

16.4 The Spectral Radius

Every criterion in Chapter 10 §10 began by looking at the largest modulus of an eigenvalue, and that section gave the number a name in passing — the spectral radius — without a numbered definition and without a single norm available to compare it with. It also left a note: “Chapter 16 measures matrices by norms, proves that the spectral radius never exceeds the norm of $\mathbf{A}$ induced by any vector norm, and recovers $\rho(\mathbf{A})$ as a limit built from the norms of the powers $\mathbf{A}^m$.” This section pays that note in full, and is careful about what is new when it does.

Throughout, matrices are complex. A real matrix is read inside $M_n(\mathbb{C})$, exactly as Chapter 10 §10 read it when it counted eigenvalues; over $\mathbb{R}$ a matrix may have no eigenvalue at all, and then there is nothing to take the largest modulus of.

16.4.1 The number Chapter 10 named

The spectral radius is the radius of the smallest disk about the origin that holds the whole spectrum.

Definition 16.4.1. Spectral Radius

Let $\mathbf{A} \in M_n(\mathbb{C})$ with $n \geq 1$. The **spectral radius** of $\mathbf{A}$ is

$$\rho(\mathbf{A}) := \max\{|\lambda| : \lambda \in \text{spec}(\mathbf{A})\}.$$

For an operator T on a nonzero finite-dimensional complex vector space, $\rho(T)$ is defined by the same formula with $\text{spec}(T)$ in place of $\text{spec}(\mathbf{A})$.

The maximum is over a **finite nonempty** set, so it exists with no analysis at all: $\text{spec}(\mathbf{A})$ is the root set of the characteristic polynomial $p_{\mathbf{A}}$, which is nonempty by the Fundamental Theorem of Algebra (Theorem 6.5.1) and has at most n elements. It is the largest modulus of an eigenvalue, **not** the largest eigenvalue: over $\mathbb{C}$ there is no order to take a largest in, and even for a real spectrum $\rho(\text{diag}(-5, 1)) = 5$, not 1. This is the number Chapter 10 §10 introduced in prose, now with a label so that later sections can cite it.

Some values, simplest first.

- **Diagonal or triangular.** The eigenvalues of a triangular matrix are its diagonal entries (Theorem 9.7.4 (b)), so ρ is the largest modulus on the diagonal. For $\text{diag}(d_1, \dots, d_n)$ it is $\max_i |d_i|$.
- **Normal.** If $\mathbf{A}$ is normal then $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$ (Corollary 12.4.3), so $\mathbf{A}^* \mathbf{A} = \mathbf{U} \text{diag}(|\lambda_i|^2) \mathbf{U}^*$ and the singular values of $\mathbf{A}$ are the moduli $|\lambda_i|$. Hence $\rho(\mathbf{A}) = \sigma_1(\mathbf{A}) = \|\mathbf{A}\|_2$ by Theorem 16.3.3 (c). For a normal matrix the spectral radius **is** the spectral norm, which is where the second name comes from.
- **A rotation.** $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues $\pm i$, so $\rho = 1$, although no real vector is left in place.
- **Nilpotent.** $\mathbf{J}_k(0)$ is triangular with zero diagonal, so by the first bullet $\rho(\mathbf{J}_k(0)) = 0$. The matrix itself is not $\mathbf{O}$ when $k \geq 2$.

The last entry is the degenerate case, and it is the important one: it says at once that ρ is not a norm.

Warning. ρ is not a norm, and it is neither subadditive nor submultiplicative. Take

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Both are nilpotent, so $\rho(\mathbf{A}) = \rho(\mathbf{B}) = 0$ although neither is the zero matrix. Their sum $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues ± 1 and their product $\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has eigenvalues 1 and 0. So $\rho(\mathbf{A} + \mathbf{B}) = 1 > 0 = \rho(\mathbf{A}) + \rho(\mathbf{B})$ and $\rho(\mathbf{AB}) = 1 > 0 = \rho(\mathbf{A})\rho(\mathbf{B})$. So positive definiteness and subadditivity both fail, and so does submultiplicativity, all on the same small pair. Only absolute homogeneity survives, since $\rho(c\mathbf{A}) = |c|\rho(\mathbf{A})$ for every scalar c .

So ρ is a genuinely different kind of measurement from the ones of the previous section. The relation between the two is the subject of the rest of this section, and it is one-sided.

16.4.2 Every norm dominates the spectral radius

Theorem 16.4.1. The Spectral Radius Is a Lower Bound for Every Norm

Let $\mathbf{A} \in M_n(\mathbb{C})$.

- (a) If $\|\cdot\|$ is the operator norm on $M_n(\mathbb{C})$ induced by a norm on $\mathbb{C}^n$ (Definition 16.3.2), then $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$.
- (b) More generally, $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$ for **every** matrix norm $\|\cdot\|$ on $M_n(\mathbb{C})$ (Definition 16.3.1).

Idea. For (a) there is nothing to do but write down the eigenvalue equation: an eigenvector is a vector on which $\mathbf{A}$ acts by multiplication by λ , so the stretching factor $|\lambda|$ is one of the factors the maximum in Definition 16.3.2 ranges over. For (b) there is no vector to feed the norm, only matrices — so manufacture one: fill an $n \times n$ matrix with copies of the eigenvector, and the eigenvalue equation becomes a matrix equation that submultiplicativity can be applied to.

