

19.1 Positive Matrices and Perron's Theorem

Chapter 9 §11 found the long-run behavior of a Markov chain by diagonalizing its transition matrix, and warned that nothing guaranteed the method would work. Chapter 16 §04 gave every square matrix a spectral radius, the largest modulus of an eigenvalue, but said nothing about whether that modulus is itself an eigenvalue. For a general matrix it need not be: the swap $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has spectral radius 1 and the eigenvalue -1 sitting on the circle beside it, and the rotation by a right angle has no real eigenvalue at all. This section shows that a single hypothesis on the **entries** removes every one of these difficulties: if every entry of $\mathbf{A}$ is positive, then $\rho(\mathbf{A})$ is an eigenvalue, it is simple, it strictly dominates every other eigenvalue in modulus, and it has an eigenvector with positive entries. That is Perron's theorem, and the rest of the chapter extends it.

Throughout, matrices and vectors have real entries, except where complex eigenvectors appear. The eigenvalues of a real matrix are those of the same matrix read in $M_n(\mathbb{C})$, exactly as in Chapter 16 §04, so that $\rho(\mathbf{A})$ is defined (Definition 16.4.1) and eigenvectors may have complex entries.

19.1.1 The entrywise order

Chapter 18 §05 wrote $\mathbf{x} \geq \mathbf{0}$ for a real vector with every entry ≥ 0 . The same comparison, made one entry at a time, makes sense for matrices of any shape, and it is the only order this chapter uses. Its hook is a recurring expression: every probability argument of Chapter 9 §11 began “the entries are nonnegative, so each entry of the product is a sum of nonnegative terms”. We name the hypothesis once and prove its consequences once.

The entrywise order compares two matrices of the same shape position by position.

Definition 19.1.1. Entrywise Order; Nonnegative and Positive Matrices

Let $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ be real $m \times n$ matrices.

- (a) $\mathbf{A} \geq \mathbf{B}$ means $a_{ij} \geq b_{ij}$ for **every** i, j , and $\mathbf{A} > \mathbf{B}$ means $a_{ij} > b_{ij}$ for **every** i, j . We write $\mathbf{B} \leq \mathbf{A}$ and $\mathbf{B} < \mathbf{A}$ for the same statements.
- (b) $\mathbf{A}$ is **nonnegative** if $\mathbf{A} \geq \mathbf{O}$, and **positive** if $\mathbf{A} > \mathbf{O}$, where $\mathbf{O}$ is the zero matrix of the same shape.
- (c) For a complex $m \times n$ matrix $\mathbf{M} = (m_{ij})$, the **entrywise absolute value** is the real matrix $|\mathbf{M}| := (|m_{ij}|)$.

Vectors are the case $n = 1$, so $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} > \mathbf{0}$ and $|\mathbf{x}| = (|x_1|, \dots, |x_n|)$ have the same meaning as in Chapter 18 §05.

In words: (a) compares every pair of corresponding entries, and asks for a strict inequality in **every** position in the strict version. So $\mathbf{A} > \mathbf{B}$ is much stronger than “ $\mathbf{A} \geq \mathbf{B}$ and $\mathbf{A} \neq \mathbf{B}$ ”, and the two must not be confused. Clause (b) is the special case $\mathbf{B} = \mathbf{O}$. Clause (c) takes the modulus of each entry separately; it has nothing to do with the absolute value $(\mathbf{A}^* \mathbf{A})^{1/2}$ of Chapter 13, and in this chapter $|\mathbf{M}|$ always means the entrywise one.

Some examples, simplest first.

- $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} > \mathbf{O}$: every entry is positive. This is the matrix of Example (Eigenvalues and Eigenspaces of a 2×2 Matrix).
- $\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \geq \mathbf{O}$, but $\mathbf{I}_2$ is **not** positive, since its off-diagonal entries are 0. Every stochastic matrix (Definition 9.11.1) is nonnegative, and so is every permutation matrix.
- **The degenerate cases.** The zero matrix is nonnegative and not positive. A 1×1 matrix (a) is positive exactly when $a > 0$, and then everything this section proves reduces to the statement that a is an eigenvalue of (a) with eigenvector (1) . It is worth keeping in mind that every theorem below must survive $n = 1$.
- For $\mathbf{M} = \begin{pmatrix} 3 & -4i \\ 0 & -1 \end{pmatrix}$, $|\mathbf{M}| = \begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix}$.

