

MATH 5524 · MATRIX THEORY

Problem Set 5

Posted Tuesday 11 April 2017. Due Tuesday 18 April 2017.

[Late work is due on Wednesday 19 April 2017.]

Complete any four problems, 25 points each.

Recall the definitions of the *numerical range* (*field of values*) of $\mathbf{A} \in \mathbb{C}^{n \times n}$,

$$W(\mathbf{A}) = \left\{ \frac{\mathbf{x}^* \mathbf{A} \mathbf{x}}{\mathbf{x}^* \mathbf{x}} : \mathbf{x} \in \mathbb{C}^n \right\},$$

and the ε -*pseudospectrum* for $\varepsilon > 0$,

$$\begin{aligned} \sigma_\varepsilon(\mathbf{A}) &= \{z \in \mathbb{C} : \|(z\mathbf{I} - \mathbf{A})^{-1}\| > 1/\varepsilon\} \\ &= \{z \in \mathbb{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbb{C}^{n \times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}. \end{aligned}$$

1. The derivative of a distinct eigenvalue λ_k of a generic matrix $\mathbf{A}(t) = \mathbf{A}_0 + t\mathbf{E}$ is given by

$$\lambda'_k(t) \Big|_{t=0} = \frac{\widehat{\mathbf{v}}_k^* \mathbf{E} \mathbf{v}_k}{\widehat{\mathbf{v}}_k^* \mathbf{v}_k},$$

where \mathbf{v}_k and $\widehat{\mathbf{v}}_k$ are the corresponding right and left eigenvectors of \mathbf{A}_0 .

- (a) Suppose $\mathbf{A} \in \mathbb{C}^{n \times n}$ has singular value decomposition $\mathbf{A} = \sum_{j=1}^n s_j \mathbf{u}_j \mathbf{v}_j^*$, and consider the matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^* & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{2n \times 2n}.$$

Show that $\sigma(\mathbf{H}) = \{\pm s_1, \dots, \pm s_n\}$. What are the corresponding right and left eigenvectors of \mathbf{H} ?

- (b) Use the result stated at the beginning of this problem to derive a formula for the derivative of a distinct nonzero *singular value* of $\mathbf{A} + t\mathbf{E}$.

This result forms the basis of curve tracing algorithms for computing $\sigma_\varepsilon(\mathbf{A})$. Given some z on the boundary of $\sigma_\varepsilon(\mathbf{A})$, one seeks to follow the contour in \mathbb{C} for which $s_n(z\mathbf{I} - \mathbf{A}) = 1/\|(z\mathbf{I} - \mathbf{A})^{-1}\| = \varepsilon$.

2. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding right eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ and left eigenvectors $\widehat{\mathbf{v}}_1, \dots, \widehat{\mathbf{v}}_n$.

The derivative of the eigenvalue $\lambda_k(t)$ of $\mathbf{A} + t\mathbf{E}$ with respect to t (for $\|\mathbf{E}\| = 1$) at $t = 0$ is bounded in magnitude by the *eigenvalue condition number*

$$\kappa(\lambda_k) = \frac{\|\widehat{\mathbf{v}}_k\| \|\mathbf{v}_k\|}{|\widehat{\mathbf{v}}_k^* \mathbf{v}_k|}.$$

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & \alpha \\ 0 & -2 \end{bmatrix},$$

where $\alpha \geq 1$ is a real constant.

- (a) Compute the eigenvalue condition numbers $\kappa(\lambda_1)$ and $\kappa(\lambda_2)$ for eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$.
(b) Compute the matrix exponential $e^{t\mathbf{A}}$.

- (c) What is $\lim_{t \rightarrow \infty} \|e^{t\mathbf{A}}\|$?
 (You do not need to explicitly compute the norm to answer this question.)
- (d) For fixed $\alpha > 4$, find the value of $t \geq 0$ that maximizes the largest entry in $e^{t\mathbf{A}}$.
- (e) Compute a formula (dependent on α) for $\frac{d}{dt}\|e^{t\mathbf{A}}\|$ at $t = 0$.
 For what values of α is this quantity positive?
- (f) What do your solutions to (c), (d) and (e) suggest about the general behavior of $\|e^{t\mathbf{A}}\|$ for $t \in (0, \infty)$ with $\alpha > 4$? How does $\max_{t \geq 0} \|e^{t\mathbf{A}}\|$ depend on α ?
 (Again, you do not need to explicitly compute any matrix norms; just describe the general behavior at small t , intermediate t , and large t .)

3. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix}$$

for real numbers $\alpha, \beta > 0$.