Proof.

- (a) Let $\lambda \in \text{spec}(\mathbf{A})$ and let $\mathbf{x} \neq \mathbf{0}$ with $\mathbf{Ax} = \lambda\mathbf{x}$. Then $\|\mathbf{Ax}\| = |\lambda| \|\mathbf{x}\|$ by homogeneity, while $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|$ by Theorem 16.3.2 (a). Since $\mathbf{x} \neq \mathbf{0}$ we have $\|\mathbf{x}\| > 0$ and may divide, giving $|\lambda| \leq \|\mathbf{A}\|$. This holds for every eigenvalue, so it holds for the largest modulus.
- (b) Let λ and $\mathbf{x}$ be as above, and let $\mathbf{X} \in M_n(\mathbb{C})$ be the matrix all of whose n columns equal $\mathbf{x}$. Then $\mathbf{X} \neq \mathbf{O}$, so $\|\mathbf{X}\| > 0$. Computing column by column, $\mathbf{AX}$ has every column equal to $\mathbf{Ax} = \lambda\mathbf{x}$, that is,

$$\mathbf{AX} = \lambda\mathbf{X}.$$

Taking norms and using homogeneity on the left and submultiplicativity on the right,

$$|\lambda| \|\mathbf{X}\| = \|\mathbf{A}\mathbf{X}\| \leq \|\mathbf{A}\| \|\mathbf{X}\|.$$

Dividing by $\|\mathbf{X}\| > 0$ gives $|\lambda| \leq \|\mathbf{A}\|$, and again the largest modulus obeys the same bound. This proves the theorem. ■

This is the inequality Chapter 10 §10 promised: the spectral radius never exceeds the norm of $\mathbf{A}$ induced by any vector norm. Part (b) says more than was promised, and costs one extra line.

The immediate use is that every computable norm is a computable *upper bound* for every eigenvalue — and two of the three formulas of Theorem 16.3.3 need no arithmetic beyond adding up moduli.

Corollary 16.4.2. Cheap Bounds on the Spectrum

Let $\mathbf{A} \in M_n(\mathbb{C})$. Then every eigenvalue λ of $\mathbf{A}$ satisfies

$$|\lambda| \leq \rho(\mathbf{A}) \leq \min\{\|\mathbf{A}\|_1, \|\mathbf{A}\|_\infty, \|\mathbf{A}\|_F\},$$

that is, $\rho(\mathbf{A})$ is at most the largest absolute column sum, at most the largest absolute row sum, and at most $\|\mathbf{A}\|_F$.

Proof. The first two are operator norms (Theorem 16.3.3 (a), (b)), so Theorem 16.4.1 (a) applies; $\|\cdot\|_F$ is a matrix norm (Example (The Frobenius norm is submultiplicative)), so Theorem 16.4.1 (b) applies. A minimum of upper bounds is an upper bound. ■

Remark. The bound $\rho(\mathbf{A}) \leq \|\mathbf{A}\|_2 = \sigma_1(\mathbf{A})$ is the same crude inequality $|\lambda| \leq \sigma_1$ that Chapter 13 §09 obtained directly from $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. It is now one instance of a general principle rather than an isolated computation, and Theorem 16.4.1 (a) is that computation with the norm left unnamed.

Example. Bounding a spectrum without finding it

Let $\mathbf{A} = \frac{1}{10} \begin{pmatrix} 2 & 5 & 1 \\ 3 & 1 & 4 \\ 1 & 2 & 5 \end{pmatrix}$. Show that $\rho(\mathbf{A}) \leq \frac{4}{5}$, and hence that $\mathbf{A}^m \rightarrow \mathbf{O}$.

Solution. Every row of $\mathbf{A}$ sums to $\frac{8}{10} = \frac{4}{5}$, so $\|\mathbf{A}\|_\infty = \frac{4}{5}$ by Theorem 16.3.3 (b), and $\rho(\mathbf{A}) \leq \frac{4}{5}$ by Corollary 16.4.2. Since $\frac{4}{5} < 1$, every eigenvalue has modulus less than 1 and $\mathbf{A}^m \rightarrow \mathbf{O}$ by Theorem 10.10.2. No characteristic polynomial was computed.

The column sums are $\frac{6}{10}$, $\frac{8}{10}$, and 1, so $\|\mathbf{A}\|_1 = 1$ and the column bound gives nothing.

It is worth computing both and keeping the smaller. In fact $\rho(\mathbf{A}) = \frac{4}{5}$ exactly, because the constant row sums make $\mathbf{1}$ an eigenvector with eigenvalue $\frac{4}{5}$, so the row bound here is attained.

Quick check. Theorem 16.4.1 says $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$ for every matrix norm. Why does it not follow that ρ is the smallest matrix norm?

Solution. Because ρ is not a matrix norm at all: ρ vanishes on every nonzero nilpotent matrix, and it fails both subadditivity and submultiplicativity, as the warning above showed. The correct statement is that $\rho(\mathbf{A})$ is a lower bound for the values $\|\mathbf{A}\|$, one matrix at a time, and the next theorem shows it is the greatest such lower bound.

