AN ELEMENTARY PROOF OF THE SPECTRAL RADIUS FORMULA FOR MATRICES

JOEL A. TROPP

Abstract. We present an elementary proof that the spectral radius of a matrix A may be obtained using the formula

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{1/n},$$

where $\|\cdot\|$ represents any matrix norm.

1. Introduction

It is a well-known fact from the theory of Banach algebras that the spectral radius of any element A is given by the formula

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{1/n}. \tag{1.1}$$

For a matrix, the spectrum is just the collection of eigenvalues, so this formula yields a technique for estimating for the top eigenvalue. The proof of Equation 1.1 is beautiful but advanced. See, for example, Rudin's treatment in his *Functional Analysis*. It turns out that elementary techniques suffice to develop the formula for matrices.

2. Preliminaries

For completeness, we shall briefly introduce the major concepts required in the proof. It is expected that the reader is already familiar with these ideas.

2.1. **Norms.** A *norm* is a mapping $\|\cdot\|$ from a vector space X into the nonnegative real numbers $\mathbb{R}^+$ which has three properties:

(1) $\|x\| = 0$ if and only if $x = 0$;
(2) $\|\alpha x\| = |\alpha|\,\|x\|$ for any scalar α and vector x; and
(3) $\|x + y\| \le \|x\| + \|y\|$ for any vectors x and y.

The most fundamental example of a norm is the Euclidean norm $\|\cdot\|_2$ which corresponds to the standard topology on $\mathbb{R}^n$. It is defined by

$$\|x\|_2 = \|(x_1, x_2, \dots, x_n)\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Date: 30 November 2001.

One particular norm we shall consider is the ℓ_∞ norm which is defined for x in $\mathbb{R}^n$ by the formula

$$\|x\|_\infty = \|(x_1, x_2, \ldots, x_n)\|_\infty = \max_i \{x_i\}.$$

For a matrix A, define the infinity norm as

$$\|A\|_\infty = \max_i \textstyle\sum_j |A_{ij}|.$$

This norm is consistent with itself and with the ℓ_∞ vector norm. That is,

$$\begin{aligned} \|AB\|_\infty &\leq \|A\|_\infty \|B\|_\infty \qquad \text{and} \\ \|Ax\|_\infty &\leq \|A\|_\infty \|x\|_\infty, \end{aligned}$$

where A and B are matrices and x is a vector.

Two norms $\|\cdot\|$ and $\|\cdot\|_\star$ on a vector space X are said be *equivalent* if there are positive constants $\underset{\sim}{C}$ and $\bar{C}$ such that

$$\underset{\sim}{C}\|x\| \leq \|x\|_\star \leq \bar{C}\|x\|$$

for every vector x. For finite-dimensional spaces, we have the following powerful result.

Theorem 2.1. *All norms on a finite-dimensional vector space are equivalent.*

Proof. We shall demonstrate that any norm $\|\cdot\|$ on $\mathbb{R}^n$ is equivalent to the Euclidean norm. Let $\{e_i\}$ be the canonical basis for $\mathbb{R}^n$, so any vector has an expression as $x = \sum x_i e_i$. First, let us check that $\|\cdot\|$ is continuous with respect to the Euclidean norm. For all pairs of vectors x and y,

$$\begin{aligned} \|x - y\| &= \|\textstyle\sum (x_i - y_i) e_i\| \\ &\leq \textstyle\sum |x_i - y_i| \, \|e_i\| \\ &\leq \max_i\{\|e_i\|\} \textstyle\sum |x_i - y_i| \\ &\leq M \left\{\textstyle\sum (x_i - y_i)^2\right\}^{1/2} \\ &= M\|x - y\|_2, \end{aligned}$$

where $M = \max_i\{\|e_i\|\}$. In other words, when two vectors are nearby with respect to the Euclidean norm, they are also nearby with respect to any other norm. Notice that the Cauchy-Schwarz inequality for real numbers has played a starring role at this stage.

Now, consider the unit sphere with respect to the Euclidean norm, $S = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$. This set is evidently closed and bounded in the Euclidean topology, so the Heine-Borel theorem shows that it is compact. Therefore, the continuous function $\|\cdot\|$ attains maximum and minimum values on S, say $\underline{C}$ and $\bar{C}$. That is,

$$\underline{C}\|x\|_2 \le \|x\| \le \bar{C}\|x\|_2$$

for any x with unit Euclidean norm. But every vector y can be expressed as $y = \alpha x$ for some x on the Euclidean unit sphere. If we multiply the foregoing inequality by $|\alpha|$ and draw the scalar into the norms, we reach

$$\underline{C}\|y\|_2 \le \|y\| \le \bar{C}\|y\|_2$$

for any vector y.