- (a) For every $z \in W(\mathbf{A})$, one can show that there exists a unit vector of the form

$$\mathbf{x} = \begin{bmatrix} t \\ e^{i\theta} \sqrt{1-t^2} \end{bmatrix}$$

with $t \in [0, 1]$ and $\theta \in [0, 2\pi)$ such that $z = \mathbf{x}^* \mathbf{A} \mathbf{x}$. (You do not need to prove this; up to a factor of a complex sign ($e^{i\phi}$), this form for \mathbf{x} is just a generic unit vector in \mathbb{C}^2 .)

Show that

$$\mathbf{x}^* \mathbf{A} \mathbf{x} = t\sqrt{1-t^2} \left((\alpha + \beta) \cos \theta + i(\alpha - \beta) \sin \theta \right)$$

and explain why this implies that $W(\mathbf{A})$ contains all the points on an *ellipse* in the complex plane.

- (b) What is the center of this ellipse? What are the major and minor axes?
- (c) For which pairs of α and β does this ellipse degenerate into a point or line segment?
 (Describe all such cases.)
- (d) Use this result to plot (in MATLAB etc.) the boundary of $W(\mathbf{A})$ and the eigenvalues of \mathbf{A} for
- (i) $\alpha = 1$ and $\beta = 0$;
 - (ii) $\alpha = 1$ and $\beta = 2$;
 - (iii) $\alpha = 1$ and $\beta = -2$.

This problem is the first step in proving that $W(\mathbf{A})$ is convex for any $\mathbf{A} \in \mathbb{C}^{n \times n}$. One can reduce the $n \times n$ problem down to a 2×2 problem, then reduce a generic 2×2 problem to this form using shifts and unitary transformations. For the complete proof, see Horn and Johnson.

4. The exercise is designed to give you some intuition for transient growth in a dynamical system.

The *Leslie matrix* arises in models for the (female) population of a given species with fixed birth rates and survivability levels. The population is divided into n brackets of y -years each, and an average member of bracket k gives birth to $b_k \geq 0$ females in the next y years, and has probability $s_k \in [0, 1]$ of surviving the next y years. Letting $p_k^{(j)}$ denote the population in the k th bracket in the j th generation, we see that the population evolves according to the matrix equation (e.g., for $n = 5$)

$$\begin{bmatrix} p_1^{(j+1)} \\ p_2^{(j+1)} \\ p_3^{(j+1)} \\ p_4^{(j+1)} \\ p_5^{(j+1)} \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 \\ s_1 & & & & \\ & s_2 & & & \\ & & s_3 & & \\ & & & s_4 & \\ & & & & s_5 \end{bmatrix} \begin{bmatrix} p_1^{(j)} \\ p_2^{(j)} \\ p_3^{(j)} \\ p_4^{(j)} \\ p_5^{(j)} \end{bmatrix},$$

with unspecified entries equal to zero. (We presume the mortality of the last age bracket.)

Denote the matrix as \mathbf{A} , so that $\mathbf{p}^{(j+1)} = \mathbf{A}\mathbf{p}^{(j)}$, and hence $\mathbf{p}^{(j)} = \mathbf{A}^j \mathbf{p}^{(0)}$.

Earlier on this problem set, we considered how transient growth in matrix powers is linked to the sensitivity of eigenvalues to perturbations in the matrix entries. This problem is designed to reinforce this connection in the context of physically meaningful transient behavior.

- (a) Design a set of parameters $b_1, \dots, b_5 > 0$ and $s_1, \dots, s_4 > 0$ (for $n = 5$) so that the population will eventually decay in size to zero ($\mathbf{A}^j \rightarrow \mathbf{0}$ as $j \rightarrow \infty$), but this will be preceded by a period of significant transient growth in the population, where $\mathbf{p}^{(j)} \gg \mathbf{p}^{(0)}$. (Think about what kind of birth and survivability values might suggest this demographic pattern.)
- (b) Plot your population for a number of generations to demonstrate the transient growth and eventual decay. (You may modify `pop.m` from the class website.)
- (c) Show that this transient growth coincides with sensitivity of the eigenvalues of your matrix. You may choose to do this in several different ways: compute the condition numbers of the eigenvalues; show that the numerical range of \mathbf{A} contains points of significantly larger magnitude than the spectral radius; or download EigTool for MATLAB (see the link on the class website) and use it to plot the pseudospectra of \mathbf{A} .