16.4.3 A norm as close to the spectral radius as we like

Theorem 16.4.1 leaves open how large the gap can be. It can be enormous for a fixed norm: $\mathbf{J}_2(0)$ has $\rho = 0$ and $\|\mathbf{J}_2(0)\|_2 = 1$, and multiplying the corner entry by 10^6 leaves $\rho = 0$ while the norm becomes 10^6 . But the norm may be chosen after the matrix, and then the gap can be made as small as we please.

Theorem 16.4.3. Norms Close to the Spectral Radius

Let $\mathbf{A} \in M_n(\mathbb{C})$ and let $\varepsilon > 0$. Then there is a norm $\|\cdot\|_t$ on $\mathbb{C}^n$ whose induced operator norm satisfies

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|_t \leq \rho(\mathbf{A}) + \varepsilon.$$

Idea. Schur-triangularize and then squash. ① By Corollary 12.1.3, $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*$ with $\mathbf{T}$ upper triangular carrying the eigenvalues on its diagonal; the row sums of $\mathbf{T}$ are the diagonal entry plus the off-diagonal junk. ② Conjugating $\mathbf{T}$ by $\mathbf{D}_t = \text{diag}(1, t, \dots, t^{n-1})$ multiplies the entry in position (i, j) by t^{i-j} , which leaves the diagonal alone and, for large t , shrinks everything strictly above it, because $i - j < 0$ there. ③ Choose t so that every off-diagonal row sum is below ε ; the largest row sum is then below $\rho(\mathbf{A}) + \varepsilon$, and the largest row sum is a norm we can compute, by Theorem 16.3.3 (b). ④ Transport that norm back to $\mathbb{C}^n$ along the invertible matrix $\mathbf{D}_t\mathbf{U}^*$.

Proof. By Corollary 12.1.3 there are a unitary $\mathbf{U}$ and an upper triangular $\mathbf{T} = (t_{ij})$ with $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*$ and $t_{ii} = \lambda_i$, the eigenvalues of $\mathbf{A}$. For $t \geq 1$ put $\mathbf{D}_t = \text{diag}(1, t, t^2, \dots, t^{n-1})$, an invertible matrix, and $\mathbf{T}_t = \mathbf{D}_t\mathbf{T}\mathbf{D}_t^{-1}$. Multiplying out,

$$(\mathbf{T}_t)_{ij} = t^{i-1} t_{ij} t^{-(j-1)} = t^{i-j} t_{ij}.$$

So $\mathbf{T}_t$ is again upper triangular with the same diagonal $\lambda_1, \dots, \lambda_n$, while for $i < j$ the entry is scaled by $t^{i-j} \leq t^{-1}$, using $t \geq 1$ and $j - i \geq 1$. Let $M = \max_{i < j} |t_{ij}|$ if $n \geq 2$, and $M = 0$ if $n = 1$. Then the i -th absolute row sum of $\mathbf{T}_t$ is at most

$$|\lambda_i| + \sum_{j>i} t^{-1} |t_{ij}| \leq \rho(\mathbf{A}) + \frac{(n-1)M}{t}.$$

Choose $t = \max(1, (n-1)M/\varepsilon)$. Then every absolute row sum of $\mathbf{T}_t$ is at most $\rho(\mathbf{A}) + \varepsilon$, so $\|\mathbf{T}_t\|_\infty \leq \rho(\mathbf{A}) + \varepsilon$ by Theorem 16.3.3 (b).

Now define, for $\mathbf{x} \in \mathbb{C}^n$,

$$\|\mathbf{x}\|_t := \|\mathbf{D}_t \mathbf{U}^* \mathbf{x}\|_\infty.$$

This is a norm: it is nonnegative, homogeneous, and subadditive because $\|\cdot\|_\infty$ is and $\mathbf{x} \mapsto \mathbf{D}_t \mathbf{U}^* \mathbf{x}$ is linear, and it vanishes only at $\mathbf{x} = \mathbf{0}$ because $\mathbf{D}_t \mathbf{U}^*$ is invertible.

It remains to compute the induced operator norm. Put $\mathbf{Q} = \mathbf{D}_t \mathbf{U}^*$, so $\|\mathbf{x}\|_t = \|\mathbf{Q}\mathbf{x}\|_\infty$. As $\mathbf{x}$ runs over $\{\|\mathbf{x}\|_t = 1\}$, the vector $\mathbf{y} = \mathbf{Q}\mathbf{x}$ runs over $\{\|\mathbf{y}\|_\infty = 1\}$, since $\mathbf{Q}$ is a bijection. Moreover $\mathbf{Q}\mathbf{A}\mathbf{Q}^{-1} = \mathbf{D}_t \mathbf{U}^* \mathbf{A} \mathbf{U} \mathbf{D}_t^{-1} = \mathbf{D}_t \mathbf{T} \mathbf{D}_t^{-1} = \mathbf{T}_t$, so

$$\|\mathbf{A}\mathbf{x}\|_t = \|\mathbf{Q}\mathbf{A}\mathbf{x}\|_\infty = \|\mathbf{T}_t \mathbf{y}\|_\infty.$$

