Crouzeix's Conjecture

Anne Greenbaum, Applied Math Dept., UW

Current Topics Seminar, Oct., 2009

Where in the complex plane does a matrix live?
(A question of L. N. Trefethen)

Translating Matrix Problems into Problems in the Complex Plane

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:

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

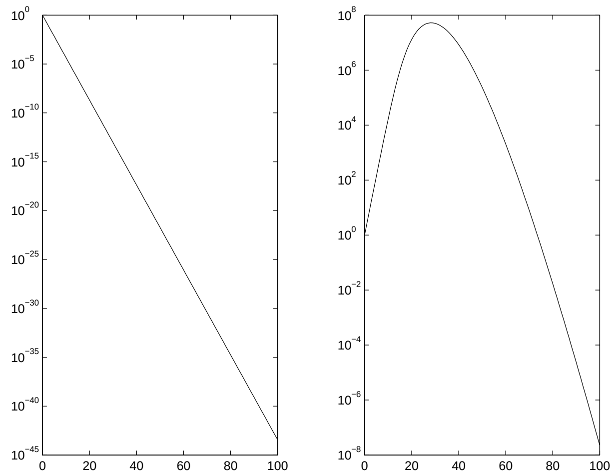
- Even if A is highly **nonnormal** (e.g., not diagonalizable, or diagonalizable but with eigenvectors that are almost linearly dependent), eigenvalues determine the *asymptotic* behavior of many functions of A:

$$\|A^k\| \to 0 \text{ as } k \to \infty \text{ iff } \rho(A) < 1.$$
$$\|e^{tA}\| \to 0 \text{ as } t \to \infty \text{ iff } \mathrm{Re}(\sigma(A)) < 0.$$

What can eigenvalues NOT do?

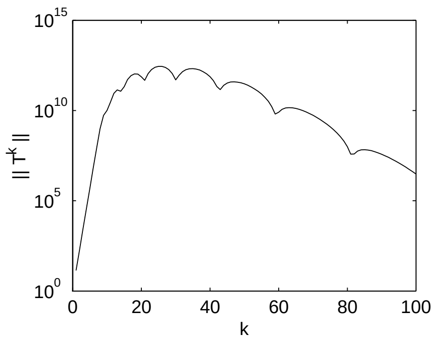
- e^{tA}: Determines the stability of $y' = Ay$.

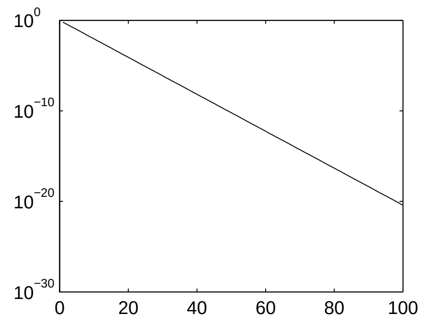
  $\lim_{t\to\infty} \|e^{tA}\| = 0$ if and only if the eigenvalues of A have negative real parts.
  But eigenvalues alone cannot distinguish:



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

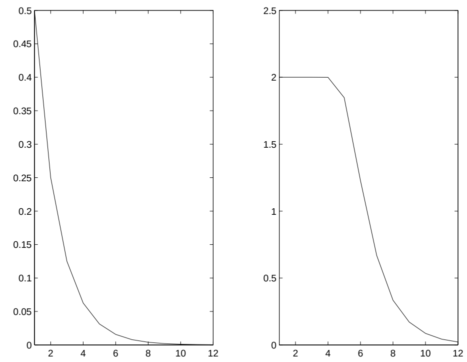
  $\lim_{k\to\infty}\|A^k\| = 0$ if and only if $\rho(A) < 1$. But eigenvalues alone cannot distinguish:





- A^k: Markov chains.

  y_0 = initial state; $A^k y_0$ = state after k steps. $A^k y_0 \rightarrow v$ = 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 or Numerical Range:

$$W(A) = \{\langle Aq, q\rangle : \ q \in \mathbf{C^n}, \ \langle q, q\rangle = 1\}.$$

- 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:

