Mathematics & Statistics
Auburn University, Alabama, USA

Jan 24, 2007

On Beruling-Gelfand's spectral radius theorem

Speaker: Tin-Yau Tam

2007 Spring Graduate Seminar

tamtiny@auburn.edu

Outline

1. Introduction

Beruling-Gelfand's spectral radius theorem

$$\lim_{m\to\infty} \|X^m\|^{1/m} = r(X),$$

where

X is a square complex matrix,

$r(X)$ is the spectral radius of X, and

$\|X\|$ is the spectral norm of X.

We will prove the result in an elementary way and discuss its extensions.

2. Basics

$\mathbb{C}^n$ = the space of complex n-tuples.

$\mathbb{C}_{n\times n}$ = the space of $n \times n$ complex matrices.

A number $\lambda \in \mathbb{C}$ is said to be an eigenvalue of $X \in \mathbb{C}_{n\times n}$ if there is a nonzero vector $v \in \mathbb{C}^n$ such that

$$Xv = \lambda v.$$

Spectral radius: $r(X)$ is the maximum eigenvalue modulus.

2-norm: $\|v\|_2 = (v^* v)^{1/2}$.

Spectral norm:

$$\|X\| := \max_{\|v\|_2=1} \|Xv\|_2$$

Fact: spectral norm is **submultiplicative**

1. $\|Xv\|_2 \leq \|X\|\,\|v\|_2$, for all $X \in \mathbb{C}_{n\times n}$ and $v \in \mathbb{C}^n$.

2. $\|AB\| \leq \|A\|\,\|B\|$, for all $A, B \in \mathbb{C}_{n\times n}$.

Proof. (1)

$$\frac{\|Xv\|_2}{\|v\|_2} = \|Xw\|_2 \leq \|X\|$$

where $w := \frac{v}{\|v\|_2}$ is a unit vector.

(2)

$$\|AB\| = \max_{\|v\|_2=1} \|ABv\|_2 \leq \max_{\|v\|_2=1} \|A\|\,\|Bv\|_2 = \|A\|\,\|B\|.$$

□

Question: But how to compute the spectral norm? The definition

$$\|X\| = \max_{\|v\|_2=1} \|Xv\|_2$$

is no good.

A matrix $H \in \mathbb{C}_{n\times n}$ is said to be Hermitian if $H^* = H$.

Fact: Eigenvalues of a Hermitian H are real.

Proof. Suppose $Hv = \lambda v$ for some unit vector $v \in \mathbb{C}^n$. Then $v^* H v = \lambda$ and by taking $*$ on both sides

$$\lambda = v^* H v = v^* H^* v = \bar{\lambda}.$$

□

Theorem 2.1. *(Rayleigh-Ritz) Let* $H \in \mathbb{C}_{n\times n}$ *be a Hermitian matrix. Then* $\max_{\|v\|_2=1} v^*Hv$ *is the largest eigenvalue of* H.

Sketch of proof: Spectral theorem for Hermitian matrix H says that there is an orthonormal basis $\{v_1, \dots, v_n\}$ such that $Hv_i = \lambda_i v_i$, $i = 1, \dots, n$.

By Rayleigh-Ritz's theorem,

$$\|X\| = \max_{\|v\|_2=1} \|Xv\|_2 = \max_{\|v\|_2=1} (v^*X^*Xv)^{1/2} = (\max_{\|v\|_2=1} v^*X^*Xv)^{1/2}$$

is the square root of the largest eigenvalue of X^*X.

Singular values of X are the square roots of eigenvalues of the positive semi-definite matrix X^*X.

So $\|X\|$ is the largest singular value of X.

Arrange the singular values of X in descending order

$$s_1(X) \geq s_2(X) \geq \cdots \geq s_n(X)$$

and the eigenvalue moduli in descending order

$$|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|.$$

So $r(X) = |\lambda_1|$.

Fact: $r(X) \le \|X^m\|^{1/m} \le \|X\|$ for all $m \in \mathbb{N}$.