Taking maxima over the two matching sets gives $\|\mathbf{A}\|_t = \|\mathbf{T}_t\|_\infty \leq \rho(\mathbf{A}) + \varepsilon$. The lower bound $\rho(\mathbf{A}) \leq \|\mathbf{A}\|_t$ is Theorem 16.4.1 (a). This proves the theorem. ■

Read together, the last two theorems say

$$\rho(\mathbf{A}) = \inf\{\|\mathbf{A}\| : \|\cdot\| \text{ an induced matrix norm}\}.$$

Remark. The infimum need not be attained. If $\mathbf{A} \neq \mathbf{0}$ is nilpotent then $\rho(\mathbf{A}) = 0$ while $\|\mathbf{A}\| > 0$ for every norm, so no norm achieves the value. At the other extreme, for a normal $\mathbf{A}$ the spectral norm achieves it exactly: $\|\mathbf{A}\|_2 = \rho(\mathbf{A})$, as computed at the start of this section. The gap between ρ and the norms is a measure of how far $\mathbf{A}$ is from being normal, which is the theme Chapter 12 §11 opened with the departure from normality.

16.4.4 Gelfand's formula

The infimum above is over norms, and to use it we must choose one. There is a second description of $\rho(\mathbf{A})$ that needs no choice: fix any norm at all, and let the powers of $\mathbf{A}$ tell you the answer.

Theorem 16.4.4. Gelfand's Formula

Let $\mathbf{A} \in M_n(\mathbb{C})$ and let $\|\cdot\|$ be any matrix norm on $M_n(\mathbb{C})$. Then

$$\lim_{k \rightarrow \infty} \|\mathbf{A}^k\|^{1/k} = \rho(\mathbf{A}).$$

In particular the limit exists and does not depend on which matrix norm is used.

Idea. Two inequalities, one for each direction, and they come from opposite ends of the chapter. ① The lower bound $\rho(\mathbf{A}) \leq \|\mathbf{A}^k\|^{1/k}$ is [Theorem 16.4.1](#) applied to $\mathbf{A}^k$, once we know $\rho(\mathbf{A}^k) = \rho(\mathbf{A})^k$ — which is the Spectral Mapping Theorem. This holds for every k , with no limit taken. ② For the upper bound, scale $\mathbf{A}$ down by a whisker more than $\rho(\mathbf{A})$. The scaled matrix has spectral radius below 1, so Chapter 10 §10's criterion makes its powers tend to $\mathbf{O}$, and once they are below 1 in norm the inequality we want has been written down. [Theorem 16.4.3](#) says the same decay is geometric, which is what the corollary below extracts.

Proof. Lower bound. Fix $k \geq 1$. By Theorem 9.10.1 applied to $q(x) = x^k$, $\text{spec}(\mathbf{A}^k) = \{\lambda^k : \lambda \in \text{spec}(\mathbf{A})\}$, so

$$\rho(\mathbf{A}^k) = \max_{\lambda \in \text{spec}(\mathbf{A})} |\lambda|^k = \rho(\mathbf{A})^k.$$

By [Theorem 16.4.1](#) (b) applied to $\mathbf{A}^k$, $\rho(\mathbf{A})^k = \rho(\mathbf{A}^k) \leq \|\mathbf{A}^k\|$. Taking k -th roots of these nonnegative reals,

$$\rho(\mathbf{A}) \leq \|\mathbf{A}^k\|^{1/k} \quad \text{for every } k \geq 1.$$

Upper bound. Let $\varepsilon > 0$, put $s = \rho(\mathbf{A}) + \varepsilon > 0$ and $\mathbf{B} = s^{-1}\mathbf{A}$. Every eigenvalue of $\mathbf{B}$ is λ/s for some $\lambda \in \text{spec}(\mathbf{A})$ (Theorem 9.10.1 with $q(x) = x/s$), so

$$\rho(\mathbf{B}) = \frac{\rho(\mathbf{A})}{\rho(\mathbf{A}) + \varepsilon} < 1.$$

By Theorem 10.10.2, $\mathbf{B}^k \rightarrow \mathbf{O}$ entrywise. Entrywise convergence of matrices is convergence in **every** norm on $M_n(\mathbb{C})$, by Corollary 16.2.4, so $\|\mathbf{B}^k\| \rightarrow 0$. A real sequence tending to 0 is eventually below 1: there is K with $\|\mathbf{B}^k\| \leq 1$ for all $k \geq K$. Since $\mathbf{B}^k = s^{-k}\mathbf{A}^k$, homogeneity gives $\|\mathbf{A}^k\| \leq s^k$, that is,

$$\|\mathbf{A}^k\|^{1/k} \leq \rho(\mathbf{A}) + \varepsilon \quad (k \geq K).$$

Conclusion. Combining, for every $\varepsilon > 0$ there is K with

$$\rho(\mathbf{A}) \leq \|\mathbf{A}^k\|^{1/k} \leq \rho(\mathbf{A}) + \varepsilon \quad (k \geq K),$$

which is the definition of $\|\mathbf{A}^k\|^{1/k} \rightarrow \rho(\mathbf{A})$. The independence of the norm is automatic, since the limit equals $\rho(\mathbf{A})$ whichever matrix norm was fixed at the start. This proves the theorem.