Nonexample by minimal change. Take the positive matrix $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ and change its $(1, 2)$ entry to 0. The result $\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}$ is still nonnegative and still $\geq \mathbf{I}_2$, but it is not positive: the clause “**every** entry > 0 ” fails in one position. That one zero is enough to lose a conclusion of this section. The new matrix is triangular, with eigenvalues 1 and 2. Its eigenvectors for 2 are the multiples of $\mathbf{e}_2$, and those for 1 the multiples of $(1, -3)$. So the dominant eigenvalue 2 has a nonnegative eigenvector, but **no** positive one. Perron’s theorem below says this cannot happen for a positive matrix.

Warning. **Positive is not positive definite, and positive definite is not positive.** The two words sound alike and measure different things. The entrywise order looks at entries; positive definiteness (Definition 13.1.1) looks at the quadratic form. The matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is positive, and symmetric, but its eigenvalues are 3 and -1 , so it is not positive definite: $\mathbf{x} = (1, -1)$ gives $\mathbf{x}^\top \mathbf{A} \mathbf{x} = -2$. The matrix $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ is positive definite, with eigenvalues 1 and 3, but it has negative entries, so it is not even nonnegative. In this book $\mathbf{A} \geq \mathbf{O}$ and $\mathbf{A} > \mathbf{O}$ always mean the entrywise order, and the Loewner order is always written $\mathbf{A} \succeq \mathbf{O}$, $\mathbf{A} \succ \mathbf{O}$.

Why this definition. The order is entrywise because the only operations the theory needs are sums and products of entries, and a sum or product of nonnegative numbers is nonnegative. The hypothesis is not coordinate-free: a change of basis destroys it at once. That is also why this chapter, unlike most of the book, cares about the **standard** basis: the hypothesis $\mathbf{A} \geq \mathbf{O}$ is a statement about the matrix, not about the operator $T_{\mathbf{A}}$. The only changes of basis that preserve it are the ones that rearrange coordinates and rescale them by positive numbers, and §02 uses the first kind throughout.

The first facts are the ones used in every argument below. They are the three rules of arithmetic with nonnegative numbers, carried to matrices.

Lemma 19.1.1. Rules of the Entrywise Order

Let $\mathbf{A}, \mathbf{B} \in M_{m \times n}(\mathbb{R})$, $\mathbf{C}, \mathbf{D} \in M_{n \times p}(\mathbb{R})$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

- (a) If $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$ and $\mathbf{0} \leq \mathbf{x} \leq \mathbf{y}$, then $\mathbf{0} \leq \mathbf{A}\mathbf{x} \leq \mathbf{B}\mathbf{y}$.

- (b) If $\mathbf{A} > \mathbf{O}$, $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{Ax} > \mathbf{0}$.
 (c) If $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$ and $\mathbf{O} \leq \mathbf{C} \leq \mathbf{D}$, then $\mathbf{O} \leq \mathbf{AC} \leq \mathbf{BD}$. In particular, for square $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$, $\mathbf{O} \leq \mathbf{A}^k \leq \mathbf{B}^k$ for every $k \geq 0$.

Idea. Each entry of a product is a sum of products of entries, so every rule comes down to the arithmetic of nonnegative numbers. For the comparison in (a), write $\mathbf{By} - \mathbf{Ax}$ as a sum of two products of nonnegative factors; for (b), one positive term in a sum of nonnegative terms is enough; and (c) is (a) applied one column at a time.

Proof.

