Pseudospectra
and
Nonnormal Dynamical Systems

Mark Embree and Russell Carden
Computational and Applied Mathematics
Rice University
Houston, Texas

ELGERSBURG

MARCH 2012

Overview of the Course

These lectures describe modern tools
for the spectral analysis of dynamical systems.
We shall cover a mix of theory, computation, and applications.

Lecture 1: Introduction to Nonnormality and Pseudospectra

Lecture 2: Functions of Matrices

Lecture 3: Toeplitz Matrices and Model Reduction

Lecture 4: Model Reduction, Numerical Algorithms, Differential Operators

Lecture 5: Discretization, Extensions, Applications

Outline for Today

Lecture 2: Functions of Matrices

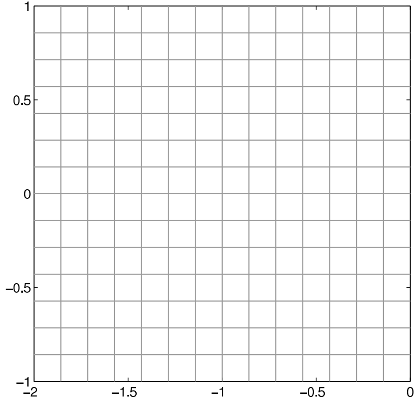
- Recap computation of $\sigma_\varepsilon(\mathbf{A})$ and $W(\mathbf{A})$
- Bounds on $\|f(\mathbf{A})\|$ using $W(\mathbf{A})$
- Behavior of the resolvent near eigenvalues
- Bounds on $\|f(\mathbf{A})\|$ using $\sigma_\varepsilon(\mathbf{A})$
- Computing the pseudospectral abscissa and radius

1(e) Computing $W(\mathbf{A})$ and $\sigma_\varepsilon(\mathbf{A})$

Computation of Pseudospectra (Dense Matrices)

Naive algorithm: $\mathcal{O}(n^3)$ per grid point

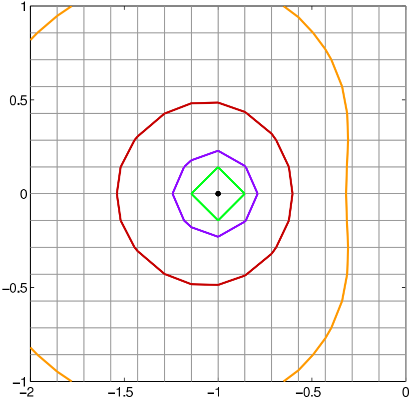
- Compute $\|(z - \mathbf{A})^{-1}\|$ using the SVD on a grid of points in $\mathbf{C}$.
  SVD costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Send data to a contour plotting routine.



Computation of Pseudospectra (Dense Matrices)

Naive algorithm: $\mathcal{O}(n^3)$ per grid point

- Compute $\|(z-\mathbf{A})^{-1}\|$ using the SVD on a grid of points in $\mathbb{C}$.
  SVD costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Send data to a contour plotting routine.



Computation of Pseudospectra (Dense Matrices)

Modern algorithm: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point [Lui 1997; Trefethen 1999]

- Compute a Schur triangularization, $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*$.
  Schur costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Recall that $\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{T})$.

Computation of Pseudospectra (Dense Matrices)

Modern algorithm: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point [Lui 1997; Trefethen 1999]

- Compute a Schur triangularization, $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*$.
  Schur costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Recall that $\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{T})$.
- One can efficiently solve linear systems $\mathbf{T}\mathbf{x} = \mathbf{b}$ for $\mathbf{x}$.
  Each solve costs $\mathcal{O}(n^2)$ for upper triangular $\mathbf{T}$.

Computation of Pseudospectra (Dense Matrices)

Modern algorithm: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point [Lui 1997; Trefethen 1999]

- Compute a Schur triangularization, $\mathbf{A} = \mathbf{UTU}^*$.
  Schur costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Recall that $\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{T})$.
- One can efficiently solve linear systems $\mathbf{Tx} = \mathbf{b}$ for $\mathbf{x}$.
  Each solve costs $\mathcal{O}(n^2)$ for upper triangular $\mathbf{T}$.
- To compute $\|(z - \mathbf{T})^{-1}\| = 1/s_{\min}(z - \mathbf{T}) = s_{\max}((z - \mathbf{T})^{-1})$,
  find the largest eigenvalue of $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$.
- To do so, apply the *Lanczos iteration* on $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$
  for each grid point z.

Computation of Pseudospectra (Dense Matrices)

Modern algorithm: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point [Lui 1997; Trefethen 1999]

- Compute a Schur triangularization, $\mathbf{A} = \mathbf{UTU}^*$.
  Schur costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Recall that $\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{T})$.
- One can efficiently solve linear systems $\mathbf{Tx} = \mathbf{b}$ for $\mathbf{x}$.
  Each solve costs $\mathcal{O}(n^2)$ for upper triangular $\mathbf{T}$.
- To compute $\|(z - \mathbf{T})^{-1}\| = 1/s_{\min}(z - \mathbf{T}) = s_{\max}((z - \mathbf{T})^{-1})$,
  find the largest eigenvalue of $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$.
- To do so, apply the *Lanczos iteration* on $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$
  for each grid point z.
- At each step, one must compute $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}\mathbf{b}$:
  This requires a lower-triangular solve with $(z - \mathbf{T})^*$ $(\mathcal{O}(n^2))$
  and an upper triangular solve with $(z - \mathbf{T})$ $(\mathcal{O}(n^2))$.
  One observes a constant number of iterations per grid point.

Computation of Pseudospectra (Dense Matrices)

Modern algorithm: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point [Lui 1997; Trefethen 1999]

- Compute a Schur triangularization, $\mathbf{A} = \mathbf{UTU}^*$.
  Schur costs $\mathcal{O}(n^3)$ for dense $\mathbf{A}$.
- Recall that $\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{T})$.
- One can efficiently solve linear systems $\mathbf{Tx} = \mathbf{b}$ for $\mathbf{x}$.
  Each solve costs $\mathcal{O}(n^2)$ for upper triangular $\mathbf{T}$.
- To compute $\|(z - \mathbf{T})^{-1}\| = 1/s_{\min}(z - \mathbf{T}) = s_{\max}((z - \mathbf{T})^{-1})$,
  find the largest eigenvalue of $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$.
- To do so, apply the *Lanczos iteration* on $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}$
  for each grid point z.
- At each step, one must compute $(z - \mathbf{T})^{-*}(z - \mathbf{T})^{-1}\mathbf{b}$:
  This requires a lower-triangular solve with $(z - \mathbf{T})^*$ $(\mathcal{O}(n^2))$
  and an upper triangular solve with $(z - \mathbf{T})$ $(\mathcal{O}(n^2))$.
  One observes a constant number of iterations per grid point.
- Total cost: $\mathcal{O}(n^3) + \mathcal{O}(n^2)$ per grid point.

Computation of Pseudospectra (Sparse Matrices)

Large-scale problems [Toh and Trefethen 1996; Wright and Trefethen 2001]

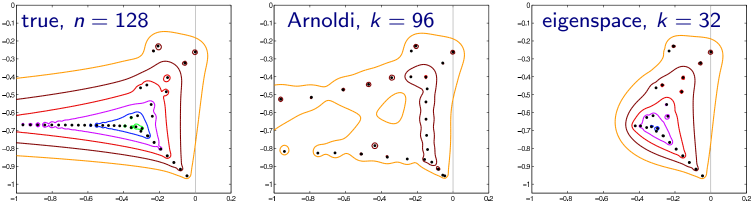
Key idea:

- Find $\mathbf{V} \in \mathbb{C}^{n\times k}$ with orthonormal columns, $\mathbf{V}^*\mathbf{AV}$, for $k \ll n$.
- The "generalized Rayleigh quotient" $\mathbf{V}^*\mathbf{AV} \in \mathbb{C}^{k\times k}$.
- *Approximate* $\sigma_\varepsilon(\mathbf{V}^*\mathbf{AV}) \approx \sigma_\varepsilon(\mathbf{A})$.

In general, $\sigma(\mathbf{V}^*\mathbf{AV}) \not\subseteq \sigma(\mathbf{A})$, so for some $\varepsilon > 0$,

$$\sigma_\varepsilon(\mathbf{V}^*\mathbf{AV}) \not\subseteq \sigma_\varepsilon(\mathbf{A}).$$

Depending on the choice of $\mathbf{V}$, the approximation $\sigma_\varepsilon(\mathbf{V}^*\mathbf{AV})$ might give a rough general impression of $\sigma_\varepsilon(\mathbf{A})$, or it might give a rather accurate approximation in one interesting region of $\sigma_\varepsilon(\mathbf{A})$.



Computation of Pseudospectra (Sparse Matrices)

First important choice for $\mathbf{V}_k$ [Toh and Trefethen 1996]:

- Projection onto Krylov subspaces (Arnoldi factorization)

$$\mathrm{Ran}(\mathbf{V}_k) = \mathrm{span}\{\mathbf{x}, \mathbf{A}\mathbf{x}, \dots, \mathbf{A}^{k-1}\mathbf{x}\}.$$

The Arnoldi process generates orthonormal bases for Krylov subspaces:

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\widetilde{\mathbf{H}}_k,$$

where $\widetilde{\mathbf{H}}_k \in \mathbb{C}^{(k+1)\times k}$ is upper Hessenberg.
Due to the upper Hessenberg structure, we have

$$s_{\min}(z - \widetilde{\mathbf{H}}_k) \geq s_{\min}(z - \mathbf{A}),$$

so one can get a *bound*

$$\sigma_\varepsilon(\widetilde{\mathbf{H}}_k) \subseteq \sigma_\varepsilon(\mathbf{A}),$$

where, for a rectangular matrix $\mathbf{R}$, we have

$$\sigma_\varepsilon(\mathbf{R}) := \{z \in \mathbb{C} : s_{\min}(z - \mathbf{R}) < \varepsilon\}.$$

Computation of Pseudospectra (Sparse Matrices)

Second important choice for $\mathbf{V}_k$ [Wright and Trefethen 2001]:

- Projection onto an invariant suspace (eigenspace)

$$\mathbf{V}_k = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$$

Suppose $\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{X}$ for some $\mathbf{X} \in \mathbb{C}^{k\times k}$.
Then $\mathrm{Ran}(\mathbf{V}_k)$ is an *invariant subspace* of $\mathbf{A}$.