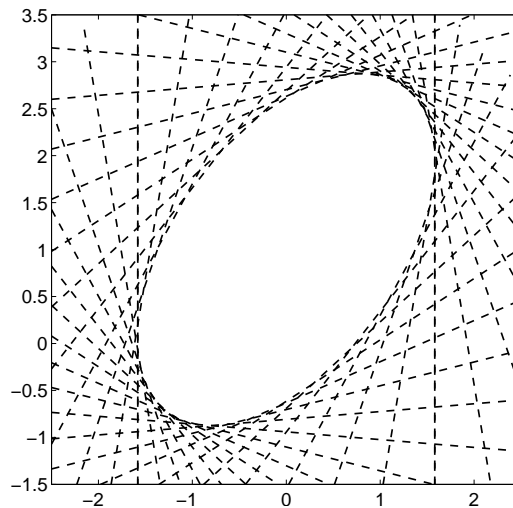
5. Johnson's algorithm for approximating the boundary of $W(\mathbf{A})$ proceeds as follows.

1. Pick m angles $0 = \theta_1 < \theta_2 < \dots < \theta_m = \pi$.
2. For $k = 1, \dots, m$
 - 2a. Define $\mathbf{H}_k := (\mathbf{e}^{i\theta_k} \mathbf{A} + \mathbf{e}^{-i\theta_k} \mathbf{A}^*)/2$
 - 2b. Compute the extreme eigenvalues $\lambda_{\min}^{(k)}$ and $\lambda_{\max}^{(k)}$ of \mathbf{H}_k
 - 2c. Plot the lines $\{e^{-i\theta_k}(\lambda_{\min}^{(k)} + i\gamma) : \gamma \in \mathbb{R}\}$ and $\{e^{-i\theta_k}(\lambda_{\max}^{(k)} + i\gamma) : \gamma \in \mathbb{R}\}$, which bound $W(\mathbf{A})$

For example, the figure below shows these lines in \mathbb{C} for 20 values of θ for the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 3i \\ 1 & 2i \end{bmatrix}.$$

The numerical range is contained between each pair of these lines (i.e., in the empty region in the center of the plot).



- (a) Implement Johnson's algorithm.
- (b) Demonstrate Johnson's algorithm (with $m = 20$, say) by producing plots for the following:
- (i) The 2×2 matrix \mathbf{A} shown above;
 - (ii) $\mathbf{A} = \text{diag}(\text{ones}(15, 1), 1)$ (a 16×16 Jordan block);
 - (iii) $\mathbf{A} = \text{diag}(\text{ones}(15, 1), 1) + \text{diag}(1, -15)$ (a circulant shift);
 - (iv) $\mathbf{A} = \text{randn}(64)/8$ (a random 64×64 matrix).
- (c) Confirm the accuracy of your bounds by including in each of your plots 100 random points in $W(\mathbf{A})$. For example, let $\mathbf{x} = \text{orth}(\text{randn}(n, 1) + 1i * \text{randn}(n, 1))$ (to get a random complex unit vector) and plot $z = \mathbf{x}^* \mathbf{A} \mathbf{x}$; repeat this for 99 more random complex unit vectors.

6. Prove the following two results that bound the ε -pseudospectrum of \mathbf{A} . Roughly speaking, they can be interpreted to mean, "the ε -pseudospectrum cannot be significantly larger than the field of values or the Gerschgorin disks, beyond a factor of order ε ."

- (a) Prove that for all $\mathbf{A} \in \mathbb{C}^{n \times n}$,

$$\sigma_\varepsilon(\mathbf{A}) \subseteq W(\mathbf{A}) + D(\varepsilon),$$

where $D(\varepsilon) = \{z \in \mathbb{C} : |z| < \varepsilon\}$ is the open disk of radius ε .

(Addition of sets is defined in the usual way: Given sets S_1 and S_2 , the sum $S_1 + S_2 = \{s_1 + s_2 : s_1 \in S_1, s_2 \in S_2\}$ is the set of all pairwise sums between elements in S_1 and S_2 .)

- (b) Use Gerschgorin's theorem to show that for all $\mathbf{A} \in \mathbb{C}^{n \times n}$

$$\sigma_\varepsilon(\mathbf{A}) \subseteq \bigcup_{j=1}^n D(a_{j,j}, r_j + \varepsilon n)$$

where $a_{j,k}$ denotes the (j, k) entry of \mathbf{A} , r_j denotes the absolute sum of the off-diagonal entries in the j th row of \mathbf{A} ,

$$r_j := \sum_{k=1, k \neq j}^n |a_{j,k}|,$$

and $D(c, r) = \{z \in \mathbb{C} : |z - c| < r\}$ is the open disk in \mathbb{C} centered at $c \in \mathbb{C}$ with radius $r > 0$.