Remark. Which analysis was used. The lower bound used none. The upper bound used Theorem 10.10.2, which is a theorem of this book, and Corollary 16.2.4, likewise; the only analysis is that a real sequence tending to 0 is eventually less than 1, and that a sequence trapped between $\rho(\mathbf{A})$ and $\rho(\mathbf{A}) + \varepsilon$ for every $\varepsilon > 0$ converges to $\rho(\mathbf{A})$. Neither is an extra import: both are the definition of a limit of a real sequence, read forwards and backwards. Nothing from the list (A1)–(A4) of the chapter introduction is needed here, and nothing outside it is used.

The formula is a genuine computation rule as well as a theorem. For a fixed k the number $\|\mathbf{A}^k\|^{1/k}$ is an *upper* estimate of $\rho(\mathbf{A})$ — that is the lower bound above, read the other way — and the estimates improve as k grows.

Example. Watching the limit converge

Let $\mathbf{A} = \begin{pmatrix} 1/2 & 10 \\ 0 & 1/3 \end{pmatrix}$. Compute $\|\mathbf{A}^k\|_\infty$ exactly and confirm Gelfand's formula.

Solution. $\mathbf{A}$ is upper triangular, so $\rho(\mathbf{A}) = \max(\frac{1}{2}, \frac{1}{3}) = \frac{1}{2}$ and the powers are upper triangular with diagonal $2^{-k}, 3^{-k}$. For the corner entry, induction on k gives

$$(\mathbf{A}^k)_{12} = 10 \sum_{j=0}^{k-1} 2^{-j} 3^{-(k-1-j)} = 60(2^{-k} - 3^{-k}),$$

the last step by summing the geometric series with ratio $3/2$. The two absolute row sums are $2^{-k} + 60(2^{-k} - 3^{-k})$ and 3^{-k} , the first being the larger, so by Theorem 16.3.3 (b)

$$\|\mathbf{A}^k\|_\infty = 61 \cdot 2^{-k} - 60 \cdot 3^{-k}.$$

At $k = 1$ this is 10.5, which overestimates $\rho(\mathbf{A}) = \frac{1}{2}$ by a factor of 21; the k -th roots run 10.5, 2.93, 1.75, 1.32, 1.11, 0.98, ... and settle toward $\frac{1}{2}$, which is what [Theorem 16.4.4](#) guarantees. The convergence is slow, and it has to be: the single number $\|\mathbf{A}\|_\infty$ cannot know that the entry 10 sits above the diagonal rather than on it.

16.4.5 What is new, and what is not

Corollary 16.4.5. Powers Tending to Zero, with a Rate

Let $\mathbf{A} \in M_n(\mathbb{C})$ and let $\|\cdot\|$ be any matrix norm on $M_n(\mathbb{C})$.

- (a) $\mathbf{A}^k \rightarrow \mathbf{O}$ if and only if $\rho(\mathbf{A}) < 1$, and this holds if and only if $\|\mathbf{A}^k\| \rightarrow 0$.
- (b) If $\rho(\mathbf{A}) < 1$, then for every q with $\rho(\mathbf{A}) < q < 1$ there is a constant $C \geq 1$, depending

on $\mathbf{A}$, on q , and on the norm, such that $\|\mathbf{A}^k\| \leq C q^k$ for **every** $k \geq 0$.

Idea. Part (a) is Theorem 10.10.2 translated into norms by Corollary 16.2.4; nothing is proved that was not proved in Chapter 10. Part (b) is the new content, and it is [Theorem 16.4.3](#) doing what it was built for: choose a norm in which $\mathbf{A}$ itself is already smaller than q , take powers there, and pay one fixed constant to return to the norm we were given.

Proof.

- (a) The first equivalence is Theorem 10.10.2, since $\rho(\mathbf{A}) < 1$ says exactly that every eigenvalue has modulus less than 1. The second is Corollary 16.2.4: entrywise convergence of matrices and convergence in the norm $\|\cdot\|$ are the same thing.
- (b) Let $\varepsilon = q - \rho(\mathbf{A}) > 0$. By [Theorem 16.4.3](#) there is an induced matrix norm $\|\cdot\|_t$ on $M_n(\mathbb{C})$ with $\|\mathbf{A}\|_t \leq \rho(\mathbf{A}) + \varepsilon = q$. Submultiplicativity (Theorem 16.3.2 (d)) gives $\|\mathbf{A}^k\|_t \leq \|\mathbf{A}\|_t^k \leq q^k$ for every $k \geq 0$. By Theorem 16.2.2, applied to the two norms $\|\cdot\|$ and $\|\cdot\|_t$ on the finite-dimensional space $M_n(\mathbb{C})$, there is $C_0 > 0$ with $\|\mathbf{M}\| \leq C_0 \|\mathbf{M}\|_t$ for every $\mathbf{M}$. Taking $C = \max(C_0, 1)$,

$$\|\mathbf{A}^k\| \leq C_0 \|\mathbf{A}^k\|_t \leq C q^k.$$

This proves the corollary. ■

It is worth being exact about the accounting. Part (a) is **not** new. Chapter 10 §10 proved that $\mathbf{A}^m \rightarrow \mathbf{O}$ if and only if every eigenvalue has modulus less than 1, and it proved it with the Jordan form and no norms whatsoever; all that has happened here is that the conclusion has been restated in a language where “small” is a number. What is new is part (b): a **rate**. Chapter 10 could say that the powers die; it could not say how fast, because it had nothing to measure a matrix with. Now every decaying matrix comes with a geometric envelope Cq^k for every q above its spectral radius, and that envelope is what later sections integrate, sum, and differentiate.