If $\mathbf{V} = [\mathbf{V}_k\ \widehat{\mathbf{V}}_k]$ is a unitary matrix, $\mathbf{V}^*\mathbf{V} = \mathbf{I}$, then

$$\mathbf{V}^*\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k & \widehat{\mathbf{V}}_k^*\mathbf{A}\mathbf{V}_k \\ \mathbf{0} & \widehat{\mathbf{V}}_k^*\mathbf{A}\widehat{\mathbf{V}}_k \end{bmatrix}.$$

Thus $\sigma_\varepsilon(\mathbf{V}_k^*\mathbf{A}\mathbf{V}_k) \subseteq \sigma_\varepsilon(\mathbf{A})$.

Compute an invariant subspace corresponding to eigenvalues of physical interest (e.g., using ARPACK).

Computation of Pseudospectra

Alternative: [Brühl 1996; Bekas and Gallopoulos, ...]
Curve tracing: follow level sets of $\|(z - \mathbf{A})^{-1}\|$.

- Given a point $z = x + iy \in \mathbf{C}$, suppose $\|z - \mathbf{A}\|^{-1} = 1/\varepsilon$.
- Suppose the smallest singular value s of $z - \mathbf{A}$ is simple, with singular vectors $\mathbf{u}$ and $\mathbf{v}$:
$$(z - \mathbf{A})\mathbf{v} = s\mathbf{u}.$$
- Brühl uses a result of Sun (1988) to obtain
$$\frac{\partial s}{\partial x} = \mathrm{Re}(\mathbf{u}^*\mathbf{v}), \qquad \frac{\partial s}{\partial y} = \mathrm{Im}(\mathbf{v}^*\mathbf{u}).$$
- Use these derivatives to follow the boundary $\partial\sigma_\varepsilon(\mathbf{A})$.

Computation of Pseudospectra

Alternative: [Brühl 1996; Bekas and Gallopoulos, ...]
Curve tracing: follow level sets of $\|(z - \mathbf{A})^{-1}\|$.

- Given a point $z = x + iy \in \mathbb{C}$, suppose $\|z - \mathbf{A}\|^{-1} = 1/\varepsilon$.
- Suppose the smallest singular value s of $z - \mathbf{A}$ is simple, with singular vectors $\mathbf{u}$ and $\mathbf{v}$:

$$(z - \mathbf{A})\mathbf{v} = s\mathbf{u}.$$

- Brühl uses a result of Sun (1988) to obtain

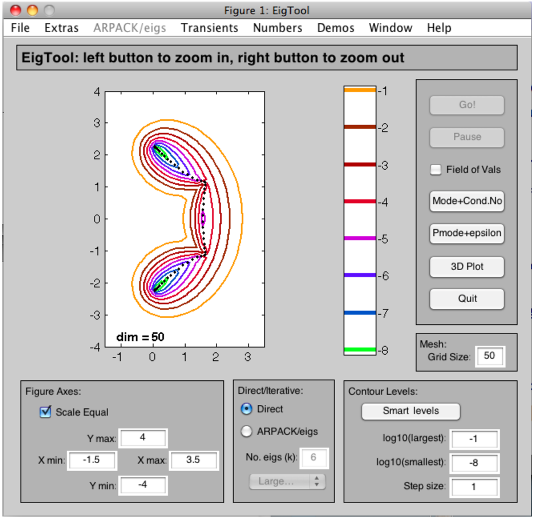
$$\frac{\partial s}{\partial x} = \mathrm{Re}(\mathbf{u}^*\mathbf{v}), \qquad \frac{\partial s}{\partial y} = \mathrm{Im}(\mathbf{v}^*\mathbf{u}).$$

- Use these derivatives to follow the boundary $\partial\sigma_\varepsilon(\mathbf{A})$.

This seems like an elegant alternative to the grid-based method, but:

- It only gives $\sigma_\varepsilon(\mathbf{A})$ for one value of ε.
- One must beware of cusps, holes, disconnected components of $\sigma_\varepsilon(\mathbf{A})$.

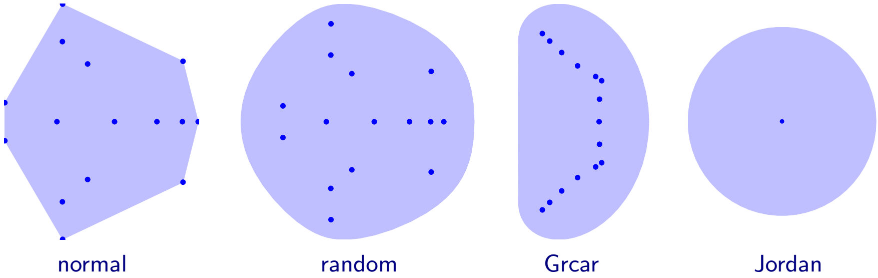
EigTool: Software for Pseudospectra Computation



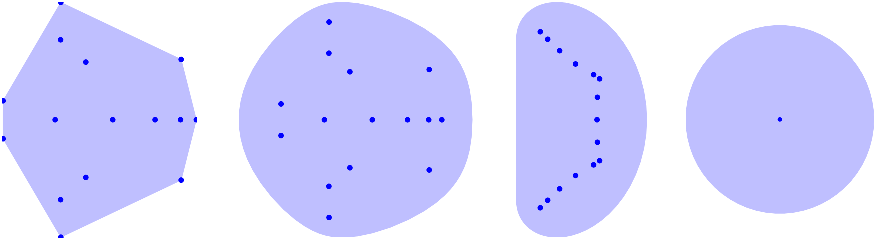
EigTool: Thomas Wright, 2002

http://www.cs.ox.ac.uk/pseudospectra/eigtool

Computing the Numerical Range



Computing the Numerical Range



normal random Grcar Jordan

If $z \in W(\mathbf{A})$, then

$$\operatorname{Re} z = \frac{z + \overline{z}}{2} = \frac{1}{2}(\mathbf{x}^*\mathbf{A}\mathbf{x} + \mathbf{x}^*\mathbf{A}^*\mathbf{x}) = \mathbf{x}^*\Big(\frac{\mathbf{A}+\mathbf{A}^*}{2}\Big)\mathbf{x}.$$

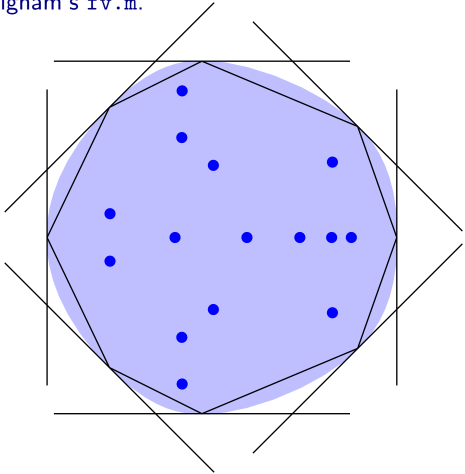
Using properties of Hermitian matrices, we conclude that

$$\operatorname{Re}(W(\mathbf{A})) = \Big[\lambda_{\min}\Big(\frac{\mathbf{A}+\mathbf{A}^*}{2}\Big), \lambda_{\max}\Big(\frac{\mathbf{A}+\mathbf{A}^*}{2}\Big)\Big].$$

Similarly, one can determine the intersection of $W(\mathbf{A})$ with any line in $\mathbb{C}$.

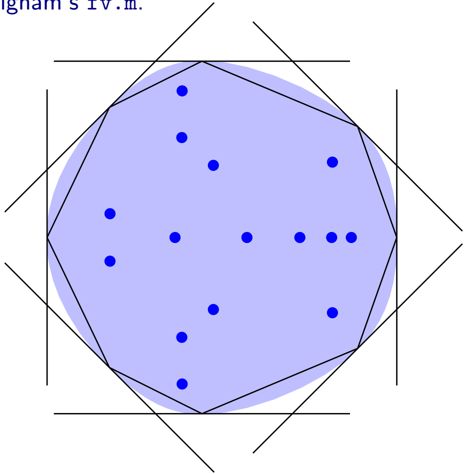
Computation of the Numerical Range

This calculation yields points on the boundary of the numerical range.
Use convexity to obtain polygonal outer and inner approximations
[Johnson 1980]; Higham's `fv.m`.



Computation of the Numerical Range

This calculation yields points on the boundary of the numerical range.
Use convexity to obtain polygonal outer and inner approximations [Johnson 1980]; Higham's `fv.m`.



Neat problem: Given $z \in W(\mathbf{A})$, find unit vector $\mathbf{x}$ such that $z = \mathbf{x}^* \mathbf{A} \mathbf{x}$ [Uhlig 2008; Carden 2009].

2. Functions of Matrices and Toeplitz Matrices

Behavior of Matrices

We are primarily interested in using pseudospectra (and other tools) to study the *behavior* of a nonnormal matrix.

In particular, we seek to quantify (or bound) how nonnormality affects the value of functions of matrices.

Much research has been devoted to functions of matrices over the past decade; see the book by Nick Higham [Hig08]. We focus on functions that are analytic on the spectrum of $\mathbf{A}$.

Theorem (Spectral Mapping Theorem)

Suppose f is analytic on $\sigma(\mathbf{A})$. Then

$$\sigma(f(\mathbf{A})) = f(\sigma(\mathbf{A})).$$

Thus $\|f(\mathbf{A})\| \geq \max_{\zeta\in\sigma(f(\mathbf{A}))} |\zeta| = \max_{\lambda\in\sigma(\mathbf{A})} |f(\lambda)|$.

See [Huhtanen, 1999] for work on lower bounds.

2(a) Bounds on $\|f(\mathbf{A})\|$ using $W(\mathbf{A})$

Numerical Range and the Matrix Exponential

What does the numerical range reveal about matrix behavior?

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\|\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{x}_0^*\mathrm{e}^{t\mathbf{A}^*}\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\Big)^{1/2} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{x}_0^*(\mathbf{I}+t\mathbf{A}^*)(\mathbf{I}+t\mathbf{A})\mathbf{x}_0\Big)^{1/2} \\ &= \frac{1}{\|\mathbf{x}_0\|}\mathbf{x}_0^*\Big(\frac{\mathbf{A}+\mathbf{A}^*}{2}\Big)\mathbf{x}_0 \end{aligned}$$

So, the rightmost point in $W(\mathbf{A})$ reveals the maximal slope of $\|\mathrm{e}^{t\mathbf{A}}\|$ at $t=0$.