(In fact, one can prove a stronger bound with $D(a_{j,j}, r_j + \varepsilon n)$ replaced by $D(a_{j,j}, r_j + \varepsilon \sqrt{n})$, but you are not required to obtain this sharper result.)

7. We have often discussed the linear, constant-coefficient dynamical system $\mathbf{x}'(t) = \mathbf{A} \mathbf{x}(t)$, whose solutions $\mathbf{x}(t)$ decay to zero as $t \rightarrow \infty$, provided all eigenvalues of \mathbf{A} have negative real part.

Is the same true for *variable-coefficient* problems? Suppose $\mathbf{x}'(t) = \mathbf{A}(t) \mathbf{x}(t)$, and that all eigenvalues of the matrix $\mathbf{A}(t) \in \mathbb{C}^{n \times n}$ have negative real part for all $t \geq 0$. Is this enough to guarantee that $\mathbf{x}(t) \rightarrow \mathbf{0}$ as $t \rightarrow \infty$? This problem asks you to explore this possibility.

- (a) Consider the matrix

$$\mathbf{U}(t) = \begin{bmatrix} \cos(\gamma t) & \sin(\gamma t) \\ -\sin(\gamma t) & \cos(\gamma t) \end{bmatrix}.$$

Show that $\mathbf{U}(t)$ is unitary for any fixed real values of γ and t .

(b) Now consider the matrix $\mathbf{A}(t) \in \mathbb{C}^{n \times n}$ defined by

$$\mathbf{A}(t) = \mathbf{U}(t)\mathbf{A}_0\mathbf{U}(t)^*, \quad \mathbf{A}_0 = \begin{bmatrix} -1 & \alpha \\ 0 & -2 \end{bmatrix}.$$

(Notice that \mathbf{A}_0 is the matrix that featured in Problem 2.)

Explain why $\sigma(\mathbf{A}(t)) = \sigma(\mathbf{A}_0)$, $W(\mathbf{A}(t)) = W(\mathbf{A}_0)$, and $\sigma_\varepsilon(\mathbf{A}(t)) = \sigma_\varepsilon(\mathbf{A}_0)$ for all real t .

(In other words, show the spectrum, numerical range, and ε -pseudospectra are identical for all t .)

(c) Now we wish to investigate the behavior of the dynamical system

$$\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t). \quad (*)$$

Define $\mathbf{y}(t) = \mathbf{U}(t)^*\mathbf{x}(t)$. Explain why equation (*) implies that

$$\mathbf{y}'(t) = (\mathbf{A}_0 + (\mathbf{U}(t)^*)'\mathbf{U}(t))\mathbf{y}(t). \quad (**)$$

(Here $(\mathbf{U}(t)^*)' \in \mathbb{C}^{n \times n}$ denotes the t -derivative of the conjugate-transpose of $\mathbf{U}(t)$.)

(d) Compute $(\mathbf{U}(t)^*)'\mathbf{U}(t)$. Does this matrix vary with t ?

(e) Define the matrix

$$\widehat{\mathbf{A}} = \mathbf{A}_0 + (\mathbf{U}(t)^*)'\mathbf{U}(t).$$

Fix $\alpha = 7$. Plot the (real) eigenvalues of $\widehat{\mathbf{A}}$ (e.g., in MATLAB) as a function of $\gamma \in [0, 7]$. Do the eigenvalues of $\widehat{\mathbf{A}}$ fall in the left half of the complex plane for all γ ?

(f) Calculate the eigenvalues of $\widehat{\mathbf{A}}$ for $\gamma = 1$ and $\alpha = 7$.

What can be said of solutions $\mathbf{y}(t)$ to the system (**) as $t \rightarrow \infty$ for these α, γ values?

What then can be said of $\|\mathbf{x}(t)\| = \|\mathbf{U}(t)\mathbf{y}(t)\|$, where $\mathbf{x}(t)$ solves (*), as $t \rightarrow \infty$?

How does this compare to the similar constant coefficient problem $\mathbf{x}'(t) = \mathbf{A}_0\mathbf{x}(t)$ (where we have seen that \mathbf{A}_0 has the same spectrum, numerical range, and pseudospectra as $\mathbf{A}(t)$ for all t)?

[Examples of this sort were perhaps first constructed by Vinograd; see Dekker and Verwer, or Lambert.]

8. Prove or disprove Crouzeix's conjecture: For any matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ and any function f analytic on $W(\mathbf{A})$,

$$\|f(\mathbf{A})\| \leq 2 \max_{z \in W(\mathbf{A})} |f(z)|.$$

[This is a challenge problem: a solution has not yet been discovered!]