Proof. Since $\|X\| := \max_{\|v\|_2=1} \|Xv\|_2$ and if $Xv = \lambda v$ for some unit vector v, then

$$|\lambda| = \|Xv\|_2 \le \|X\|.$$

It is not hard to see that $r(X^m) = r(X)^m$. So for all $m \in \mathbb{N}$,

$$r(X)^m = r(X^m) \le \|X^m\|.$$

By submultiplicativeness of $\| \cdot \|$,

$$\|X^m\| \le \|X\|^m.$$

□

3. Proof

Theorem 3.1. (Beruling, 1938, Gelfand, 1941) Let $A \in \mathbb{C}_{n\times n}$.

$$\lim_{m\to\infty} \|X^m\|^{1/m} = r(X). \tag{1}$$

where $r(X)$ is the spectral radius of X and $\|X\|$ is the spectral norm of A.

We now provide an elementary proof which is different from those in the literature.

Some ingredients:

(1) Schur triangularization theorem asserts that if $X \in \mathbb{C}_{n\times n}$, then there is a unitary matrix U ($U^* = U^{-1}$) such that $T := U^*XU$ is upper triangular and the diagonal entries of T are the eigenvalues of X and can be preordered. Proof: Induction on n.

(2) $r(T) = r(X)$ Proof: Similar matrices have the same eigenvalues.

(3) $T^m = U^*X^mX$ so that $\|T^m\| = \|U^*X^mU\| = \|X^m\|$ (exercise! Hint: $\|Uv\|_2 = \|v\|$ for all $v \in \mathbb{C}^n$).

(4) $\|X\| \le \sum_{i=1}^n \sum_{j=1}^n |x_{ij}|$ (exercise!).

We may assume that $X = T$ is upper triangular with ascending diagonal moduli

$$|t_{11}| \le \cdots \le |t_{nn}|.$$

When T is nilpotent, that is, $r(T) = 0$, we have a strictly upper triangular matrix T. Thus $T^m = 0$ for $m \ge n$ and hence (1) is obviously true. So we assume that T is not nilpotent so that $r(X) = |t_{nn}| \ne 0$. Since both sides of (1) are homogenous, by appropriate scaling, we may assume that $|t_{nn}| \ge 1$.

Write $T^m = [t_{ij}^{(m)}] \in \mathbb{C}_{n\times n}$ which is also upper triangular. For $1 \le i \le j \le n$,

$$t_{ij}^{(m)} = \sum_{i=p_0\le p_1\le\cdots\le p_m=j} \prod_{\ell=1}^{m} t_{p_{\ell-1}p_\ell}. \tag{2}$$

Clearly $t_{ii}^{(m)} = t_{ii}^m$, $i = 1, \ldots, n$. Let us estimate $|t_{ij}^{(m)}|$ for fixed $i < j$.

The number of $(m+1)$-tuples $(p_0, p_1, \cdots, p_m)$, where $i = p_0 \le p_1 \le \cdots \le p_m = j$ are integers, is equal to $\binom{j-i+m-1}{m-1}$. For each of such $(p_0, \cdots, p_m)$, there are at most $j-i$ numbers ℓ's in $\{1, \cdots, m\}$ such that $p_{\ell-1} \ne p_\ell$.

Let

$$c := \max_{1 \le p,q \le n} |t_{pq}| \ge 1$$

denote the maximal entry modulus of T. By (2) and the fact that $|t_{11}| \le \cdots \le |t_{nn}|$, when $m \ge n$,

$$\begin{aligned}
|t_{ij}^{(m)}| &\le \sum_{i=p_0 \le p_1 \le \cdots \le p_m = j} \prod_{\ell=1}^{m} |t_{p_{\ell-1}p_\ell}| \\
&\le \sum_{i=p_0 \le p_1 \le \cdots \le p_m = j} c^{j-i}\,|t_{jj}|^{m-j+i} \\
&= \binom{j-i+m-1}{m-1} c^{j-i} |t_{jj}|^{m-j+i} \qquad (3) \\
&\le \binom{n+m-2}{m-1} c^{n-1} |t_{nn}|^{m}.
\end{aligned}$$