Numerical Range and the Matrix Exponential

What does the numerical range reveal about matrix behavior?

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\|\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{x}_0^*\mathrm{e}^{t\mathbf{A}^*}\mathrm{e}^{t\mathbf{A}}\mathbf{x}_0\right)^{1/2} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{x}_0^*(\mathbf{I}+t\mathbf{A}^*)(\mathbf{I}+t\mathbf{A})\mathbf{x}_0\right)^{1/2} \\
&= \frac{1}{\|\mathbf{x}_0\|}\mathbf{x}_0^*\left(\frac{\mathbf{A}+\mathbf{A}^*}{2}\right)\mathbf{x}_0
\end{aligned}$$

So, the rightmost point in $W(\mathbf{A})$ reveals the maximal slope of $\|\mathrm{e}^{t\mathbf{A}}\|$ at $t = 0$.

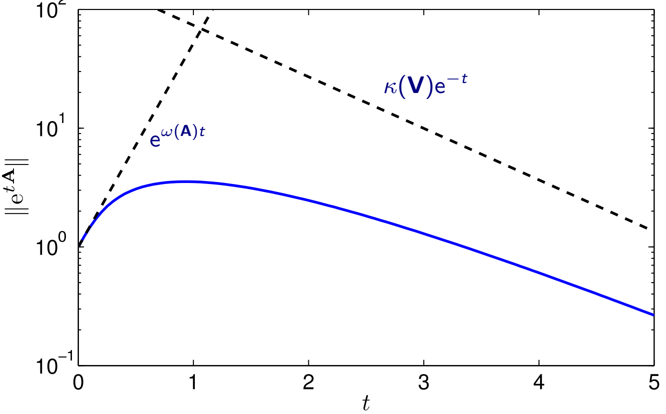
Definition (numerical abscissa)

The *numerical abscissa* is the rightmost in $W(\mathbf{A})$:

$$\omega(\mathbf{A}) := \max_{z \in W(\mathbf{A})} \operatorname{Re} z.$$

Initial Transient Growth via Numerical Abscissa

$$\mathbf{A} = \begin{bmatrix} -1.1 & 10 \\ 0 & -1 \end{bmatrix}$$



Numerical Range and Matrix Norm

Definition

The *numerical radius* of $\mathbf{A} \in \mathbb{C}^{n \times n}$ is the largest magnitude of a point in $W(\mathbf{A})$:

$$\mu(\mathbf{A}) := \max_{z \in W(\mathbf{A})} |z|.$$

Theorem

$$\frac{1}{2}\mu(\mathbf{A}) \leq \|\mathbf{A}\| \leq \mu(\mathbf{A}).$$

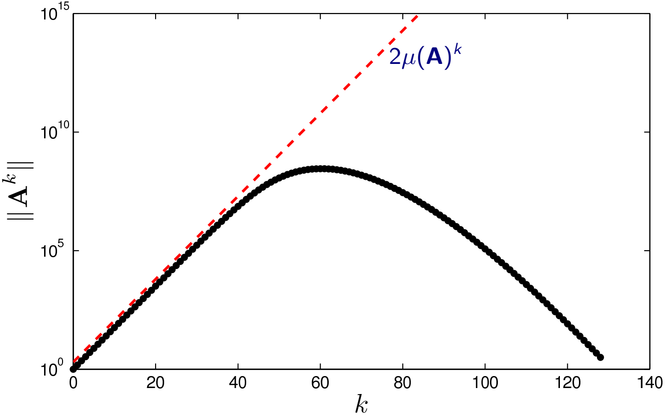
[To be proved during the Exercises.]

Numerical Range and Matrix Powers

Theorem (Berger, 1965)

$$\|\mathbf{A}^k\| \le 2\mu(\mathbf{A})^k.$$

For $\mathbf{A} = \text{tridiag}(0, 1/2, 1)$ of dimension $n = 32$:



Crouzeix's Theorem

The numerical range can bound the size of general matrix functions.

Theorem (Crouzeix, 2007)

Let f be a function analytic on $W(\mathbf{A})$. Then

$$\|f(\mathbf{A})\| \leq 11.08 \max_{z \in W(\mathbf{A})} |f(z)|.$$

Crouzeix has conjectured that the constant 11.08 can be improved...

Crouzeix's Theorem

The numerical range can bound the size of general matrix functions.

Theorem (Crouzeix, 2007)

Let f be a function analytic on $W(\mathbf{A})$. Then

$$\|f(\mathbf{A})\| \leq 11.08 \max_{z \in W(\mathbf{A})} |f(z)|.$$

Crouzeix has conjectured that the constant 11.08 can be improved...

- In some situations, the physical problem restricts the extent of $W(\mathbf{A})$, e.g. coercivity in finite element problems.

Crouzeix's Theorem

The numerical range can bound the size of general matrix functions.

Theorem (Crouzeix, 2007)

Let f be a function analytic on $W(\mathbf{A})$. Then

$$\|f(\mathbf{A})\| \leq 11.08 \max_{z \in W(\mathbf{A})} |f(z)|.$$

Crouzeix has conjectured that the constant 11.08 can be improved...

- In some situations, the physical problem restricts the extent of $W(\mathbf{A})$, e.g. coercivity in finite element problems.
- However, in many other applications, $W(\mathbf{A})$ is so large that it prevents this bound from being meaningful. For example, in these three situations, we want $\max_{z \in W(\mathbf{A})} |f(z)|$ small:
  $f(z) = z^k$, but $W(\mathbf{A})$ contains z with $|z| \geq 1$;
  $f(z) = e^{tz}$, but $W(\mathbf{A})$ contains z with $\operatorname{Re} z \geq 0$;
  $f(z) = p(z)$ with $p(0) = 1$, but $W(\mathbf{A})$ contains $z = 0$.

2(b) Behavior of the resolvent near eigenvalues

Spectral Representation of a Matrix

Theorem (See Kato, 1980)

Any matrix with m distinct eigenvalues can be written in the form

$$\mathbf{A} = \sum_{j=1}^{m} \lambda_j \mathbf{P}_j + \mathbf{D}_j$$

where, for Jordan curves Γ_j surrounding λ_j and no other eigenvalues,

- $\mathbf{P}_j = \frac{1}{2\pi i} \int_{\Gamma_j} (z - \mathbf{A})^{-1} \, \mathrm{d}z$ *is a spectral projector;*
- $\mathbf{D}_j = \frac{1}{2\pi i} \int_{\Gamma_j} (z - \lambda_j)(z - \mathbf{A})^{-1} \, \mathrm{d}z$ *is nilpotent;*
- $\mathbf{P}_j \mathbf{A} = \mathbf{A} \mathbf{P}_j = \lambda_j \mathbf{P}_j + \mathbf{D}_j$;
- $\mathbf{P}_j \mathbf{P}_k = \mathbf{0}$ *if* $j \neq k$;
- $\mathbf{D}_j = \mathbf{0}$ *if* λ_j *is not defective.*

The resolvent plays a fundamental role in the structure of the matrix $\mathbf{A}$.

Functions of a Diagonalizable Matrix

If $\mathbf{A}$ is diagonalizable (i.e., no defective eigenvalues), then

$$\mathbf{A} = \sum_{j=1}^{m} \lambda_j \mathbf{P}_j.$$

If f is any function that is analytic on $\sigma(\mathbf{A})$ and on/inside all contours Γ_j, then

$$f(\mathbf{A}) = \sum_{j=1}^{m} f(\lambda_j) \mathbf{P}_j.$$

Equivalently, for

$$\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{v}}_1^* \\ \vdots \\ \widehat{\mathbf{v}}_n^* \end{bmatrix},$$

we have

$$f(\mathbf{A}) = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{v}}_1^* \\ \vdots \\ \widehat{\mathbf{v}}_n^* \end{bmatrix}.$$

Functions of a Nondiagonalizable Matrix

For all matrices, we have a more general formula. If the Jordan form of $\mathbf{A}$ has m Jordan blocks,

$$\mathbf{A} = \begin{bmatrix} \mathbf{V}_1 & \cdots & \mathbf{V}_m \end{bmatrix} \begin{bmatrix} \mathbf{J}_1 & & \\ & \ddots & \\ & & \mathbf{J}_m \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{V}}_1^* \\ \vdots \\ \widehat{\mathbf{V}}_m^* \end{bmatrix},$$

then

$$f(\mathbf{A}) = \begin{bmatrix} \mathbf{V}_1 & \cdots & \mathbf{V}_m \end{bmatrix} \begin{bmatrix} f(\mathbf{J}_1) & & \\ & \ddots & \\ & & f(\mathbf{J}_m) \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{V}}_1^* \\ \vdots \\ \widehat{\mathbf{V}}_m^* \end{bmatrix},$$

where for

$$\mathbf{J} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}, \qquad f(\mathbf{J}) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(d-1)}(\lambda)}{(d-1)!} \\ & f(\lambda) & \ddots & \vdots \\ & & \ddots & f'(\lambda) \\ & & & f(\lambda) \end{bmatrix}.$$

Cauchy Integral Formula for Matrices

Recall the standard result from complex analysis.

Theorem (Cauchy Integral Formula)

Let Γ be a finite union of Jordan curves containing $a \in C$ in its interior, and suppose f is a function analytic on Γ and its interior. Then

$$f(a) = \frac{1}{2\pi i} \int_\Gamma \frac{f(z)}{z-a}\, \mathrm{d}z.$$

This formula holds for matrices (sometimes called the Dunford–Taylor integral).

Cauchy Integral Formula for Matrices

Recall the standard result from complex analysis.

Theorem (Cauchy Integral Formula)

Let Γ *be a finite union of Jordan curves containing* $a \in C$ *in its interior, and suppose* f *is a function analytic on* Γ *and its interior. Then*

$$f(a) = \frac{1}{2\pi i} \int_\Gamma \frac{f(z)}{z-a}\,\mathrm{d}z.$$

This formula holds for matrices (sometimes called the Dunford–Taylor integral).