It remains to check that the constants $\underline{C}$ and $\bar{C}$ are positive. They are clearly nonnegative since $\|\cdot\|$ is nonnegative, and $\underline{C} \le \bar{C}$ by definition. Assume that $\underline{C} = 0$, which implies the existence of a point x on the Euclidean unit sphere for which $\|x\| = 0$. But then $x = 0$, a contradiction. □

2.2. **The spectrum of a matrix.** For an n-dimensional matrix A, consider the equation

$$Ax = \lambda x, \tag{2.1}$$

where x is a nonzero vector and λ is a complex number. Numbers λ which satisfy Equation 2.1 are called *eigenvalues* and the corresponding x are called *eigenvectors*. Nonzero vector solutions to this equation exist if and only if

$$\det(A - \lambda I) = 0, \tag{2.2}$$

where I is the identity matrix. The left-hand side of Equation 2.2 is called the *characteristic polynomial* of A because it is a polynomial in λ of degree n whose solutions are identical with the eigenvalues of A. The *algebraic multiplicity* of an eigenvalue λ is the multiplicity of λ as a root of the characteristic polynomial. Meanwhile, the *geometric multiplicity* of λ is the number of linearly independent eigenvectors corresponding to this eigenvalue. The geometric multiplicity of an eigenvalue never exceeds the algebraic multiplicity. Now, the collection of eigenvalues of a matrix, along with their geometric and algebraic multiplicities, completely determines the *eigenstructure* of the matrix. It turns out that all matrices with the same eigenstructure are similar to each other. That is, if A and B have the same eigenstructure, there exists a nonsingular matrix S such that $S^{-1}AS = B$.

We call the set of all eigenvalues of a matrix A its *spectrum*, which is written as $\sigma(A)$. The *spectral radius* $\rho(A)$ is defined by

$$\rho(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\}.$$

In other words, the spectral radius measures the largest magnitude attained by any eigenvalue.

2.3. **Jordan canonical form.** We say that a matrix is in *Jordan canonical form* if it is block-diagonal and each block has the form

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdots & \cdot & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}^{d\times d}.$$

It can be shown that the lone eigenvalue of this *Jordan block* is λ. Moreover, the geometric multiplicity of λ is exactly one and the algebraic multiplicity of λ is exactly d, the block size. The eigenvalues of a block-diagonal matrix are simply the eigenvalues of its blocks with the algebraic and geometric multiplicities of identical eigenvalues summed across the blocks. Therefore, a diagonal matrix composed of Jordan blocks has its eigenstructure laid bare. Using the foregoing facts, it is easy to construct a matrix in Jordan canonical form which has any eigenstructure whatsoever. Therefore, every matrix is similar to a matrix in Jordan canonical form.

Define the choose function $\binom{n}{k}$ according to the following convention:

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!} & \text{when } k = 0, \ldots, n \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 2.1. *If J is a Jordan block with eigenvalue λ, then the components of its nth power satisfy*

$$(J^n)_{ij} = \binom{n}{j-i} \lambda^{n-j+i}. \tag{2.3}$$

Proof. For $n = 1$, it is straightforward to verify that Equation 2.3 yields

$$J^1 = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdots & \cdot & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}^{d\times d}$$

Now, for arbitrary indices i and j, use induction to calculate that

$$\begin{aligned}\left(J^{n+1}\right)_{ij} &= \sum_{k=1}^{d} \left(J^n\right)_{ik} J_{kj} \\ &= \sum_{k=1}^{d} \left\{ \binom{n}{k-i} \lambda^{n-k+i} \right\} \left\{ \binom{1}{j-k} \lambda^{1-j+k} \right\} \\ &= \binom{n}{j-i} \lambda^{(n-j+i)+1} + \binom{n}{j-(i+1)} \lambda^{n-j+(i+1)} \\ &= \binom{n+1}{j-i} \lambda^{(n+1)-j+i}\end{aligned}$$

as advertised. □

3. The spectral radius formula

First, we prove the following special case.

Theorem 3.1. *For any matrix A, the spectral radius formula holds for the infinity matrix norm:*

$$\|A^n\|_\infty^{1/n} \longrightarrow \rho(A).$$

Proof. Throughout this argument, we shall denote the ℓ_∞ vector and matrix norms by $\|\cdot\|$.