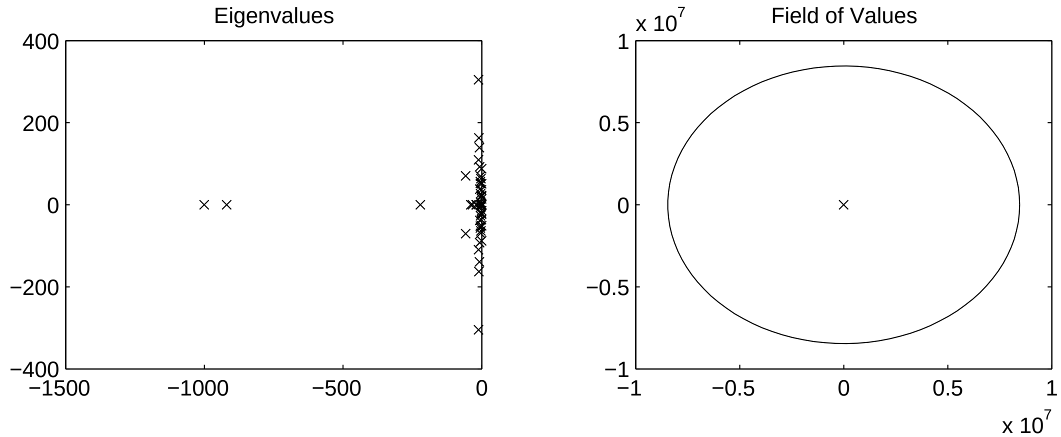
$$\mathcal{H}_k(A) = \{z \in \mathbf{C} : \ \|p(A)\| \geq |p(z)| \ \forall p \in \mathcal{P}_k\}.$$

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

- If $0 \notin W(A)$, then
$$\min_{\substack{p\in\mathcal{P}_1\\ p(0)=1}} \|p(A)\| \leq \sqrt{1 - d^2/\|A\|^2},$$
  where d is the distance from 0 to $W(A)$.

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)

- 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.)

  May suggest a direction for proof.

- For more information and interesting open problems, see:

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

Known Results

- Von Neumann's Inequality (1951):

$$\|p(A)\| \le \max_{z \in \mathcal{D}_{\|A\|}} |p(z)|.$$

- Power Inequality (Berger/Pearcy, 1966):

$$\|A^k\| \le 2 \max_{z \in W(A)} |z^k|.$$

More precisely, $\nu(A^k) \le \nu(A)^k$, where $\nu(A)$ is the numerical radius: $\max_{z \in W(A)} |z|$.

- Badea (2004), based on Ando (1973):

$$\|p(A)\| \leq 2 \max_{z \in \mathcal{D}_{\nu(A)}} |p(z)|.$$

- Crouzeix (2004 $->$):

  The conjecture is true for 2 by 2 matrices. For general n by n matrices,

$$\|p(A)\| \leq 11.08 \max_{z \in W(A)} |p(z)|$$

  If A is a 2 by 2 matrix and $W(A)$ is a disk, then best constant is 2; if $W(A)$ is an ellipse with eccentricity ϵ, then the best constant is

$$2 \exp\left(- \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} \frac{2}{1 + \rho^{4n}} \right), \quad \text{where } \rho = \frac{1 + \sqrt{1 - \epsilon^2}}{\epsilon}$$

  Method of Proof: Explicitly map $W(A)$ to $\bar{\mathcal{D}}$.

Von Neumann's Inequality

If $\|A\| \leq 1$, it has a *unitary dilation*; e.g.,

$$B = \begin{pmatrix} A & (I - AA^*)^{1/2} & 0 & \dots & 0 \\ 0 & 0 & I & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & I \\ -(I - A^*A)^{1/2} & A^* & 0 & \dots & 0 \end{pmatrix}, \quad p(B) = \begin{pmatrix} p(A) & * \\ * & * \end{pmatrix}.$$

$\|p(A)\| \leq \|p(B)\| \leq \sup_{z \in \mathcal{D}} |p(z)|$.

For general A, apply to $A/\|A\|$. If $q(z) = p(z/\|A\|)$, then

$$\|q(A)\| = \|p(A/\|A\|)\| \leq \sup_{z \in \mathcal{D}} |p(z)| = \sup_{z \in \mathcal{D}_{\|A\|}} |q(z)|.$$

Badea's Result

Ando: If $\nu(A) \leq 1$, then there is a Hermitian matrix B and a unitary matrix U such that:

$$A = 2\cos(B)U\sin(B).$$

Claim: A is similar to a contraction via a similarity transformation with condition number ≤ 2.