Theorem (Cauchy Integral Formula for Matrices)

Let Γ *be a finite union of Jordan curves containing* $\sigma(\mathbf{A})$ *in its interior, and suppose* f *is a function analytic on* Γ *and its interior. Then*

$$f(\mathbf{A}) = \frac{1}{2\pi i} \int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z.$$

Resolvent Bounds

Suppose the eigenvalues of $\mathbf{A} \in \mathbb{C}^{n\times n}$ are distinct.

Apply the previous formula to $f(\zeta) = (z - \zeta)^{-1}$:

$$f(\mathbf{A}) = (z - \mathbf{A})^{-1} = \sum_{j=1}^{n} \frac{1}{z - \lambda_j} \mathbf{P}_j.$$

If λ_j has right eigenvector $\mathbf{v}_j$ and left eigenvector $\widehat{\mathbf{v}}_j$, then

$$\mathbf{P}_j = \frac{\mathbf{v}_j \widehat{\mathbf{v}}_j^*}{\widehat{\mathbf{v}}_j^* \mathbf{v}_j}$$

and the norm of $\mathbf{P}_j$ is

$$\kappa(\lambda_j) := \|\mathbf{P}_j\| = \frac{\|\mathbf{v}_j\| \|\widehat{\mathbf{v}}_j\|}{|\widehat{\mathbf{v}}_j^* \mathbf{v}_j|},$$

which is called the *condition number of the eigenvalue λ_j*.

Hence for z near λ_j,

$$\|(z - \mathbf{A})^{-1}\| \approx \frac{\kappa(\lambda_j)}{|z - \lambda_j|}.$$

Containment Regions for Pseudospectra

Hence for small $\varepsilon > 0$ and diagonalizable matrices, we can approximate

$$\sigma_\varepsilon(\mathbf{A}) \approx \bigcup_{j=1}^{n} \lambda_j + \Delta_{\varepsilon\kappa(\lambda_j)},$$

where $\Delta_r := \{z \in \mathbf{C} : |z| < r\}$.

Theorem (Bauer–Fike, 1963)

If $\mathbf{A} \in C^{n\times n}$ is diagonalizable, then for all $\varepsilon > 0$,

$$\sigma_\varepsilon(\mathbf{A}) \subseteq \bigcup_{j=1}^{n} \lambda_j + \Delta_{n\varepsilon\kappa(\lambda_j)}.$$

Unlike the earlier version of Bauer–Fike, the radii of the disks vary with j.

Pseudospectra near a Defective Eigenvalue

For $\mathbf{A}$ in Jordan form with m distinct Jordan blocks, we write

$$\mathbf{A} = \sum_{j=1}^{m} \mathbf{V}_j \mathbf{J}_j \widehat{\mathbf{V}}_j^*, \qquad f(\mathbf{A}) = \sum_{j=1}^{m} \mathbf{V}_j f(\mathbf{J}_j) \widehat{\mathbf{V}}_j^*.$$

The resolvent follows with $f(\zeta) = (z - \zeta)^{-1}$,

$$(z - \mathbf{A})^{-1} = \sum_{j=1}^{m} \mathbf{V}_j (z - \mathbf{J}_j)^{-1} \widehat{\mathbf{V}}_j^*,$$

Pseudospectra near a Defective Eigenvalue

For $\mathbf{A}$ in Jordan form with m distinct Jordan blocks, we write

$$\mathbf{A} = \sum_{j=1}^{m} \mathbf{V}_j \mathbf{J}_j \widehat{\mathbf{V}}_j^*, \qquad f(\mathbf{A}) = \sum_{j=1}^{m} \mathbf{V}_j f(\mathbf{J}_j) \widehat{\mathbf{V}}_j^*.$$

The resolvent follows with $f(\zeta) = (z-\zeta)^{-1}$,

$$(z-\mathbf{A})^{-1} = \sum_{j=1}^{m} \mathbf{V}_j (z-\mathbf{J}_j)^{-1} \widehat{\mathbf{V}}_j^*,$$

with, for a $d \times d$ Jordan block (index d eigenvalue),

$$\mathbf{J} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}, \qquad (z-\mathbf{J})^{-1} = \begin{bmatrix} \frac{1}{z-\lambda} & \frac{-1}{(z-\lambda)^2} & \cdots & \frac{(-1)^{d+1}}{(z-\lambda)^d} \\ & \frac{1}{z-\lambda} & \ddots & \vdots \\ & & \ddots & \frac{1}{(z-\lambda)^2} \\ & & & \frac{1}{z-\lambda} \end{bmatrix}.$$

Pseudospectra near a Defective Eigenvalue

For $\mathbf{A}$ in Jordan form with m distinct Jordan blocks, we write

$$\mathbf{A} = \sum_{j=1}^{m} \mathbf{V}_j \mathbf{J}_j \widehat{\mathbf{V}}_j^*, \qquad f(\mathbf{A}) = \sum_{j=1}^{m} \mathbf{V}_j f(\mathbf{J}_j) \widehat{\mathbf{V}}_j^*.$$

The resolvent follows with $f(\zeta) = (z-\zeta)^{-1}$,

$$(z-\mathbf{A})^{-1} = \sum_{j=1}^{m} \mathbf{V}_j (z-\mathbf{J}_j)^{-1} \widehat{\mathbf{V}}_j^*,$$

with, for a $d \times d$ Jordan block (index d eigenvalue),

$$\mathbf{J} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}, \qquad (z-\mathbf{J})^{-1} = \begin{bmatrix} \frac{1}{z-\lambda} & \frac{-1}{(z-\lambda)^2} & \cdots & \frac{(-1)^{d+1}}{(z-\lambda)^d} \\ & \frac{1}{z-\lambda} & \ddots & \vdots \\ & & \ddots & \frac{1}{(z-\lambda)^2} \\ & & & \frac{1}{z-\lambda} \end{bmatrix}.$$

So near an eigenvalue of index d, Rellich's perturbation theory requires that the pseudospectrum behave like a disk whose radius scales like $\varepsilon^{1/d}$ as $\varepsilon \longrightarrow 0$.

2(c) Bounds on $\|f(\mathbf{A})\|$ using $\sigma_\varepsilon(\mathbf{A})$

Bounds on $\|f(\mathbf{A})\|$

Suppose $\mathbf{A}$ is diagonalizable, $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$. Then

$$f(\mathbf{A}) = \sum_{j=1}^{n} f(\lambda_j)\mathbf{P}_j = \mathbf{V} f(\boldsymbol{\Lambda})\mathbf{V}^{-1}.$$

This immediately suggests several upper bounds on $\|f(\mathbf{A})\|$:

$$\begin{aligned} \|f(\mathbf{A})\| = \|\mathbf{V} f(\boldsymbol{\Lambda})\mathbf{V}^{-1}\| &\leq \|\mathbf{V}\|\|\mathbf{V}^{-1}\| \max_{\lambda \in \sigma(\mathbf{A})} |f(\lambda)| \\ &= \kappa(\mathbf{V}) \max_{\lambda \in \sigma(\mathbf{A})} |f(\lambda)|; \end{aligned}$$

$$\|f(\mathbf{A})\| \leq \sum_{j=1}^{n} |f(\lambda_j)|\|\mathbf{P}_j\| = \sum_{j=1}^{n} \kappa(\lambda_j)|f(\lambda_j)|.$$

We seek bounds that provide a more flexible way of handling nonnormality.

Pseudospectral Bounds on $\|f(\mathbf{A})\|$

Theorem (Cauchy Integral Formula for Matrices)

$$f(\mathbf{A}) = \frac{1}{2\pi i} \int_\Gamma f(z)(z - \mathbf{A})^{-1}\, \mathrm{d}z.$$

Take norms of the expression for $f(\mathbf{A})$:

$$\|f(\mathbf{A})\| \;=\; \left\| \frac{1}{2\pi i} \int_\Gamma f(z)(z - \mathbf{A})^{-1}\, \mathrm{d}z \right\| \;\leq\; \frac{1}{2\pi} \int_\Gamma |f(z)| \|(z - \mathbf{A})^{-1}\| |dz|.$$

Pseudospectral Bounds on $\|f(\mathbf{A})\|$

Theorem (Cauchy Integral Formula for Matrices)

$$f(\mathbf{A}) = \frac{1}{2\pi i}\int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z.$$

Take norms of the expression for $f(\mathbf{A})$:

$$\|f(\mathbf{A})\| \;=\; \left\|\frac{1}{2\pi i}\int_\Gamma f(z)(z-\mathbf{A})^{-1}\,\mathrm{d}z\right\| \;\le\; \frac{1}{2\pi}\int_\Gamma |f(z)|\|(z-\mathbf{A})^{-1}\||dz|.$$

Now pick Γ to be the boundary of $\sigma_\varepsilon(\mathbf{A})$

$$\begin{aligned}
\|f(\mathbf{A})\| &\le \frac{1}{2\pi}\int_{\partial\sigma_\varepsilon} |f(z)|\|(z-\mathbf{A})^{-1}\||dz| \\
&= \frac{1}{2\pi\varepsilon}\int_{\partial\sigma_\varepsilon} |f(z)||dz| \\
&\le \frac{1}{2\pi\varepsilon}\sup_{z\in\sigma_\varepsilon(\mathbf{A})}|f(z)|\int_{\partial\sigma_\varepsilon}|dz| \;\le\; \frac{L_\varepsilon}{2\pi\varepsilon}\sup_{z\in\sigma_\varepsilon(\mathbf{A})}|f(z)|
\end{aligned}$$

where L_ε denotes the arc-length of $\partial\sigma_\varepsilon(\mathbf{A})$ [Trefethen 1990].

Pseudospectral Bounds on $\|f(\mathbf{A})\|$

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \;\le\; \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

Some key observations:

- This should be regarded as a *family of bounds* that vary with ε;
- The best choice for ε will depend on the problem;
- Sometimes it is excellent; usually it is decent; on occasion it is poor;
- The choice of ε has *nothing* to do with rounding errors; do not expect the bound to be most descriptive when $\varepsilon = \varepsilon_{\mathrm{mach}}$ or $\varepsilon = \|\mathbf{A}\|\varepsilon_{\mathrm{mach}}$.

Bounds on the Matrix Exponential

To understand behavior of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ [and LTI control systems], we wish to use pseudospectra to bound $\|e^{t\mathbf{A}}\|$.