By (3), for $m \geq n$,

$$\begin{aligned} |t_{nn}|^m = r(T)^m &\leq \|T^m\| \\ &\leq \sum_{j=1}^{n} \sum_{i=1}^{n} |t_{ij}^{(m)}| \\ &\leq n^2 c^{n-1} \binom{n+m-2}{m-1} |t_{nn}|^m \\ &\leq n^2 c^{n-1} (n+m-2)^{n-1} |t_{nn}|^m \end{aligned}$$

Taking m-th roots on all sides and taking limits for $m \to \infty$ lead to

$$\lim_{m\to\infty} \|T^m\|^{1/m} = |t_{nn}| = r(T).$$

4. Extensions

Theorem 4.1. Let $A, B, X \in \mathbb{C}_{n\times n}$ such that A, B are nonsingular. Then

$$\lim_{m\to\infty} \|AX^m B\|^{1/m} = r(X) \qquad (4)$$

Remark: All norms on $\mathbb{C}_{n\times n}$ are equivalent, i.e., if $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are norms $\mathbb{C}^n$, there are constants $c_1, c_2 > 0$ such that

$$c_1\|X\|_\alpha \le \|X\|_\beta \le c_2\|X\|_\alpha,$$

for all $X \in \mathbb{C}_{n\times n}$. So the theorem is true for all norm, not just the spectral norm.

Proof. It suffices to prove (4) for $B = I_n$ and for all $X \in \mathbb{C}_{n\times n}$ since

$$AX^mB = AB(B^{-1}XB)^m$$

and the spectrum of X and $B^{-1}XB$ are identical. Since $\|\cdot\|$ is submultiplicative, for all $m \in \mathbb{N}$,

$$\frac{1}{\|A^{-1}\|^{1/m}}\|X^m\|^{1/m} \le \|AX^m\|^{1/m} \le \|A\|^{1/m}\|X^m\|^{1/m}.$$

Since $\|A^{-1}\|^{1/m}$ and $\|A\|^{1/m}$ converge to 1, we have the desired result. □

Rewrite Beurling-Gelfand :

$$\lim_{m\to\infty} \|X^m\|_2^{1/m} = r(X) \Leftrightarrow \lim_{m\to\infty} [s_1(X^m)]^{1/m} = |\lambda_1(X)|$$

Yamamoto (1967):

$$\lim_{m\to\infty} [s_i(X^m)]^{1/m} = |\lambda_i(X)|, \quad i = 1, \dots, n.$$

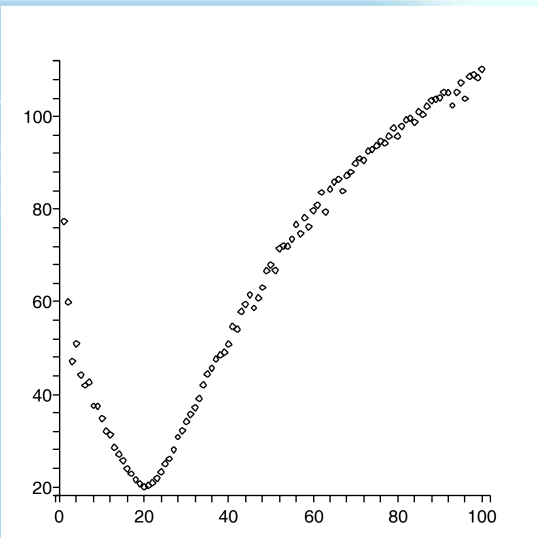
- a natural generalization of Beurling-Gelfand (finite dim. case).
- Loesener (1976) rediscovered Yamamoto
- Mathias (1990 another proof)
- Johnson and Nylen (1990 generalized singular values)
- Nylen and Rodman (1990 Banach algebra)

Numerical experiments:

Computing the discrepancy between

$$[s(X^m)]^{1/m} \quad \text{and} \quad |\lambda(X)|$$

of randomly generated $X \in \mathrm{GL}_n(\mathbb{C})$. Here $s(X) := \mathrm{diag}\,(s_1(X), \ldots, s_n(X))$ and $\lambda(X) := \mathrm{diag}\,(\lambda_1(X), \ldots, \lambda_n(X))$.



