Where in the complex plane does a matrix live?

(A question of L. N. Trefethen)

Connections Between Matrix Theory and Complex Analysis

What can eigenvalues do?

- If A is **normal** (e.g., real symmetric) or **near normal** (well-conditioned eigenvectors) then eigenvalues describe behavior in spectral norm perfectly or almost perfectly:

$$\|f(A)\| \approx \max_{\lambda \in \sigma(A)} |f(\lambda)|.$$

- Even if A is highly **nonnormal** (e.g., not diagonalizable, or diagonalizable but with eigenvectors that are almost linearly dependent), eigenvalues determine the *asymptotic* behavior of many functions of A:

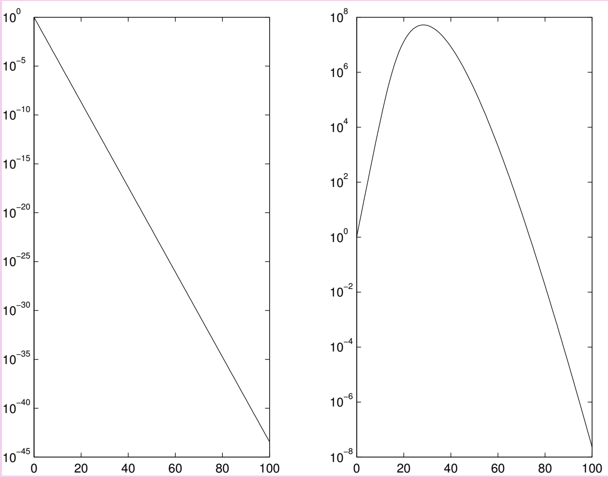
$$\|A^k\| \to 0 \quad \text{as } k \to \infty \quad \text{iff } \rho(A) < 1.$$

$$\|e^{tA}\| \to 0 \quad \text{as } t \to \infty \quad \text{iff } \mathrm{Re}(\sigma(A)) < 0.$$

What can eigenvalues NOT do?

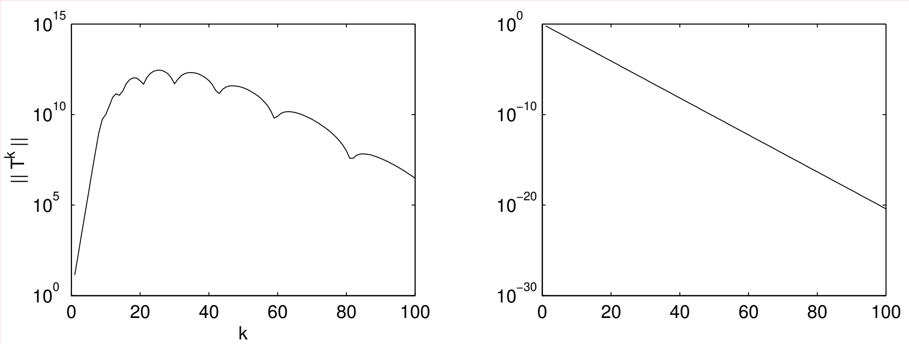
- e^{tA}: Determines the stability of $y' = Ay$.

  $\lim_{t\to\infty} \|e^{tA}\| = 0$ if and only if the eigenvalues of A have negative real parts. But eigenvalues alone cannot distinguish:



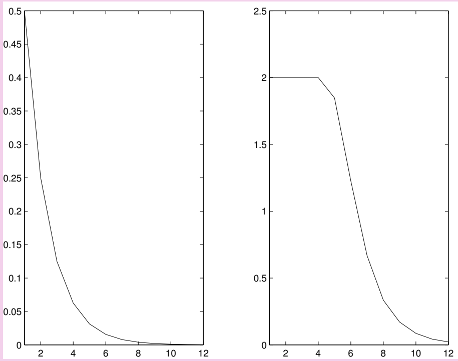
- A^k: Determines stability of finite difference schemes; determines the convergence of stationary iterative methods for linear systems.

  $\lim_{k\to\infty} \|A^k\| = 0$ if and only if $\rho(A) < 1$. But eigenvalues alone cannot distinguish:



- A^k: Markov chains.

  y_0 = initial state; $A^k y_0$ = state after k steps. $A^k y_0 \to v$ = eigenvector corresponding to eigenvalue 1. For k large, convergence rate is determined by second largest eigenvalue. But eigenvalues cannot distinguish:



- $\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(A)\|$: Residual norm in ideal GMRES.

  Any possible convergence behavior of GMRES can be attained with a matrix having any given eigenvalues. (G., Pták, Strakoš, '96)

Given an n by n matrix A, find a set $S \subset \mathbf{C}$ that can be associated with A to give more information than the spectrum alone can provide about the 2-norm of functions of A.