Let $g(x) = \max\{1, 2|\cos x|\}$, and define $H = g(B)$, $T = H^{-1}AH$. Then

$$\|H\| \leq 2, \quad \|H^{-1}\| \leq 1, \quad \|\sin(B)H\| \leq 1, \quad 2\|H^{-1}\cos(B)\| \leq 1.$$

Thus $\|T\| \leq 1$. $\square$

By von Neumann's inequality,

$$\|p(A)\| \leq \|H\|\ \|p(T)\|\ \|H^{-1}\| \leq 2\|p\|_{\mathcal{L}^\infty(\mathcal{D})}.$$

What p maximizes $\|p(A)\|/\|p\|_{\mathcal{L}^\infty(W(A))}$? Don't know, but ...

$p(A)$ is completely determined by the values of p (and perhaps some of its derivatives) at the eigenvalues of A. Hence conjecture is equivalent to:

$$\|p(A)\| \leq 2 \inf\{\|f\|_{\mathcal{L}^\infty(W(A))} : f(A) = p(A)\}$$

Finding this infimum is a **Pick-Nevanlinna interpolation problem**.

Map $W(A)$ conformally to $\bar{\mathcal{D}}$. Infimum is achieved by a function $\tilde{f}$ that is a scalar multiple of a finite **Blaschke product**:

$$\tilde{f}(z) = \mu \prod_{k=0}^{n-1} \frac{z - \alpha_k}{1 - \bar{\alpha}_k z} = \mu \frac{\gamma_0 + \gamma_1 z + \ldots + \gamma_{n-1} z^{n-1}}{\bar{\gamma}_{n-1} + \bar{\gamma}_{n-2} z + \ldots + \bar{\gamma}_0 z^{n-1}}$$

Using second representation, Glader and Lindström showed how to compute $\tilde{f}$ and $\|\tilde{f}\|_{\mathcal{L}^\infty(\mathcal{D})}$ by solving a simple eigenvalue problem. Determine μ, $\gamma_0, \ldots, \gamma_{n-1}$ from conditions $\tilde{f}(\hat{\lambda}_j) = p(\lambda_j)$, $j = 1, \ldots, n$, where $\hat{\lambda}_j$'s are the mapped eigenvalues.

Let V be the Vandermonde matrix for $\hat{\lambda}_1, \ldots, \hat{\lambda}_n$:

$$V^T = \begin{pmatrix} 1 & \hat{\lambda}_1 & \ldots & \hat{\lambda}_1^{n-1} \\ 1 & \hat{\lambda}_2 & \ldots & \hat{\lambda}_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \hat{\lambda}_n & \ldots & \hat{\lambda}_n^{n-1} \end{pmatrix}.$$

If $\gamma = (\gamma_0, \ldots, \gamma_{n-1})^T$, and Π is the permutation matrix with 1's on its skew diagonal, then these conditions are:

$$V^{-T} p(\Lambda) V^T \Pi \bar{\gamma} = \mu \gamma.$$

Largest real μ for which this holds for some nonzero vector γ is $\|\tilde{f}\|_{\mathcal{L}^\infty(\mathcal{D})}$. This is a **coneigenvalue** problem; equate real and imaginary parts to get a $2n$ by $2n$ eigenvalue problem.

Numerical Testing of Crouzeix's Conjecture

$$\|p(A)\| \overset{?}{\leq} 2 \ \inf\{\|f\|_{\mathcal{L}^\infty(W(A))} : f(A) = p(A)\}$$

- Given A, compute eigendecomposition $A = S\Lambda S^{-1}$, field of values $W(A)$, conformal mapping $g : W(A) \to \bar{\mathcal{D}}$, and $g(\Lambda)$.

- Try values $w_1, \ldots, w_n$ for $p(\lambda_1), \ldots, p(\lambda_n)$. Compute $\|p(A)\| = \|Sp(\Lambda)S^{-1}\|$, and find

$$\mu \equiv \inf\{\|f\|_{\mathcal{L}^\infty(\mathcal{D})} : f(g(\lambda_j)) = w_j, \ j = 1, \ldots, n\}$$

  by solving eigenvalue problem.

- Vary $w_1, \ldots, w_n$ to minimize $\mu / \|p(A)\|$. If $< \frac{1}{2}$, conjecture is false.

  Experiments show that for some problems (e.g. 3×3 perturbed Jordan block with small ξ) need (almost) exact $W(A)$ and $g : W(A) \to \mathcal{D}$ to obtain constant 2.