Definition

The *spectral abscissa* is the rightmost point in the spectrum:

$$\alpha(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} \operatorname{Re} z.$$

The *ε-pseudospectral abscissa* is the supremum of the real parts of $z \in \sigma_\varepsilon(\mathbf{A})$:

$$\alpha_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} \operatorname{Re} z.$$

Applying the Cauchy integral bound to $f(z) = e^{tz}$ gives an upper bound.

Bounds on the Matrix Exponential

To understand behavior of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ [and LTI control systems], we wish to use pseudospectra to bound $\|e^{t\mathbf{A}}\|$.

Definition

The *spectral abscissa* is the rightmost point in the spectrum:

$$\alpha(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} \operatorname{Re} z.$$

The *ε-pseudospectral abscissa* is the supremum of the real parts of $z \in \sigma_\varepsilon(\mathbf{A})$:

$$\alpha_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} \operatorname{Re} z.$$

Applying the Cauchy integral bound to $f(z) = e^{tz}$ gives an upper bound.

Theorem (Upper Bound on $\|e^{t\mathbf{A}}\|$)

For any $\mathbf{A} \in C^{n\times n}$ and $\varepsilon > 0$,

$$\|e^{t\mathbf{A}}\| \leq \frac{L_\varepsilon}{2\pi\varepsilon} e^{t\alpha_\varepsilon(\mathbf{A})},$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

Upper Bound on the Matrix Exponential

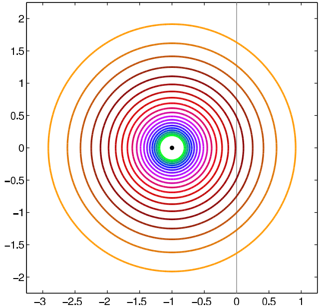
$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

Question: can you estimate $\sigma_\varepsilon(\mathbf{A})$ for this matrix?

Upper Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

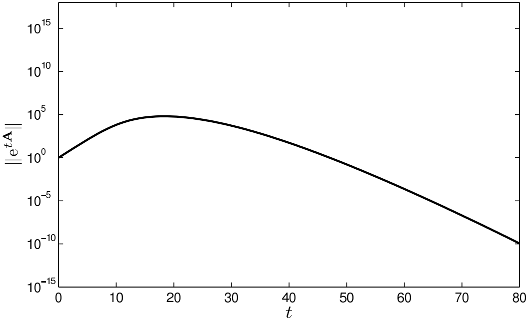
Question: can you estimate $\sigma_\varepsilon(\mathbf{A})$ for this matrix?



$\sigma_\varepsilon(\mathbf{A})$ for $\varepsilon = 10^{-20}, 10^{-19}, \ldots, 10^{-1}$

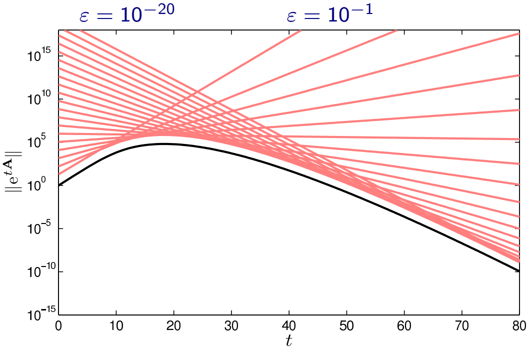
Upper Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Upper Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Computing the Bound Based on the Cauchy Integral Formula

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \;\leq\; \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

To compute/estimate this bound:

- Compute resolvent norm on a grid that contains all $\sigma_\varepsilon(\mathbf{A})$ of interest, e.g., using EigTool: `[x,y,z] = eigtool(A);`

Computing the Bound Based on the Cauchy Integral Formula

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \;\le\; \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

To compute/estimate this bound:

- Compute resolvent norm on a grid that contains all $\sigma_\varepsilon(\mathbf{A})$ of interest, e.g., using EigTool: `[x,y,z] = eigtool(A);`
- Send resolvent norm data to a contour plotting routine, e.g., `c = contour(x,y,z,epsilon);`

Computing the Bound Based on the Cauchy Integral Formula

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \;\le\; \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

To compute/estimate this bound:

- Compute resolvent norm on a grid that contains all $\sigma_\varepsilon(\mathbf{A})$ of interest, e.g., using EigTool: `[x,y,z] = eigtool(A);`
- Send resolvent norm data to a contour plotting routine, e.g., `c = contour(x,y,z,epsilon);`
- Extract the contour data from `c` for each value of `epsilon`;

Computing the Bound Based on the Cauchy Integral Formula

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \;\leq\; \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

To compute/estimate this bound:

- Compute resolvent norm on a grid that contains all $\sigma_\varepsilon(\mathbf{A})$ of interest, e.g., using EigTool: `[x,y,z] = eigtool(A);`
- Send resolvent norm data to a contour plotting routine, e.g., `c = contour(x,y,z,epsilon);`
- Extract the contour data from `c` for each value of `epsilon`;
- Estimate L_ε using linear interpolation between points on each contour;

Computing the Bound Based on the Cauchy Integral Formula

Theorem

Let f be analytic on $\sigma_\varepsilon(\mathbf{A})$ for some $\varepsilon > 0$. Then

$$\|f(\mathbf{A})\| \;\leq\; \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$.

To compute/estimate this bound:

- Compute resolvent norm on a grid that contains all $\sigma_\varepsilon(\mathbf{A})$ of interest, e.g., using EigTool: `[x,y,z] = eigtool(A);`
- Send resolvent norm data to a contour plotting routine, e.g., `c = contour(x,y,z,epsilon);`
- Extract the contour data from `c` for each value of `epsilon`;
- Estimate L_ε using linear interpolation between points on each contour;
- Using the maximum principle, $\sup_{z\in\sigma_\varepsilon(\mathbf{A})} |f(z)| = \sup_{z\in\partial\sigma_\varepsilon(\mathbf{A})} |f(z)|$, so take the max of $|f(z)|$ evaluated at all points z on each contour.

Lower Bounds on the Matrix Exponential

We would like to *guarantee* the potential for transient growth.

Theorem (Lower bound on $\|e^{t\mathbf{A}}\|$)

Suppose $\alpha(\mathbf{A}) < 0$*. Then for all* $\varepsilon > 0$*,*

$$\sup_{t \ge 0} \|e^{t\mathbf{A}}\| \ge \frac{\alpha_\varepsilon(\mathbf{A})}{\varepsilon}.$$

Lower Bounds on the Matrix Exponential

We would like to *guarantee* the potential for transient growth.

Theorem (Lower bound on $\|e^{t\mathbf{A}}\|$)

Suppose $\alpha(\mathbf{A}) < 0$. Then for all $\varepsilon > 0$,

$$\sup_{t\geq 0} \|e^{t\mathbf{A}}\| \geq \frac{\alpha_\varepsilon(\mathbf{A})}{\varepsilon}.$$

Proof. Recall the formula for the Laplace transform of $e^{t\alpha}$: For $s > \alpha$,

$$\int_0^\infty e^{t\alpha} e^{-st} \, dt = \frac{1}{s-\alpha}.$$

This formula generalizes for matrices, e.g., when $\alpha(\mathbf{A}) < 0$ and $\operatorname{Re} z > 0$:

$$\int_0^\infty e^{t\mathbf{A}} e^{-zt} \, dt = (z - \mathbf{A})^{-1}.$$

Lower Bounds on the Matrix Exponential

We would like to *guarantee* the potential for transient growth.

Theorem (Lower bound on $\|e^{t\mathbf{A}}\|$)

Suppose $\alpha(\mathbf{A}) < 0$. *Then for all* $\varepsilon > 0$,

$$\sup_{t\geq 0} \|e^{t\mathbf{A}}\| \geq \frac{\alpha_\varepsilon(\mathbf{A})}{\varepsilon}.$$

Proof.

$$\int_0^\infty e^{t\mathbf{A}} e^{-zt} \mathrm{d}t = (z - \mathbf{A})^{-1}.$$

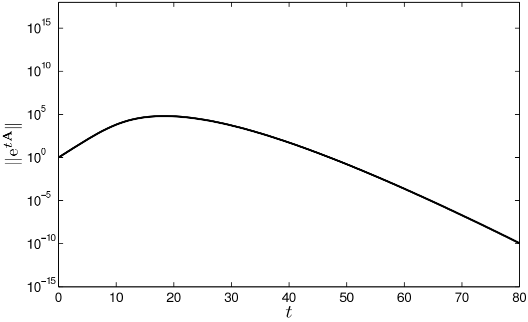
Since $\mathbf{A}$ is stable, let $M := \sup_{t\geq 0} \|e^{t\mathbf{A}}\|$. Then for any $z \in \sigma_\varepsilon(\mathbf{A})$, $\mathrm{Re}\, z > 0$:

$$\begin{aligned} \frac{1}{\varepsilon} < \|(z - \mathbf{A})^{-1}\| &= \left\| \int_0^\infty e^{t\mathbf{A}} e^{-zt}\, \mathrm{d}t \right\| \\ &\leq \int_0^\infty \|e^{t\mathbf{A}}\| |e^{-zt}|\, \mathrm{d}t \leq M \int_0^\infty e^{-(\mathrm{Re}\, z)t}\, \mathrm{d}t = \frac{M}{\mathrm{Re}\, z}. \end{aligned}$$

Hence $M \geq (\mathrm{Re}\, z)/\varepsilon$. Take the sup over all $z \in \sigma_\varepsilon(\mathbf{A})$ to get the bound. $\square$

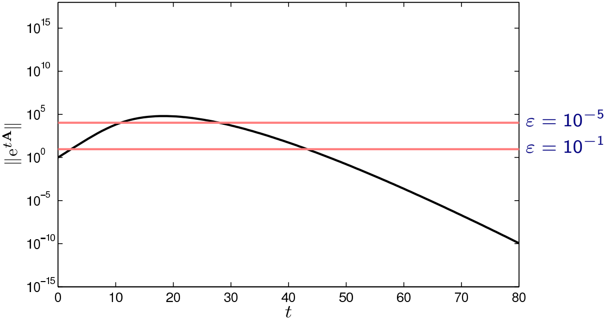
Lower Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Lower Bound on the Matrix Exponential

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & & & \\ & -1 & \ddots & & \\ & & \ddots & 2 & \\ & & & -1 & 2 \\ & & & & -1 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Zabczyk's Example

Pseudospectra illuminate an important example in semigroup theory.