- (a) Each entry $(\mathbf{Ax})_i = \sum_j a_{ij}x_j$ is a sum of products of nonnegative numbers, so $\mathbf{Ax} \geq \mathbf{0}$. Next, $\mathbf{By} - \mathbf{Ax} = \mathbf{B}(\mathbf{y} - \mathbf{x}) + (\mathbf{B} - \mathbf{A})\mathbf{x}$, and both terms are nonnegative by what was just shown, because $\mathbf{B} \geq \mathbf{O}$, $\mathbf{y} - \mathbf{x} \geq \mathbf{0}$, $\mathbf{B} - \mathbf{A} \geq \mathbf{O}$ and $\mathbf{x} \geq \mathbf{0}$.
 (b) Since $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{x} \geq \mathbf{0}$, some entry x_k is positive. For every i , $(\mathbf{Ax})_i = \sum_j a_{ij}x_j \geq a_{ik}x_k > 0$, because the other terms are ≥ 0 and $a_{ik} > 0$.
 (c) By Theorem 0.9.2, column l of $\mathbf{AC}$ is $\mathbf{Ac}_l$, where $\mathbf{c}_l$ is column l of $\mathbf{C}$, and likewise for $\mathbf{BD}$. Since $\mathbf{0} \leq \mathbf{c}_l \leq \mathbf{d}_l$, part (a) gives $\mathbf{0} \leq \mathbf{Ac}_l \leq \mathbf{Bd}_l$ for every l . The statement about powers follows by induction on k : $\mathbf{A}^0 = \mathbf{B}^0 = \mathbf{I}$, and $\mathbf{A}^{k+1} = \mathbf{A}^k \mathbf{A} \leq \mathbf{B}^k \mathbf{B} = \mathbf{B}^{k+1}$ by the case just proved. ■

The absolute value interacts with products through the triangle inequality, which is the second tool.

Lemma 19.1.2. Absolute Value of a Product

Let $\mathbf{A} \in M_{m \times n}(\mathbb{C})$, $\mathbf{B} \in M_{n \times p}(\mathbb{C})$ and $\mathbf{x} \in \mathbb{C}^n$. Then

- (a) $|\mathbf{Ax}| \leq |\mathbf{A}| |\mathbf{x}|$;
 (b) $|\mathbf{AB}| \leq |\mathbf{A}| |\mathbf{B}|$;
 (c) if $m = n$, then $|\mathbf{A}^k| \leq |\mathbf{A}|^k$ for every $k \geq 0$.

Idea. Each entry of $\mathbf{Ax}$ is a sum of products, and the triangle inequality bounds its modulus by the sum of the moduli of the products: that is (a). Then (b) is (a) one column at a time, and (c) follows from (b) by induction on k , with the order rules to replace $|\mathbf{A}^k|$ by the larger $|\mathbf{A}|^k$.

Proof.

- (a) For each i , the triangle inequality (Theorem 0.6.4, extended to finitely many terms by induction) and the multiplicativity of the modulus (Theorem 0.6.2 (d)) give

$$|(\mathbf{Ax})_i| = \left| \sum_j a_{ij}x_j \right| \leq \sum_j |a_{ij}| |x_j| = (|\mathbf{A}| |\mathbf{x}|)_i.$$

- (b) Apply (a) to each column of $\mathbf{B}$, using Theorem 0.9.2 as in the previous proof.
 (c) Induction on k . For $k = 0$, $|\mathbf{I}| = \mathbf{I}$. If $|\mathbf{A}^k| \leq |\mathbf{A}|^k$, then by (b) and Lemma 19.1.1 (c),

$$|\mathbf{A}^{k+1}| = |\mathbf{A}^k \mathbf{A}| \leq |\mathbf{A}^k| |\mathbf{A}| \leq |\mathbf{A}|^k |\mathbf{A}| = |\mathbf{A}|^{k+1}.$$

This proves the lemma. ■

For a nonnegative $\mathbf{A}$ we have $|\mathbf{A}| = \mathbf{A}$, and (a) reads $|\mathbf{A}\mathbf{x}| \leq \mathbf{A}|\mathbf{x}|$. That is the form used below: whatever cancellation happens inside $\mathbf{A}\mathbf{x}$ can only make it smaller.

