Bounds on Norms of Functions of Matrices Using the Field of Values

Anne Greenbaum, Applied Math Dept., UW

PNWNAS, Oct., 2010
Definitions

• $A$ an $n$ by $n$ matrix (bounded linear operator on a Hilbert space),

• $p$ a given polynomial (analytic function),

• $\|p(A)\| \equiv \max_{\|v\|_2=1} \|p(A)v\|_2$, which is the largest singular value of $p(A)$. 
If $A$ is diagonalizable, $A = V \Lambda V^{-1}$, then $p(A) = V p(\Lambda) V^{-1}$, and

$$\max_{\lambda \in \sigma(A)} |p(\lambda)| \leq \|p(A)\| \leq \|V\| \cdot \|V^{-1}\| \cdot \max_{\lambda \in \sigma(A)} |p(\lambda)|.$$ 

If $\kappa(V) \equiv \|V\| \cdot \|V^{-1}\|$ is of moderate size, then $\sigma(A)$ determines $\|p(A)\|$.

What if $\kappa(V)$ is huge?
Even if $\kappa(V)$ is huge, eigenvalues still determine \textit{asymptotic} behavior.

- $\|e^{tA}\|$ determines stability of $y'(t) = Ay(t), \ t > 0$.
  $$\lim_{t \to \infty} \|e^{tA}\| = 0 \iff \Re(\sigma(A)) < 0.$$  

- $\|(I + (\Delta t)A)^k\|$ determines stability of finite difference method (Euler’s method) for approximating the solution.
  $$\lim_{k \to \infty} \|B^k\| = 0 \iff \rho(B) < 1.$$  

What about transient behavior?
Field of Values or Numerical Range

\[ W(A) = \{ q^* Aq : q \in \mathbb{C}^n, \|q\| = 1 \} . \]

- \( W(\alpha I + A) = \alpha + W(A) \) since \( q^*(\alpha I + A)q = \alpha + q^*Aq \).
- \( W(cA) = cW(A) \) since \( q^*(cA)q = cq^*Aq \).
- For \( A \) finite dimensional, \( W(A) \) is closed (continuous image of compact unit ball in \( \mathbb{C}^n \)), not necessarily so in infinite dimensions.
- \( W(A) \) (\( W(A) \)) contains \( \sigma(A) \):
  \[\begin{align*}
  Av &= \lambda v, \|v\| = 1 \implies v^*Av = \lambda \in W(A).
\end{align*}\]
• Toeplitz-Hausdorff theorem (1918-1919). $W(A)$ is convex.

Standard method of proof: Reduce to $2 \times 2$ case.

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

![Eigenvalues](image1.png) ![Field of Values](image2.png)
Crouzeix’s Conjecture

For any square matrix $A$ and any polynomial $p$ (or any function analytic in $W(A)$),

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

- **Constant 2 can be attained:**
  
  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

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

- **Crouzeix proved:**
  
  $$\|p(A)\| \leq 11.08 \max_{z \in W(A)} |p(z)|.$$  

- For more information and interesting open problems, see:

  http://perso.univ-rennes1.fr/michel.crouzeix
Related Results

• Von Neumann’s Inequality (1951):

\[ \|p(A)\| \leq \max_{z \in D(\|A\|,0)} |p(z)|. \]

Proof: \( A/\|A\| \) has norm 1. Any matrix \( C \) with \( \|C\| \leq 1 \) has a unitary dilation:

\[
Z = \begin{pmatrix}
C & (I - CC^*)^{1/2} \\
-(I - C^*C)^{1/2} & C^*
\end{pmatrix}.
\]

In fact, there is a unitary dilation \( U \) such that \( p(U) \) is a dilation of \( p(C) \):

\[
U = \begin{pmatrix}
C & (I - CC^*)^{1/2} \\
0 & I \\
-(I - C^*C)^{1/2} & C^* \\
\end{pmatrix}, \quad p(U) = \begin{pmatrix}
p(C) & * \\
* & * \\
\end{pmatrix}.
\]
\[ \|p(C)\| \leq \|p(U)\| = \max_{\lambda \in \sigma(U)} |p(\lambda)| \leq \max_{z \in D(1,0)} |p(z)|. \]

If \( q(z) = \|A\|z \), then \( p(A) = (p \circ q)(A/\|A\|) \), so

\[ \|p(A)\| \leq \max_{z \in D(1,0)} |(p \circ q)(z)| = \max_{\zeta \in D(\|A\|,0)} |p(\zeta)|. \]

\[\square\]
Numerical radius: \( r(A) \equiv \max_{z \in W(A)} |z|. \)

- **Power Inequality (Berger/Pearcy, 1966):**
  \[
  r(A^k) \leq [r(A)]^k, \quad k = 1, 2, \ldots.
  \]

- **Claim:** \( \frac{1}{2} \|A\| \leq r(A) \leq \|A\|. \)
  
  **Proof:**
  \[
  |q^*Aq| \leq \|q^*\| \cdot \|A\| \cdot \|q\| = \|A\| \text{ if } \|q\| = 1. \text{ Therefore } r(A) \leq \|A\|. \]
  