Zabczyk (1975) proposed a semi-infinite matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{J}_1 & & & & \\ & \mathbf{J}_2 & & & \\ & & \mathbf{J}_3 & & \\ & & & \mathbf{J}_4 & \\ & & & & \ddots \end{bmatrix},$$

an unbounded operator on $\ell^2(\mathbf{N})$, where

$$\mathbf{J}_k = \begin{bmatrix} -1+\mathrm{i}k & 1 & & \\ & -1+\mathrm{i}k & \ddots & \\ & & \ddots & 1 \\ & & & -1+\mathrm{i}k \end{bmatrix} \in \mathbf{C}^{k\times k}.$$

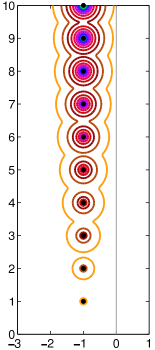
Zabczyk's Example

$$\mathbf{A} = \begin{bmatrix} \mathbf{J}_1 & & & & \\ & \mathbf{J}_2 & & & \\ & & \mathbf{J}_3 & & \\ & & & \mathbf{J}_4 & \\ & & & & \ddots \end{bmatrix}$$

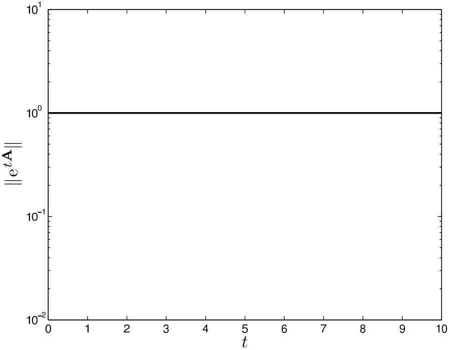
Spectral abscissa: $\alpha(\mathbf{A}) = -1$

However, $\|e^{t\mathbf{A}}\| = 1$ for all $t \geq 0$.

Growth bound *not* determined by $\alpha(\mathbf{A}) = -1$.



$\sigma_\varepsilon(\mathbf{A})$ for $\varepsilon = 10^{-10}, \ldots, 10^{-1}$



Matrix Powers

Definition

The *spectral radius* is the largest point in the spectrum:

$$\rho(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} |z|.$$

The *ε-pseudospectral radius* is the supremum of magnitudes of points in $\sigma_\varepsilon(\mathbf{A})$:

$$\rho_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |z|.$$

Matrix Powers

Definition

The *spectral radius* is the largest point in the spectrum:

$$\rho(\mathbf{A}) := \max_{\lambda \in \sigma(\mathbf{A})} |z|.$$

The *ε-pseudospectral radius* is the supremum of magnitudes of points in $\sigma_\varepsilon(\mathbf{A})$:

$$\rho_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |z|.$$

Theorem (Upper Bound)

For any $\mathbf{A} \in C^{n \times n}$ and $\varepsilon > 0$,

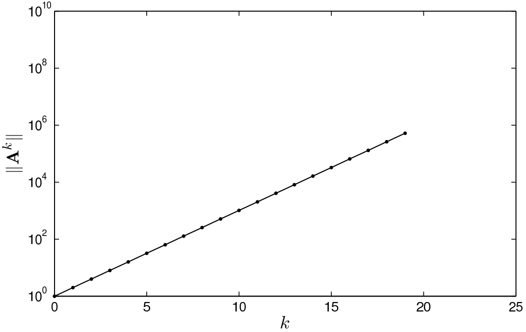
$$\|\mathbf{A}^k\| \leq \frac{\rho_\varepsilon(\mathbf{A})^{k+1}}{\varepsilon}$$

Proof: Apply the Cauchy integral bound, taking Γ to be the circle of radius $\rho_\varepsilon(\mathbf{A})$ centered at the origin. $\square$

Upper Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbb{C}^{20\times 20}.$$

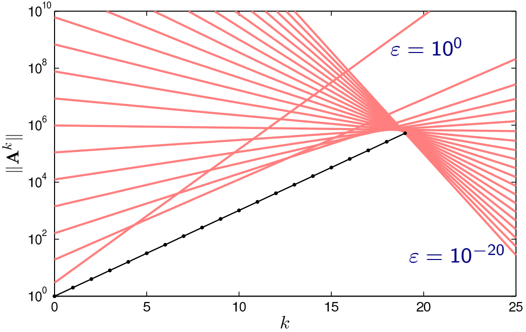
$\mathbf{A}^{20} = \mathbf{0}$, but lower powers show transient growth.



Upper Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$

$\mathbf{A}^{20} = \mathbf{0}$, but lower powers show transient growth.



Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$. *Then for all* $\varepsilon > 0$,

$$\sup_{k \geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$*. Then for all* $\varepsilon > 0$*,*

$$\sup_{k \geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

Proof. Since $\rho(\mathbf{A}) < 1$, $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$. Let M denote the maximum value of $\|\mathbf{A}^k\|$, $k \geq 0$, and suppose $z \in \sigma_\varepsilon(\mathbf{A})$ for $|z| > 1$. Then

$$\frac{1}{\varepsilon} \;<\; \|(z - \mathbf{A})^{-1}\| \;=\; \Big\| \frac{1}{z}\Big(1 + \frac{1}{z}\mathbf{A} + \frac{1}{z^2}\mathbf{A}^2 + \cdots\Big)\Big\|$$

Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$*. Then for all* $\varepsilon > 0$*,*

$$\sup_{k\geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

Proof. Since $\rho(\mathbf{A}) < 1$, $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$. Let M denote the maximum value of $\|\mathbf{A}^k\|$, $k \geq 0$, and suppose $z \in \sigma_\varepsilon(\mathbf{A})$ for $|z| > 1$. Then

$$\begin{aligned} \frac{1}{\varepsilon} < \|(z - \mathbf{A})^{-1}\| &= \left\| \frac{1}{z}\left(1 + \frac{1}{z}\mathbf{A} + \frac{1}{z^2}\mathbf{A}^2 + \cdots\right) \right\| \\ &\leq \frac{1}{|z|}\left(M + \frac{M}{|z|} + \frac{M}{|z|^2} + \cdots\right) \end{aligned}$$

Lower Bounds on Matrix Powers

Theorem (Lower bound on power growth)

Suppose $\rho(\mathbf{A}) < 1$. Then for all $\varepsilon > 0$,

$$\sup_{k\geq 0} \|\mathbf{A}^k\| \geq \frac{\rho_\varepsilon(\mathbf{A}) - 1}{\varepsilon}.$$

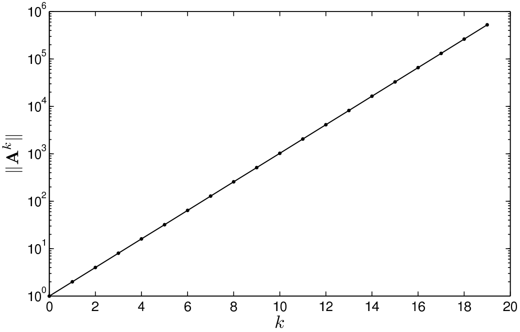
Proof. Since $\rho(\mathbf{A}) < 1$, $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$. Let M denote the maximum value of $\|\mathbf{A}^k\|$, $k \geq 0$, and suppose $z \in \sigma_\varepsilon(\mathbf{A})$ for $|z| > 1$. Then

$$\begin{aligned}
\frac{1}{\varepsilon} < \|(z-\mathbf{A})^{-1}\| &= \left\|\frac{1}{z}\left(1 + \frac{1}{z}\mathbf{A} + \frac{1}{z^2}\mathbf{A}^2 + \cdots\right)\right\| \\
&\leq \frac{1}{|z|}\left(M + \frac{M}{|z|} + \frac{M}{|z|^2} + \cdots\right) \\
&= \frac{M}{|z|}\,\frac{1}{1 - 1/|z|} = \frac{M}{|z| - 1}.
\end{aligned}$$

Rearrange to obtain $M \geq (|z| - 1)/\varepsilon$ for all $z \in \sigma_\varepsilon(\mathbf{A})$ with $|z| > 1$.
Take the supremum of $|z|$ over all $z \in \sigma_\varepsilon(\mathbf{A})$ to get the result. $\square$

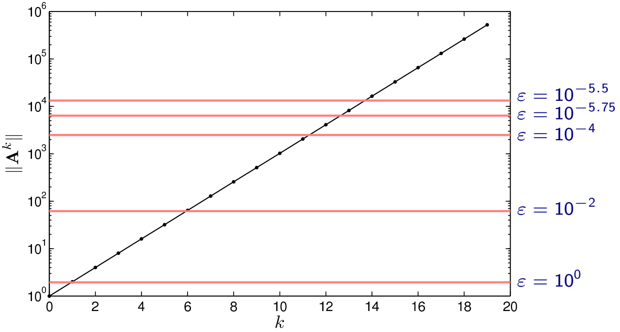
Lower Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Lower Bound on Matrix Powers

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & & & \\ & 0 & \ddots & & \\ & & \ddots & 2 & \\ & & & 0 & 2 \\ & & & & 0 \end{bmatrix} \in \mathbf{C}^{20\times 20}.$$



Pseudospectra are not a Panacea

Key question: "Do pseudospectra determine behavior of a matrix?"
[Greenbaum & Trefethen, 1993].

Greenbaum and Trefethen define "behavior" to mean "norms of polynomials".

They prove that pseudospectra *do not* determine behavior.

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & \alpha \\ & & & & 0 \end{bmatrix}.$$

If $\alpha \in (1, \sqrt{2}]$, then $\sigma_\varepsilon(\mathbf{A}_1) = \sigma_\varepsilon(\mathbf{A}_2)$ *for all* $\varepsilon > 0$,

$$1 = \|\mathbf{A}_1\| \neq \|\mathbf{A}_2\| = \sqrt{2}.$$

Pseudospectra are not a Panacea (in the 2-norm)

Key question: "Do pseudospectra determine behavior of a matrix?" [Greenbaum & Trefethen, 1993].

Greenbaum and Trefethen define "behavior" to mean "norms of polynomials".

They prove that pseudospectra *do not* determine behavior.