Let S be a similarity transform such that $S^{-1}AS$ has Jordan form:

$$J = S^{-1}AS = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}.$$

Using the consistency of $\|\cdot\|$, we develop the following bounds.

$$\begin{aligned}\|A^n\|^{1/n} &= \|SJ^nS^{-1}\|^{1/n} \\ &\le \left\{\|S\|\|S^{-1}\|\right\}^{1/n} \|J^n\|^{1/n},\end{aligned}$$

and

$$\begin{aligned}\|A^n\|^{1/n} &= \left\{ \frac{\|S^{-1}\|\|SJ^nS^{-1}\|\|S\|}{\|S\|\|S^{-1}\|} \right\}^{1/n} \\ &\ge \left\{\|S\|\|S^{-1}\|\right\}^{-1/n} \|J^n\|^{1/n}.\end{aligned}$$

In each inequality, the former term on the right-hand side tends toward one as n approaces infinity. Therefore, we need only investigate the behavior of $\|J^n\|^{1/n}$.

Now, the matrix J is block-diagonal, so its powers are also block-diagonal with the blocks exponentiated individually:

$$J^n = \begin{bmatrix} J_1^n & & \\ & \ddots & \\ & & J_s^n \end{bmatrix}.$$

Since we are using the infinity norm,

$$\|J^n\| = \max_k \{\|J_k^n\|\}.$$

The nth root is monotonic, so we may draw it inside the maximum to obtain

$$\|J^n\|^{1/n} = \max_k \left\{\|J_k^n\|^{1/n}\right\}.$$

What is the norm of an exponentiated Jordan block? Recall the fact that the infinity norm of a matrix equals the greatest absolute row sum, and apply it to the explicit form of J_k^n provided in Lemma 2.1.

$$\begin{aligned} \|J_k^n\| &= \sum_{j=1}^{d_k} |(J_k^n)_{1j}| \\ &= \sum_{j=1}^{d_k} \binom{n}{j-1} |\lambda_k|^{n-j+1} \\ &= |\lambda_k|^n \left\{ |\lambda_k|^{1-d_k} \sum_{j=1} d_k \binom{n}{j-1} |\lambda_k|^{d_k - j} \right\}, \end{aligned}$$

where λ_k is the eigenvalue of block J_k and d_k is the block size. Bound the choose function above and below with $1 \leq \binom{n}{j-1} \leq n^{d_k}$, and write $M_k = |\lambda_k|^{1-d_k} \sum_j |\lambda_k|^{d_k - j}$ to obtain the relation

$$M_k \, |\lambda_k|^n \leq \|J_k^n\| \leq M_k \, n^{d_k} \, |\lambda_k|^n.$$

Extracting the nth root and taking the limit as n approaches infinity, we reach

$$\lim_{n\to\infty} \|J_k^n\|^{1/n} = |\lambda_k|.$$

A careful reader will notice that the foregoing argument does not apply to a Jordan block J_k with a zero eigenvalue. But such a matrix is nilpotent: placing a large exponent on J_k yields the zero matrix. The norm of a zero matrix is zero, so we have

$$\lim_{n\to\infty} \|J_k^n\|^{1/n} = 0.$$

Combining these facts, we conclude that

$$\lim_{n\to\infty} \|J_n\|^{1/n} = \max_k \{|\lambda_k|\} = \rho(J).$$

The spectrum of the Jordan form J is identical with the spectrum of A, which completes the proof. □

It is quite easy to bootstrap the general result from this special case.

Corollary 3.1. *The spectral radius formula holds for any matrix and any norm:*

$$\|A^n\|^{1/n} \longrightarrow \rho(A).$$

Proof. Theorem 2.1 on the equivalence of norms yields the inequality

$$\underline{C}\|A^n\|_\infty \le \|A^n\| \le \bar{C}\|A^n\|_\infty$$

for positive numbers $\underline{C}$ and $\bar{C}$. Extract the nth root of this inequality, and take the limit. The root drives the constants toward one, which leaves the relation

$$\lim_{n\to\infty} \|A^n\|_\infty^{1/n} \le \lim_{n\to\infty} \|A^n\|^{1/n} \le \lim_{n\to\infty} \|A^n\|_\infty^{1/n}.$$

Apply Theorem 3.1 to the upper and lower bounds to reach

$$\lim_{n\to\infty} \|A^n\|^{1/n} = \rho(A).$$

□