Two honest limitations of the envelope deserve saying out loud.

Warning. $\rho(\mathbf{A}) < 1$ **promises eventual decay, not immediate decay**. The constant C in [Corollary 16.4.5](#) (b) is generally larger than 1 and grows as q approaches $\rho(\mathbf{A})$. For

$$\mathbf{A} = \begin{pmatrix} 9/10 & 10 \\ 0 & 9/10 \end{pmatrix}, \quad \mathbf{A}^k = \begin{pmatrix} (9/10)^k & 10k(9/10)^{k-1} \\ 0 & (9/10)^k \end{pmatrix},$$

we have $\rho(\mathbf{A}) = 9/10 < 1$, yet $\|\mathbf{A}^k\|_\infty = (9/10)^{k-1}(9/10 + 10k)$ **rises** from 10.9 at $k = 1$ to about 39.13 at $k = 9$ before beginning its decay, and does not fall below 1 until $k = 63$. A transient hump of this kind is invisible to the spectral radius and is exactly what the constant C is paying for.

Quick check. For which matrices can one take $C = 1$ and $q = \rho(\mathbf{A})$ in the bound of [Corollary 16.4.5](#) (b), with the spectral norm?

Solution. For normal $\mathbf{A}$, among others. If $\mathbf{A}$ is normal then so is $\mathbf{A}^k$ (its adjoint is $(\mathbf{A}^*)^k$, and powers of commuting matrices commute), so $\|\mathbf{A}^k\|_2 = \rho(\mathbf{A}^k) = \rho(\mathbf{A})^k$ by the computation at the start of this section together with Theorem 9.10.1. There is then no transient at all: the norm of the powers is exactly geometric. The hump in the warning is possible only because that matrix is far from normal, with departure from normality 10 in the sense of Definition 12.11.1.

16.4.6 Exercises

16.4.6.1 A. Check your understanding

Exercise. A1

- Define the **spectral radius**, and say why the maximum in the definition exists.
- State [Theorem 16.4.1](#) and [Theorem 16.4.4](#).
- Determine whether the following statement is correct, and justify your answer: ρ is a matrix norm on $M_n(\mathbb{C})$.
- Determine whether the following statement is correct, and justify your answer: if $\rho(\mathbf{A}) < 1$ then $\|\mathbf{A}\|_2 < 1$.
- Chapter 10 §10 already proved that $\mathbf{A}^m \rightarrow \mathbf{O}$ exactly when $\rho(\mathbf{A}) < 1$. What does [Corollary 16.4.5](#) add to that?

Solution.

- $\rho(\mathbf{A}) = \max\{|\lambda| : \lambda \in \text{spec}(\mathbf{A})\}$ ([Definition 16.4.1](#)). The set $\text{spec}(\mathbf{A})$ is finite and nonempty — it is the root set of $p_{\mathbf{A}}$, nonempty by Theorem 6.5.1 — so the maximum of the finitely many moduli exists.
- For every matrix norm on $M_n(\mathbb{C})$, $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$; and for every matrix norm, $\|\mathbf{A}^k\|^{1/k} \rightarrow \rho(\mathbf{A})$ as $k \rightarrow \infty$.
- Incorrect. ρ vanishes on nonzero nilpotent matrices, so it is not even a norm, and it is neither subadditive nor submultiplicative: see the warning after [Definition 16.4.1](#).
- Incorrect. $\mathbf{A} = \begin{pmatrix} 0 & 10 \\ 0 & 0 \end{pmatrix}$ has $\rho(\mathbf{A}) = 0$ and $\|\mathbf{A}\|_2 = 10$. The correct statement is that *some* norm has $\|\mathbf{A}\| < 1$, by [Theorem 16.4.3](#).
- Nothing about the criterion, which is Theorem 10.10.2 restated; what it adds is the quantitative envelope $\|\mathbf{A}^k\| \leq Cq^k$ for every $q \in (\rho(\mathbf{A}), 1)$, a rate of decay rather than the bare fact of it.

16.4.6.2 B. Practice

Exercise. B1: Four numbers again

Let $\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 1/4 & 0 \end{pmatrix}$. Compute $\rho(\mathbf{A})$, $\|\mathbf{A}\|_1$, $\|\mathbf{A}\|_\infty$, and $\|\mathbf{A}\|_2$, and verify [Theorem 16.4.1](#) in each case.

Solution. $p_{\mathbf{A}}(x) = x^2 - \frac{1}{2}$, so $\text{spec}(\mathbf{A}) = \{\pm 1/\sqrt{2}\}$ and $\rho(\mathbf{A}) = 1/\sqrt{2} \approx 0.707$.