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & \alpha \\ & & & & 0 \end{bmatrix}.$$

If $\alpha \in (1, \sqrt{2}]$, then $\sigma_\varepsilon(\mathbf{A}_1) = \sigma_\varepsilon(\mathbf{A}_2)$ *for all* $\varepsilon > 0$,

$$1 = \|\mathbf{A}_1\| \neq \|\mathbf{A}_2\| = \sqrt{2}.$$

However: Greenbaum and Trefethen [1993] do show that pseudospectra determine behavior in the Frobenius norm ($\sigma_\varepsilon(\mathbf{A})$ defined via resolvent norms).

Pseudospectra are not a Panacea in the 2-norm

Ransford and colleagues have constructed a number of surprising examples.

A sequence $\eta_2, \ldots, \eta_m$ is *submultiplicative* if $\eta_{j+k} \leq \eta_j \eta_k$
for all $j, k \in \{2, \ldots, m\}$ such that $j + k \leq m$.

Theorem (Ransford, 2007)

*Suppose $\alpha_2, \ldots, \alpha_m$ and $\beta_2, \ldots, \beta_m$ are submultiplicative sequences.
Then there exist matrices $\mathbf{A}_1, \mathbf{A}_2 \in C^{(2m+3)\times(2m+3)}$ with $\sigma_\varepsilon(\mathbf{A}_1) = \sigma_\varepsilon(\mathbf{A}_2)$
for all $\varepsilon > 0$, yet $\|\mathbf{A}_1^k\| = \alpha_k$ and $\|\mathbf{A}_2^k\| = \beta_k$ for $k = 2, \ldots, m$.*

Pseudospectra are not a Panacea in the 2-norm

Ransford and Rostand (2011) construct matrices with *simple eigenvalues*

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 180 & -360 & 0 & 0 \\ -90+120\sqrt{5} & 180+60\sqrt{5} & 120\sqrt{5} & 0 \\ 450 & -180 & -360 & 216\sqrt{5} \end{bmatrix}$$

and

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 120 & -360 & 0 & 0 \\ 45\sqrt{130}-15\sqrt{26} & 45\sqrt{26}+15\sqrt{130} & 120\sqrt{5} & 0 \\ 30\sqrt{130} & 10\sqrt{130} & 80\sqrt{5} & 216\sqrt{5} \end{bmatrix}$$

Theorem (Ransford and Rostand, 2011)

The above matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ have superidentical pseudospectra (i.e., all singular values of $z-\mathbf{A}_1$ and $z-\mathbf{A}_2$ match for all $z \in C$), yet $\|\mathbf{A}_1^2\| \neq \|\mathbf{A}_2^2\|$.

2(d) Computing the pseudospectral abscissa $\alpha_\varepsilon(\mathbf{A})$ and radius $\rho_\varepsilon(\mathbf{A})$

Distance to Instability

The spectral abscissa and spectral radius

$$\alpha(\mathbf{A}) := \sup_{z \in \sigma(\mathbf{A})} \operatorname{Re} z, \qquad \rho(\mathbf{A}) := \sup_{z \in \sigma(\mathbf{A})} |z|$$

are one important measure of stability, but these quantities can both be sensitive to perturbations. We seek a more robust measure.

Definition (Distance to Instability)

The *distance to instability* of a stable system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is

$$\min\{\|\mathbf{E}\| : iy \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } y \in \mathbb{R}\}.$$

The *distance to instability* of a stable system $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}$ is

$$\min\{\|\mathbf{E}\| : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } |z| = 1\}.$$

See also [Hinrichsen and Pritchard, 1990...], especially in the context of structured perturbations.

Distance to Instability

$$\min\{\|\mathbf{E}\| : iy \in \sigma(\mathbf{A}+\mathbf{E}) \text{ for some } y \in \mathbb{R}\}$$

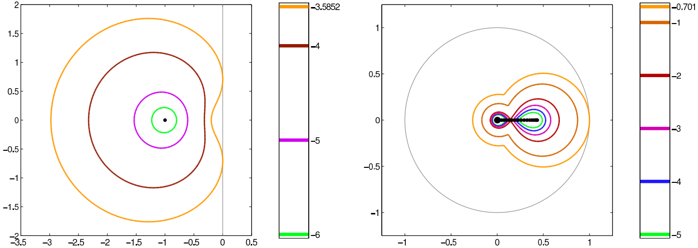
$$\min\{\|\mathbf{E}\| : z \in \sigma(\mathbf{A}+\mathbf{E}) \text{ for some } |z| = 1\}$$

The continuous-time distance to instability is thus

$$\sup\{\varepsilon > 0 : \sigma_\varepsilon(\mathbf{A}) \text{ is contained in the left half plane}\}$$

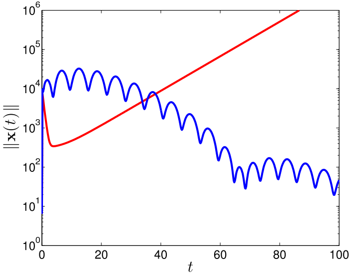
and the discrete-time distance to instability is thus

$$\sup\{\varepsilon > 0 : \sigma_\varepsilon(\mathbf{A}) \text{ is contained in the unit circle}\}.$$



Optimizing $\alpha_\varepsilon(\mathbf{A})$ and $\rho_\varepsilon(\mathbf{A})$ for Robust Stability

- As seen with the Boeing 767 example yesterday, optimizing $\alpha(\mathbf{A})$ can give asymptotically stable models with considerable transient growth.



- Burke, Lewis, Overton, and Mengi have undertaken an extensive program:
  (a) to characterize smoothness of $\sigma_\varepsilon(\mathbf{A})$, $\alpha_\varepsilon(\mathbf{A})$, and $\rho_\varepsilon(\mathbf{A})$
  (b) and develop algorithms to compute $\alpha_\varepsilon(\mathbf{A})$ and $\rho_\varepsilon(\mathbf{A})$
  (c) toward optimization of these quantities within a given class of models.

Here we focus on efficient algorithms for computing

$$\alpha_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} \operatorname{Re} z, \qquad \rho_\varepsilon(\mathbf{A}) := \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |z|.$$

Toward Computing $\alpha_\varepsilon(\mathbf{A})$ and $\rho_\varepsilon(\mathbf{A})$

These algorithms will be based on this fundamental result.

Theorem (Byers, 1988)

The matrix $(x + iy) - \mathbf{A}$ *has a singular value* ε *if and only if* iy *is an eigenvalue of the Hamiltonian matrix*

$$\begin{bmatrix} x - \mathbf{A}^* & \varepsilon \mathbf{I} \\ -\varepsilon \mathbf{I} & \mathbf{A} - x \end{bmatrix}.$$

Question: Since this matrix has at most $2n$ eigenvalues, what can you deduce about the boundary of $\sigma_\varepsilon(\mathbf{A})$?

Toward Computing $\alpha_\varepsilon(\mathbf{A})$ and $\rho_\varepsilon(\mathbf{A})$

These algorithms will be based on this fundamental result.

Theorem (Byers, 1988)

The matrix $(x + iy) - \mathbf{A}$ has a singular value ε if and only if iy is an eigenvalue of the Hamiltonian matrix

$$\begin{bmatrix} x - \mathbf{A}^* & \varepsilon \mathbf{I} \\ -\varepsilon \mathbf{I} & \mathbf{A} - x \end{bmatrix}.$$

Question: Since this matrix has at most $2n$ eigenvalues, what can you deduce about the boundary of $\sigma_\varepsilon(\mathbf{A})$?

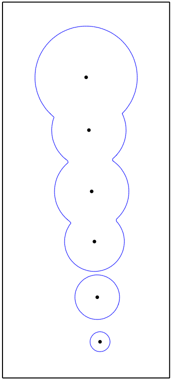
The boundary $\partial\sigma_\varepsilon(\mathbf{A})$ intersects any vertical line in $\mathbf{C}$ at no more than $2n$ distinct points. Even more, this implies that $\partial\sigma_\varepsilon(\mathbf{A})$ contains no vertical line segment – nor, by substituting $e^{i\theta}\mathbf{A}$ for $\mathbf{A}$, a segment of any straight line.

Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

This algorithm was proposed by [Burke, Lewis, Overton 2003].

Given a fixed value of $\varepsilon > 0$ and the corresponding pseudospectrum $\sigma_\varepsilon(\mathbf{A})$:

1. Find λ, the rightmost eigenvalue of $\mathbf{A}$.

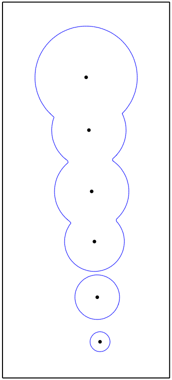


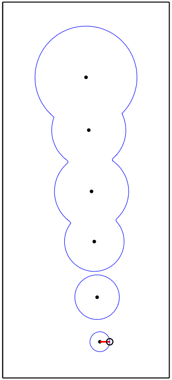
Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

This algorithm was proposed by [Burke, Lewis, Overton 2003].

Given a fixed value of $\varepsilon > 0$ and the corresponding pseudospectrum $\sigma_\varepsilon(\mathbf{A})$:

1. Find λ, the rightmost eigenvalue of $\mathbf{A}$.
2. Find the rightmost point in $\partial\sigma_\varepsilon(\mathbf{A})$ that intersects the *horizontal* line $\{z \in \mathbf{C} : \operatorname{Im} z = \operatorname{Im} \lambda\}$ through λ. Call this point z_1.

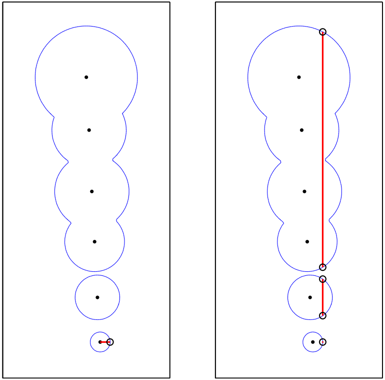




Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *vertical* line through z_k, $\{z \in \mathbf{C} : \operatorname{Re} z = \operatorname{Re} z_k\}$. These points determine the intervals where $\sigma_\varepsilon(\mathbf{A})$ intersects the line.