The graph of

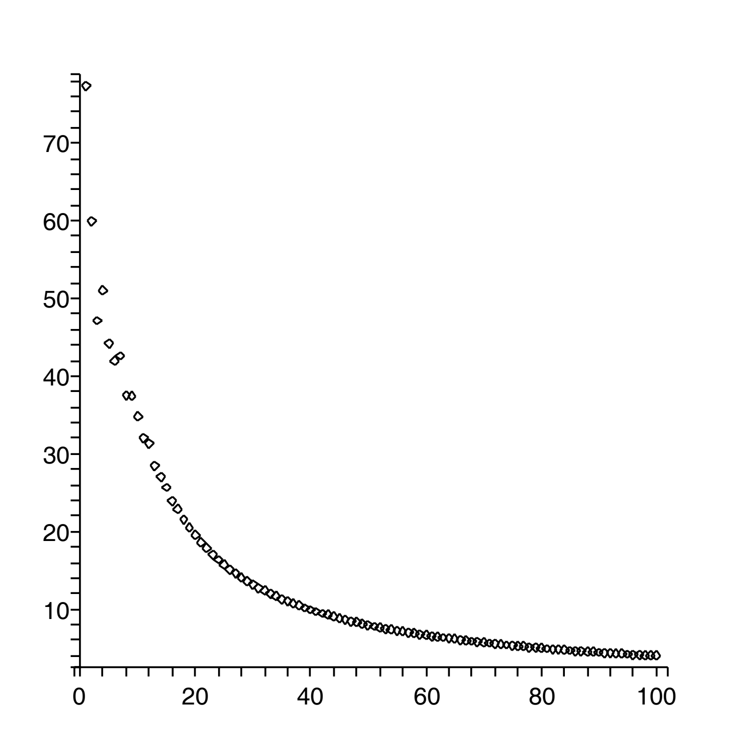
$$\|[s(X^m)]^{1/m} - |\lambda(X)|\|_2$$

versus m ($m = 1, \ldots, 100$)

If we consider

$$|s_1(X^m)^{1/m} - |\lambda_1(X)||$$

for the above example, convergence occurs. But divergence occurs for $|s_2(X^m)^{1/m} - |\lambda_2(X)||$



Theorem 4.2. Let $A, B, X \in \mathbb{C}_{n\times n}$ such that A, B are nonsingular. Then

$$\lim_{m\to\infty} [s_k(AX^m B)]^{1/m} = |\lambda_k(X)|, \quad k = 1, \dots, n.$$

Proof. One may use the inequality (needs a proof)

$$s_i(AB) \le s_1(A) s_i(B),$$

$A, B \in \mathbb{C}_{n\times n}$ and continuity argument to have

$$s_n(A)s_i(X^m) \le s_i(AX^m) \le s_1(A)s_i(X^m), \quad i = 1, \dots, n.$$

By Yamamoto we get the result. □

Banach algebra

A Banach algebra $\mathcal{A}$ is a complex algebra and a Banach space with respect to a norm that satisfies the submultiplicative property:

$$\|xy\| \leq \|x\| \, \|y\|, \qquad x, y \in \mathcal{A}$$

and $\mathcal{A}$ contains a unit element e such that

$$\|e\| = 1$$

and

$$xe = ex = x$$

for all $x \in \mathcal{A}$.

Example: $\mathbb{C}_{n\times n}$ is a Banach algebra with I_n as the identity element with respect to the spectral norm.

Example: H = a Banach space. Then $B(H)$ = the algebra of all bounded linear operators on H is a Banach algebra with respect to the usual operator norm. The identity operator I is the unit element. Every closed subalgebra of $B(H)$ that contains I is also a Banach algebra.

Spectrum $\sigma(x)$ of $x \in \mathcal{A}$ is the set of all $\lambda \in \mathbb{C}$ such that $\lambda e - x$ is not invertible.

Spectral radius of x:

$$r(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}$$

Remark: The spectrum and the spectral radius of an $x \in \mathcal{A}$ are defined in terms of the algebraic structure of $\mathcal{A}$, regardless of any metric (or topological) consideration.