The proof of Perron's theorem also needs to know when the triangle inequality is an equality, a case Chapter 0 did not record. The answer is the expected one: when all the terms point in the same direction.

Lemma 19.1.3. Equality in the Triangle Inequality

Let $z_1, \dots, z_m \in \mathbb{C}$, put $s = z_1 + \dots + z_m$, and suppose $|s| = |z_1| + \dots + |z_m|$ and $s \neq 0$. Then $z_j = |z_j|c$ for every j , where $c = s/|s|$ has $|c| = 1$.

Idea. Rotate the sum onto the positive real axis: multiplying by $d = \bar{s}/|s|$ turns s into the real number $|s|$. Then $|s|$ is the real part of $\sum_j dz_j$, which is at most $\sum_j |z_j|$ term by term. Equality of the totals forces equality in every term, and a complex number whose real part equals its modulus is a real number ≥ 0 .

Proof. Recall from the proof of Theorem 0.6.4 that $\operatorname{Re} u = x \leq |x| \leq \sqrt{x^2 + y^2} = |u|$ for $u = x + yi$. Equality $\operatorname{Re} u = |u|$ forces equality in both steps, that is $y^2 = 0$ and $x = |x|$, so u is a real number ≥ 0 and $u = |u|$.

Put $d = \bar{s}/|s|$. Then $|d| = 1$ and $ds = |s|^2/|s| = |s|$ by Theorem 0.6.2 (c), (d). Hence

$$\sum_j |z_j| = |s| = ds = \operatorname{Re} \sum_j dz_j = \sum_j \operatorname{Re}(dz_j) \leq \sum_j |dz_j| = \sum_j |z_j|.$$

So the inequality in the middle is an equality. It is a sum of the inequalities $\operatorname{Re}(dz_j) \leq |dz_j|$, so each of them is an equality, and by the first paragraph $dz_j = |dz_j| = |z_j|$. Multiplying by $c = s/|s|$, which satisfies $cd = s\bar{s}/|s|^2 = 1$, gives $z_j = |z_j|c$. ■

19.1.2 The spectral radius is monotone

A first payoff joins the entrywise order to Chapter 16. Enlarging the entries of a nonnegative matrix should not shrink its eigenvalues, and for the spectral radius this is true.

Proposition 19.1.4. Monotonicity of the Spectral Radius

- (a) If $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{R})$ and $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$, then $\rho(\mathbf{A}) \leq \rho(\mathbf{B})$.
 (b) More generally, if $\mathbf{A} \in M_n(\mathbb{C})$, $\mathbf{B} \in M_n(\mathbb{R})$ and $|\mathbf{A}| \leq \mathbf{B}$, then $\rho(\mathbf{A}) \leq \rho(|\mathbf{A}|) \leq \rho(\mathbf{B})$.

Idea. Eigenvalues do not see the entries directly, but Gelfand's formula (Theorem 16.4.4) expresses ρ through the norms of the powers, and the largest absolute row sum $\|\cdot\|_\infty$ is a norm that does see the entries, monotonically. So compare the powers entrywise, take norms, take k -th roots, and pass to the limit.

Proof.

- (b) First, $|\mathbf{A}^k| \leq |\mathbf{A}|^k \leq \mathbf{B}^k$ for every $k \geq 1$: the first inequality is Lemma 19.1.2 (c), and the second is Lemma 19.1.1 (c) applied to $\mathbf{O} \leq |\mathbf{A}| \leq \mathbf{B}$. By Theorem 16.3.3 (b), $\|\mathbf{M}\|_\infty$ is the largest absolute row sum of $\mathbf{M}$, so it depends only on $|\mathbf{M}|$ and does not decrease when the entries of $|\mathbf{M}|$ increase. Hence

$$\|\mathbf{A}^k\|_\infty \leq \| |\mathbf{A}|^k \|_\infty \leq \|\mathbf{B}^k\|_\infty \quad (k \geq 1).$$

The k -th root is increasing on the nonnegative reals, so the same inequalities hold between the k -th roots. The operator norm $\|\cdot\|_\infty$ is a matrix norm (Theorem 16.3.2 (d)), so by Theorem 16.4.4 the three sequences of k -th roots converge to $\rho(\mathbf{A})$, $\rho(|\mathbf{A}|)$ and $\rho(\mathbf{B})$. A nonstrict inequality survives a limit, which gives $\rho(\mathbf{A}) \leq \rho(|\mathbf{A}|) \leq \rho(\mathbf{B})$.