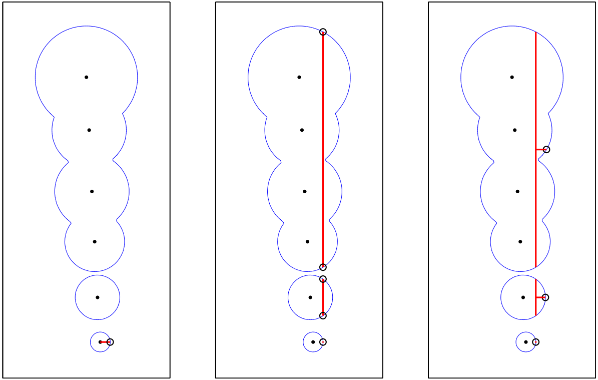


Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *vertical* line through z_k, $\{z \in \mathbf{C} : \operatorname{Re} z = \operatorname{Re} z_k\}$. These points determine the intervals where $\sigma_\varepsilon(\mathbf{A})$ intersects the line.

4. Find the midpoints of these vertical segments. From each midpoint, search horizontally to find $\partial\sigma_\varepsilon(\mathbf{A})$. Call the rightmost of these points z_{k+1}.

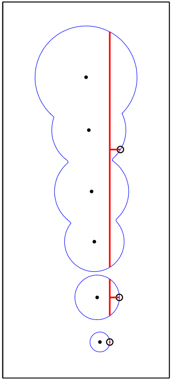


Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *vertical* line through z_k, $\{z \in \mathbf{C} : \mathrm{Re}\, z = \mathrm{Re}\, z_k\}$. These points determine the intervals where $\sigma_\varepsilon(\mathbf{A})$ intersects the line.
4. Find the midpoints of these vertical segments. From each midpoint, search horizontally to find $\partial\sigma_\varepsilon(\mathbf{A})$. Call the rightmost of these points z_{k+1}.

$k = 1$



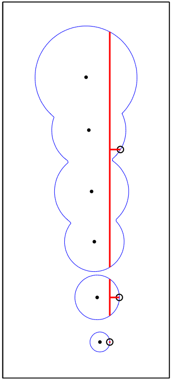
Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

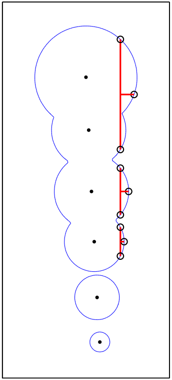
3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *vertical* line through z_k, $\{z \in \mathbf{C} : \operatorname{Re} z = \operatorname{Re} z_k\}$. These points determine the intervals where $\sigma_\varepsilon(\mathbf{A})$ intersects the line.

4. Find the midpoints of these vertical segments. From each midpoint, search horizontally to find $\partial\sigma_\varepsilon(\mathbf{A})$. Call the rightmost of these points z_{k+1}.

$k = 1$



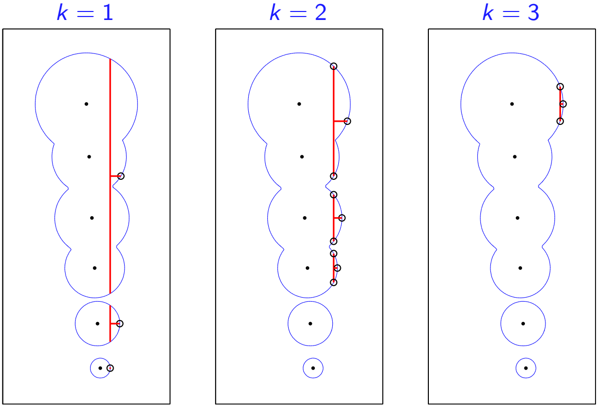
$k = 2$



Criss-Cross Algorithm for $\alpha_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *vertical* line through z_k, $\{z \in \mathbf{C} : \mathrm{Re}\, z = \mathrm{Re}\, z_k\}$. These points determine the intervals where $\sigma_\varepsilon(\mathbf{A})$ intersects the line.
4. Find the midpoints of these vertical segments. From each midpoint, search horizontally to find $\partial\sigma_\varepsilon(\mathbf{A})$. Call the rightmost of these points z_{k+1}.



Criss–Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

This algorithm will rely on the following adaptation of the result of Byers.

Theorem (Mengi & Overton, 2005)

The matrix $re^{i\theta} - \mathbf{A}$ *has a singular value* ε *if and only if* $i\,r$ *is an eigenvalue of the Hamiltonian matrix*

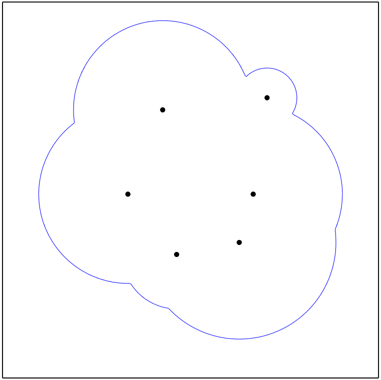
$$\begin{bmatrix} ie^{i\theta}\mathbf{A}^* & -\varepsilon\mathbf{I} \\ \varepsilon\mathbf{I} & ie^{-i\theta}\mathbf{A} \end{bmatrix}.$$

Criss-Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

This algorithm was proposed by [Mengi & Overton, 2005].

Given a fixed value of $\varepsilon > 0$ and the corresponding pseudospectrum $\sigma_\varepsilon(\mathbf{A})$:

1. Find λ, the eigenvalue of $\mathbf{A}$ having largest magnitude.

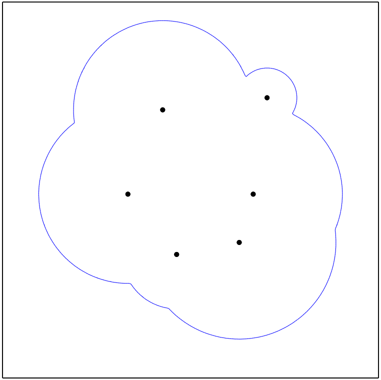


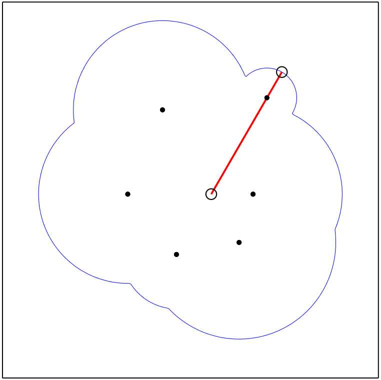
Criss-Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

This algorithm was proposed by [Mengi & Overton, 2005].

Given a fixed value of $\varepsilon > 0$ and the corresponding pseudospectrum $\sigma_\varepsilon(\mathbf{A})$:

1. Find λ, the eigenvalue of $\mathbf{A}$ having largest magnitude.
2. Find the largest magnitude point in $\partial\sigma_\varepsilon(\mathbf{A})$ that intersects the line connecting the origin to λ, i.e., $\{z \in \mathbf{C} : \arg z = \arg \lambda\}$. Call this point z_1.

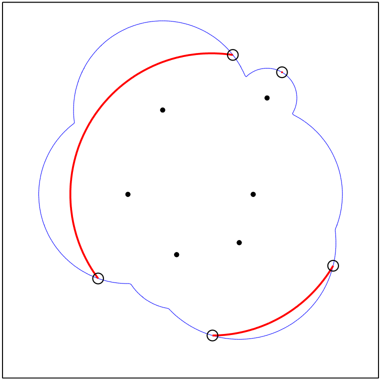




Criss-Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

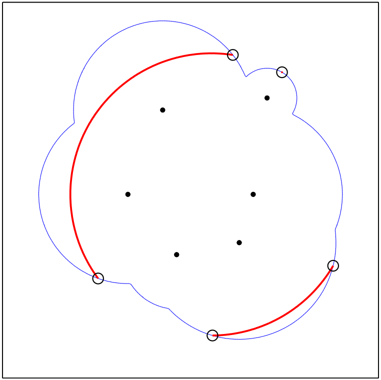
3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *circle* of radius $|z_k|$.
   These points determine arcs where $\sigma_\varepsilon(\mathbf{A})$ intersects this circle.
4. Find the midpoints of each arc. From each arc, search radially for $\partial\sigma_\varepsilon(\mathbf{A})$.
   Call the point of intersection with largest magnitude z_{k+1}.

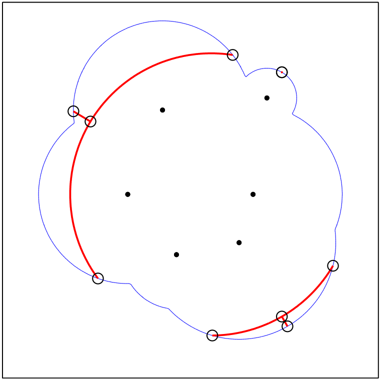


Criss-Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *circle* of radius $|z_k|$. These points determine arcs where $\sigma_\varepsilon(\mathbf{A})$ intersects this circle.
4. Find the midpoints of each arc. From each arc, search radially for $\partial\sigma_\varepsilon(\mathbf{A})$. Call the point of intersection with largest magnitude z_{k+1}.

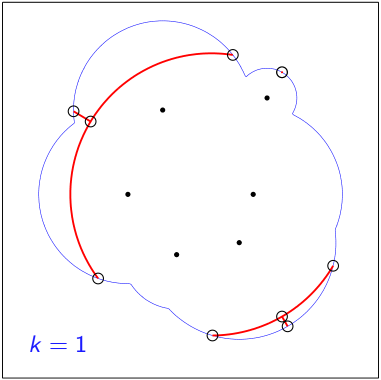




Criss-Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *circle* of radius $|z_k|$.
   These points determine arcs where $\sigma_\varepsilon(\mathbf{A})$ intersects this circle.
4. Find the midpoints of each arc. From each arc, search radially for $\partial\sigma_\varepsilon(\mathbf{A})$.
   Call the point of intersection with largest magnitude z_{k+1}.



Criss-Cross Algorithm for $\rho_\varepsilon(\mathbf{A})$

For $k = 1, 2, \ldots$ until convergence

3. Find all points where $\partial\sigma_\varepsilon(\mathbf{A})$ intersects the *circle* of radius $|z_k|$.
These points determine arcs where $\sigma_\varepsilon(\mathbf{A})$ intersects this circle.

4. Find the midpoints of each arc. From each arc, search radially for $\partial\sigma_\varepsilon(\mathbf{A})$.
Call the point of intersection with largest magnitude z_{k+1}.