Theorem 4.3. *Let $\mathcal{A}$ be a Banach algebra and $x \in \mathcal{A}$.*

1. *The spectrum $\sigma(x)$ is compact and nonempty.*
2. $$\lim_{m\to\infty} \|x^m\|^{1/m} = \inf_{m\geq 1} \|x^m\|^{1/m} = r(x)$$

Remark: $r(x) \leq \|x\|$ is contained in the formula.

Remark: $\lim_{m\to\infty} \|x^m\|^{1/m}$ depends obviously on metric properties of A. This is a very remarkable feature of the spectral radius formula.

Summary:

$$\lim_{m\to\infty} \|x^m\|^{1/m} = \inf_{m\geq 1} \|x^m\|^{1/m} = r(x)$$

$\mathbb{C}_{n\times n} \longrightarrow$ Banach algebra

matrix $\longrightarrow$ Banach algebra element

spectral norm $\longrightarrow$ norm on a Banach algebra

Remark: However Yamamoto's type result doesn't make sense in Banach algebra. Even we have * operation (then a C^* algebra) so that we have $\sigma(x^*x)$, the spectrum is not finite nor discrete in general.

Review Yamamoto's result:

$$\lim_{m\to\infty}[s(X^m)]^{1/m} = |\lambda(X)|,$$

$s(X) := \operatorname{diag}(s_1(X),\ldots,s_n(X))$ and $\lambda(X) := \operatorname{diag}(\lambda_1(X),\ldots,\lambda_n(X))$.

Singular value decomposition (SVD):

Given $X \in \mathbb{C}_{n\times n}$ there are unitary matrices U, V such that

$$X = USV$$

where $S = \operatorname{diag}(s_1,\ldots,s_n)$.

$\mathrm{GL}_n(\mathbb{C})$, called the general linear group, is the group of nonsingular matrices in $\mathbb{C}_{n\times n}$. It is a Lie group.

Lie groups

- $\mathfrak{g}$ = real semisimple Lie algebra with connected noncompact Lie group G.
- $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ a fixed (algebra) Cartan decomposition of $\mathfrak{g}$
- $K \subset G$ the connected subgroup with Lie algebra $\mathfrak{k}$.
- $\mathfrak{a} \subset \mathfrak{p}$ a maximal abelian subspace.
- Fix a *closed* Weyl chamber $\mathfrak{a}_+$ in $\mathfrak{a}$ and set

$$A_+ := \exp \mathfrak{a}_+, \quad A := \exp \mathfrak{a}$$

An example

Each $A \in \mathbb{C}_{n\times n}$ has Hermitian decomposition:

$$A = \frac{A - A^*}{2} + \frac{A + A^*}{2}.$$

(Group) Cartan decomposition

$$G = KA_+K$$

• $k_1, k_2 \in K$ not unique in $g = k_1 a_+(g) k_2$, the element $a_+(g) \in A_+$ is unique.

CMJD for real semisimple G:

• $h \in G$ is hyperbolic if $h = \exp(X)$ where $X \in \mathfrak{g}$ is real semisimple, that is, $\mathrm{ad}\, X \in \mathrm{End}\,(\mathfrak{g})$ is diagonalizable over $\mathbb{R}$.

• $u \in G$ is unipotent if $u = \exp(N)$ where $N \in \mathfrak{g}$ is nilpotent, that is, $\mathrm{ad}\, N \in \mathrm{End}\,(\mathfrak{g})$ is nilpotent.

• $e \in G$ is elliptic if $\mathrm{Ad}(e) \in \mathrm{Aut}\,(\mathfrak{g})$ is diagonalizable over $\mathbb{C}$ with eigenvalues of modulus 1.

Each $g \in G$ can be uniquely written as

$$g = ehu,$$

where e, h, u commute.

Extension of Yamamoto:

Write $g = e(g)h(g)u(g)$

Fact: $h(g)$ is conjugate to $b(g) \in A_+$.

Theorem 4.4. (Huang and Tam, 2006) Given $g \in G$, let $b(g) \in A_+$ be the unique element in A_+ conjugate to the hyperbolic part $h(g)$ of g. Then

$$\lim_{m \to \infty} [a_+(g^m)]^{1/m} = b(g).$$

Journal of Japan Math. Soc. 58 (2006) 1197-1202.

THANK YOU FOR YOUR ATTENTION