  To see that \( \|A\| \leq 2r(A) \), write \( A = (A + A^*)/2 + (A - A^*)/2: \)
  \[
  \|A\| \leq \|(A + A^*)/2\| + \|(A - A^*)/2\|
  = \max_{\|q\|=1} |\text{Re}(q^*Aq)| + \max_{\|w\|=1} |\text{Im}(w^*Aw)|
  \leq \max_{\|q\|=1} |q^*Aq| + \max_{\|w\|=1} |w^*Aw| = 2r(A). \quad \square
  \]
Therefore follows from power inequality that

\[ \|A^k\| \leq 2 \max_{z \in W(A)} |z^k|; \]

i.e., Crouzeix’s conjecture holds for \( p(A) = A^k \).
• Okubo and Ando (1975):

\[ \|p(A)\| \leq 2 \max_{z \in D(r(A),0)} |p(z)|. \]

By scaling and translating,

\[ \|p(A)\| \leq 2 \max \{|p(z)| : z \in \text{smallest disk containing } W(A) \}. \]
Lemma. If $r(A) \leq 1$, then there is a Hermitian matrix $H$ and a unitary matrix $U$ such that $A = 2 \cos(H)U \sin(H)$.

Theorem. If $r(A) \leq 1$ then $A$ is similar to a contraction via a similarity transformation with condition number less than or equal to 2; i.e., $A = SCS^{-1}$, where $\|C\| \leq 1$ and $\kappa(S) \leq 2$.

Proof of Theorem. Let $g(x) = \max\{1, 2|\cos x|\}$, and define $S = g(H)$, $C = S^{-1}AS$. Then

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

Therefore

$$\|C\| = \|2S^{-1}\cos(H)U \sin(H)S\| \leq 2\|S^{-1}\cos(H)\| \cdot \|U\| \cdot \|\sin(H)S\| \leq 1. \quad \Box$$

Crouzeix’s conjecture holds if $W(A)$ is a disk.
Stronger (equivalent?) conjecture:
There is a conformal mapping $g$ from $W(A)$ to $D$ such that $g(A) = SCS^{-1}$ where $\|C\| \leq 1$ and $\kappa(S) \leq 2$.

$\implies$ Crouzeix’s conjecture:

$$p(A) = (p \circ g^{-1})(g(A)) = S(p \circ g^{-1})(C)S^{-1}.$$  

$$\|p(A)\| \leq \kappa(S) \max_{\zeta \in D} |(p \circ g^{-1})(\zeta)|$$ (von Neumann)

$$= \kappa(S) \max_{z \in W(A)} |p(z)|.$$  

In other words, any matrix $A$ can be written as $A = Sf(C)S^{-1}$, where $f$ is a conformal mapping from $D$ to $W(A)$, $\|C\| \leq 1$, and $\kappa(S) \leq 2$. 
Perturbed Jordan block:

\[ J_\nu = \begin{pmatrix}
0 & 1 \\
\vdots & \ddots & \ddots \\
\nu & \ddots & \ddots & 1 \\
\nu & \ddots & \ddots & 0
\end{pmatrix}, \quad \nu \in (0, 1). \]

- \( \nu = 0 \): Jordan block, \( W(J_0) \) is a disk. Therefore Crouzeix’s conjecture holds for \( J_0 \).

- \( \nu = 1 \): normal matrix. \( \| p(A) \| = \max_{\lambda \in \sigma(A)} |p(\lambda)| \). Therefore conjecture holds for \( J_1 \). In fact, if \( \nu \geq 2^{-n/(n-1)} \), then \( \kappa(V) \leq 2 \).

Eigenvalues of \( J_\nu \) are \( n \)th roots of \( \nu \). Field of values is complicated.
But, by symmetry, conformal mapping \( g \) from \( W(J_\nu) \) to \( D \) with \( g(0) = 0 \) just multiplies eigenvalues by a constant \( c \).

\[
g(J_\nu) = cJ_\nu.
\]

- \( g(J_\nu) = D^{-1}J_{cn_\nu}D \), where \( D = \text{diag}(1, c, \ldots, c^{n-1}) \).
- **If** \( c \leq 2^{1/(n-1)} \), then \( \kappa(D) \leq 2 \) and \( \|J_{cn_\nu}\| = 1 \).
- **If** \( c > 2^{1/(n-1)} \), then \( \|g(J_\nu)^{n-1}\| = c^{n-1} > 2 \), but \( \max_{z \in W(J_\nu)} |g(z)^{n-1}| = 1 \).

So, conjecture is true for perturbed Jordan blocks *if and only if* \( c \leq 2^{1/(n-1)} \).
• $c \leq \frac{1}{\cos\left(\frac{\pi}{n}\right)}$ (because field of values of $n - 1$ by $n - 1$ principal submatrix is a disk of this radius), and this is $< \frac{2^{1/(n-1)}}{n}$ for $n > 6$.

• Can also show $c \leq \frac{2^{1/(n-1)}}{n}$ for $n = 6$.

• Still do not know for $n = 3, 4, 5$. Difficult for small $n$ because no room for approximations.
Related Areas to Explore

• Crouzeix’s conjecture has implications in a number of areas; e.g., convergence rate of GMRES when \( 0 \not\in W(A) \), and even, by considering the \textit{minimal norm interpolating function}, when \( 0 \in W(A) \).

• Probably better to study the conjecture: Any matrix \( A \) can be written in the form \( A = Sf(C)S^{-1} \), where \( f \) is a conformal mapping from \( D \) to \( W(A) \), \( \|C\| \leq 1 \), and \( \kappa(S) \leq 2 \).

• What class of matrices are similar via a similarity transformation with condition number at most 2?