- Field of values:

$$W(A) = \{\langle Aq, q\rangle : \ \langle q, q\rangle = 1\}.$$

- ϵ-pseudospectrum:

$$\sigma_\epsilon(A) = \{z \in \mathbf{C} : \ z \text{ is an eigenvalue of } A + E \text{ for some } E \text{ with } \|E\| < \epsilon\}.$$

- Polynomial numerical hull of degree k:

$$\mathcal{H}_k(A) = \{z \in \mathbf{C} : \ \|p(A)\| \geq |p(z)| \ \forall p \in \mathcal{P}_k\}.$$

Field of Values or Numerical Range

- $W(A)$ is closed if A is finite dimensional (continuous image of compact unit ball); not necessarily so if A is an operator on infinite dimensional Hilbert space.

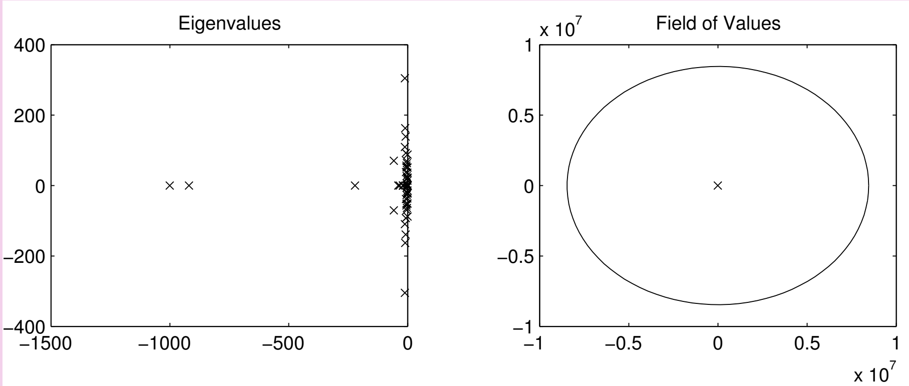
- $\sigma(A) \subset \overline{W(A)}$.

  Proof for eigenvalues: $Aq = \lambda q$, $\|q\| = 1 \quad \Rightarrow \langle Aq, q \rangle = \lambda$.

- $W(A)$ is a **convex** set (Toeplitz-Hausdorf theorem, 1918).

  Method of Proof: Reduce to the 2 by 2 case.

- If A is normal then $\overline{W(A)}$ is the convex hull of $\sigma(A)$; if A is nonnormal $W(A)$ contains more.



- If $\mathbf{y}' = A\mathbf{y}$ then for certain initial data, $\|\mathbf{y}(t)\|$ initially increases if $W(A)$ extends into rhp; $\|\mathbf{y}(t)\|$ decreases monotonically if $W(A)$ lies in lhp.

  Proof:

  $$\frac{d}{dt}\langle \mathbf{y}(t), \mathbf{y}(t)\rangle = 2\text{Re}\langle \mathbf{y}'(t), \mathbf{y}(t)\rangle = 2\text{Re}\langle A\mathbf{y}, \mathbf{y}\rangle.$$

- If $0 \notin W(A)$, then

  $$\min_{\substack{p\in\mathcal{P}_1 \\ p(0)=1}} \|p(A)\| \le \sqrt{1 - d^2/\|A\|^2},$$

  where d is the distance from 0 to $W(A)$.

ϵ-Pseudospectrum

- 1974, Landau, Varah, Godunov; more recently *Trefethen*.

$$\begin{aligned}\Lambda_\epsilon(A) &= \{z \in \mathbf{C}:\ \|(zI-A)^{-1}\| > \epsilon^{-1}\}\\ &= \{z \in \mathbf{C}:\ z \in \sigma(A+E) \text{ for some } E\\ &\qquad \text{with } \|E\| < \epsilon\}.\end{aligned}$$

- Cauchy integral formula:

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

$(zI-A)^{-1}$ is called the **resolvent**; ϵ-pseudospectra are level curves of the resolvent norm. Take $\Gamma = \partial\Lambda_\epsilon$:

$$\|f(A)\| \le \frac{1}{2\pi}\frac{\mathcal{L}(\partial\Lambda_\epsilon)}{\epsilon}\max_{z\in\Lambda_\epsilon}|f(z)|.$$

Crouzeix's Conjecture: For any polynomial p,

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

- "If true it would be astounding." (Peter Lax)
- Need only consider $p = B \circ g$ where g is a conformal mapping from $W(A)$ to $\mathcal{D}$ and B is a finite Blaschke product. Show $\|B(g(A))\| \leq 2$.

- Crouzeix proved
  $\|p(A)\| \leq 11.08 \max_{z \in W(A)} |p(z)|$, but proof is *complicated* and does not appear to be extendable to yield smaller constant.

- Constant 2 can be attained:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

  $W(A)$ is disk of radius 1/2 about 0.
  $\|A\| = 1 = 2 \max_{z \in \mathcal{D}_{1/2}} |z|$.

- Another open question: If constant 2 is attained, is $W(A)$ necessarily a disk? (Yes, for 2 by 2 matrices.)

- For more information and interesting open problems, see:

  `http://perso.univ-rennes1.fr/michel.crouzeix`