The absolute column sums are $\frac{1}{4}$ and 2, so $\|\mathbf{A}\|_1 = 2$; the absolute row sums are 2 and $\frac{1}{4}$, so $\|\mathbf{A}\|_\infty = 2$. For the spectral norm, $\mathbf{A}^\top \mathbf{A} = \text{diag}(\frac{1}{16}, 4)$, whose eigenvalues are $\frac{1}{16}$ and 4; hence $\sigma_1 = 2$ and $\|\mathbf{A}\|_2 = 2$ by Theorem 16.3.3 (c).

So $\rho(\mathbf{A}) = 0.707 \cdots \leq 2$ in all three cases, and the gap is a factor of nearly 3 even though $\mathbf{A}$ is a very small matrix. The reason is the mismatch between the two off-diagonal entries: $\mathbf{A}$ is far from normal.

Exercise. B2: Gelfand exactly

With $\mathbf{A}$ as in [Exercise \(B1: Four numbers again\)](#), compute $\mathbf{A}^k$ for every k and evaluate $\lim_k \|\mathbf{A}^k\|_\infty^{1/k}$ directly.

Solution. $\mathbf{A}^2 = \frac{1}{2}\mathbf{I}$, so by induction $\mathbf{A}^{2l} = 2^{-l}\mathbf{I}$ and $\mathbf{A}^{2l+1} = 2^{-l}\mathbf{A}$. Hence

$$\|\mathbf{A}^{2l}\|_\infty = 2^{-l}, \quad \|\mathbf{A}^{2l+1}\|_\infty = 2^{-l} \cdot 2 = 2^{1-l}.$$

For even $k = 2l$, $\|\mathbf{A}^k\|_\infty^{1/k} = (2^{-l})^{1/(2l)} = 2^{-1/2}$, exactly $\rho(\mathbf{A})$ for every l . For odd $k = 2l + 1$, $\|\mathbf{A}^k\|_\infty^{1/k} = 2^{(1-l)/(2l+1)}$, and $(1-l)/(2l+1) \rightarrow -\frac{1}{2}$, so this too tends to $2^{-1/2}$. Both subsequences converge to $1/\sqrt{2} = \rho(\mathbf{A})$, confirming [Theorem 16.4.4](#).

Exercise. B3: Deciding convergence without eigenvalues

For each matrix, decide whether $\mathbf{A}^m \rightarrow \mathbf{O}$, using only [Corollary 16.4.2](#) and Theorem 10.10.2.

(a) $\mathbf{A} = \frac{1}{4} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix}.$

(b) $\mathbf{A} = \begin{pmatrix} 1/2 & 3 \\ 0 & 2 \end{pmatrix}.$

Solution.

- (a) The absolute row sums are $\frac{3}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$, so $\|\mathbf{A}\|_\infty = \frac{3}{4}$ and $\rho(\mathbf{A}) \leq \frac{3}{4} < 1$. By Theorem 10.10.2, $\mathbf{A}^m \rightarrow \mathbf{O}$.
- (b) Here the row sums are $\frac{7}{2}$ and 2 and the column sums are $\frac{1}{2}$ and 5, so the bounds give only $\rho(\mathbf{A}) \leq \frac{7}{2}$ and decide nothing. But the bound is an upper bound, and here the matrix is triangular, so the eigenvalues are visible: $\frac{1}{2}$ and 2, giving $\rho(\mathbf{A}) = 2 > 1$. By Theorem 10.10.2 the powers do **not** tend to $\mathbf{O}$. The moral is that [Corollary 16.4.2](#) can

only ever certify convergence, never rule it out.

16.4.6.3 C. Going deeper

Exercise. C1: A norm that sees the decay

Let $\mathbf{A} \in M_n(\mathbb{C})$.

- Prove that if $\|\mathbf{A}\| < 1$ for **some** matrix norm, then $\mathbf{A}^k \rightarrow \mathbf{O}$.
- Give a matrix with $\mathbf{A}^k \rightarrow \mathbf{O}$ and $\|\mathbf{A}\|_2 > 1$, so that the converse of (a) fails for a fixed norm.
- Prove that the converse of (a) nevertheless holds if the norm is allowed to depend on $\mathbf{A}$.

Solution.

- Submultiplicativity gives $\|\mathbf{A}^k\| \leq \|\mathbf{A}\|^k$ by induction on k . With $r = \|\mathbf{A}\| < 1$ the right-hand side tends to 0, so $\|\mathbf{A}^k\| \rightarrow 0$, and by Corollary 16.2.4 this is the same as $\mathbf{A}^k \rightarrow \mathbf{O}$ entrywise.
- $\mathbf{A} = \begin{pmatrix} 0 & 10 \\ 0 & 0 \end{pmatrix}$ satisfies $\mathbf{A}^2 = \mathbf{O}$, so certainly $\mathbf{A}^k \rightarrow \mathbf{O}$; but $\|\mathbf{A}\|_2 = 10 > 1$. The same happens for $\begin{pmatrix} 1/2 & 10 \\ 0 & 1/2 \end{pmatrix}$, which is not even nilpotent.
- Suppose $\mathbf{A}^k \rightarrow \mathbf{O}$. By Theorem 10.10.2, $\rho(\mathbf{A}) < 1$. Apply Theorem 16.4.3 with $\varepsilon = \frac{1}{2}(1 - \rho(\mathbf{A})) > 0$: it produces an induced matrix norm $\|\cdot\|_t$ with

$$\|\mathbf{A}\|_t \leq \rho(\mathbf{A}) + \varepsilon = \frac{1}{2}(1 + \rho(\mathbf{A})) < 1.$$

So a norm witnessing the decay always exists, though it must be chosen after $\mathbf{A}$.

