analysis,M.Ainsworth,J.Levesley, and M.Marletta,eds.,Springer-Verlag,Berlin,1999,pp.217-250.

[10] L.N.Trefethen,*Computation of pesudospectra.*In Acta numerical 1999,pp.247-295,Cambridge University Press,Cambridge

[11] T.G.Wright and L.N.Trefethen,*Computation of pseudospectra using ARPACK and eigs.*SIAM J.Sci.Comput., 23(2001),pp.591-605

[12] T.G.Wright and L.N.Trefethen,*Eigenvalues and pseudospectra of rectangular matrices* ,Tech. report 01/13, Oxford University Computing Laboratory Numerical Analysis Group