Exercise. C2: Spectral radius zero

Let $\mathbf{A} \in M_n(\mathbb{C})$.

- Prove that $\rho(\mathbf{A}) = 0$ if and only if $\mathbf{A}$ is nilpotent.
- Deduce that for a nilpotent $\mathbf{A}$ and any matrix norm, $\|\mathbf{A}^k\|^{1/k} \rightarrow 0$, and verify this directly for $\mathbf{A} = \mathbf{J}_n(0)$.
- Explain why (a) shows that no inequality of the form $\|\mathbf{A}\| \leq c\rho(\mathbf{A})$, with c depending only on n , can hold.

Solution.

- ($\Rightarrow$) If $\rho(\mathbf{A}) = 0$ then every eigenvalue is 0, so $p_{\mathbf{A}}(x) = x^n$ because $p_{\mathbf{A}}$ is monic of degree n and splits over $\mathbb{C}$ with all roots 0 (Theorem 6.5.1). By the Cayley–Hamilton theorem (Theorem 9.6.1), $\mathbf{A}^n = p_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$, so $\mathbf{A}$ is nilpotent (Definition 10.2.1).
($\Leftarrow$) If $\mathbf{A}^d = \mathbf{O}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{0} = \mathbf{A}^d\mathbf{x} = \lambda^d\mathbf{x}$, so $\lambda^d = 0$ and $\lambda = 0$. Hence $\text{spec}(\mathbf{A}) = \{0\}$ and $\rho(\mathbf{A}) = 0$.
- By Theorem 16.4.4 the limit is $\rho(\mathbf{A}) = 0$. Directly for $\mathbf{A} = \mathbf{J}_n(0)$: $\mathbf{A}^k = \mathbf{O}$ for $k \geq n$,

so $\|\mathbf{A}^k\|^{1/k} = 0$ for all $k \geq n$, and the sequence is eventually constant at 0.

- (c) Such an inequality would force $\|\mathbf{A}\| = 0$, hence $\mathbf{A} = \mathbf{O}$, for every nilpotent $\mathbf{A}$. But $\mathbf{J}_2(0) \neq \mathbf{O}$ is nilpotent. So the two quantities cannot be compared in that direction at all; [Theorem 16.4.3](#) is the best available substitute, and it buys its conclusion by changing the norm.

Exercise. C3: When the infimum is attained

Call $\mathbf{A} \in M_n(\mathbb{C})$ **radius-attaining** if $\|\mathbf{A}\| = \rho(\mathbf{A})$ for some matrix norm.

- (a) Prove that every normal $\mathbf{A}$ is radius-attaining.
- (b) Prove that a nonzero nilpotent $\mathbf{A}$ is not.
- (c) Deduce that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not radius-attaining, even though it is invertible.

Hint for (c): consider $\mathbf{A} - \mathbf{I}$.

Solution.

- (a) Let $\mathbf{A}$ be normal, so $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$ (Corollary 12.4.3). Then $\mathbf{A}^*\mathbf{A} = \mathbf{U} \text{diag}(|\lambda_i|^2) \mathbf{U}^*$, so the singular values of $\mathbf{A}$ are the $|\lambda_i|$ in decreasing order and $\|\mathbf{A}\|_2 = \sigma_1 = \max_i |\lambda_i| = \rho(\mathbf{A})$ by Theorem 16.3.3 (c).
- (b) If $\mathbf{A} \neq \mathbf{O}$ is nilpotent then $\rho(\mathbf{A}) = 0$ by [Exercise \(C2: Spectral radius zero\)](#) (a), while $\|\mathbf{A}\| > 0$ for every norm, since a norm vanishes only at $\mathbf{O}$ (Definition 16.1.1). So no norm can give $\|\mathbf{A}\| = \rho(\mathbf{A})$.
- (c) Put $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, whose only eigenvalue is 1, so $\rho(\mathbf{A}) = 1$. Suppose some matrix norm had $\|\mathbf{A}\| = 1$. Then $\|\mathbf{A}^k\| \leq \|\mathbf{A}\|^k = 1$ for every k , so the sequence $(\mathbf{A}^k)$ is bounded in that norm, hence bounded entrywise by Corollary 16.2.4, whose second sentence covers boundedness. But $\mathbf{A}^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, whose $(1, 2)$ -entry is unbounded. Contradiction. (The hint points at the same phenomenon: $\mathbf{A} - \mathbf{I}$ is a nonzero nilpotent, so $\mathbf{A}$ is a unipotent matrix, and it is the nilpotent part that the spectral radius cannot see.)