- (c) If $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$, then $|\mathbf{A}| = \mathbf{A}$, and (b) gives $\rho(\mathbf{A}) \leq \rho(\mathbf{B})$. ■

The inequality need not be strict, even when $\mathbf{A} \neq \mathbf{B}$: $\text{diag}(1, 0) \leq \mathbf{I}_2$ and both have spectral radius 1. For positive matrices it is strict, which is Exercise C3. The hypothesis $\mathbf{O} \leq \mathbf{A}$ cannot be dropped from (a): $\text{diag}(-3, 0) \leq \mathbf{O}$ entrywise, yet $\rho(\text{diag}(-3, 0)) = 3 > 0 = \rho(\mathbf{O})$.

Example. Enlarging one entry

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$. Compute both spectral radii and compare with Proposition 19.1.4.

Solution. $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$, with a difference only in the $(2, 2)$ entry. By the 2×2 formula $p(x) = x^2 - (\text{tr})x + \det$, $p_{\mathbf{A}}(x) = x^2 - x - 2 = (x - 2)(x + 1)$, so $\rho(\mathbf{A}) = 2$, and $p_{\mathbf{B}}(x) = x^2 - 2x - 1$, with roots $1 \pm \sqrt{2}$, so $\rho(\mathbf{B}) = 1 + \sqrt{2} \approx 2.414$. Indeed $\rho(\mathbf{A}) \leq \rho(\mathbf{B})$. In both cases the eigenvalue of largest modulus is positive, which Perron's theorem below (Theorem 19.1.5) guarantees for the positive $\mathbf{B}$.

19.1.3 Perron's theorem

Here is the theorem this section is built around.

Theorem 19.1.5. Perron's Theorem

Let $n \geq 1$ and let $\mathbf{A} \in M_n(\mathbb{R})$ be **positive**, $\mathbf{A} > \mathbf{O}$. Write $r = \rho(\mathbf{A})$. Then:

- (a) $r > 0$, and r is an eigenvalue of $\mathbf{A}$ with an eigenvector $\mathbf{v} > \mathbf{0}$;
- (b) r is **algebraically simple**: $a_{\mathbf{A}}(r) = 1$. In particular $E_r(\mathbf{A}) = \text{span}(\mathbf{v})$;
- (c) every eigenvalue $\lambda \in \mathbb{C}$ of $\mathbf{A}$ with $\lambda \neq r$ satisfies $|\lambda| < r$;
- (d) every nonnegative eigenvector of $\mathbf{A}$ is a positive multiple of $\mathbf{v}$. In particular, no eigenvector of $\mathbf{A}$ for an eigenvalue other than r is nonnegative;
- (e) $\mathbf{A}^\top$ is positive with $\rho(\mathbf{A}^\top) = r$, so (a)–(d) hold for $\mathbf{A}^\top$: there is $\mathbf{w} > \mathbf{0}$ with $\mathbf{w}^\top \mathbf{A} = r\mathbf{w}^\top$.

The route. There are several proofs of Perron's theorem. We take the one that finds r as the answer to an optimization problem: **how large can t be if $\mathbf{A}\mathbf{x} \geq t\mathbf{x}$ for some probability vector $\mathbf{x}$?** Its advantage is that it produces the positive eigenvector and the eigenvalue together, and that compactness is used exactly once. The quotient form of this question, maximizing $\min_i(\mathbf{A}\mathbf{x})_i/x_i$ over $\mathbf{x} \in \Delta_n$, is the Collatz–Wielandt formula, which §04 develops for all irreducible matrices.

The quotient itself is awkward on the boundary of the simplex. Where some $x_i = 0$, the i -th ratio is undefined, and the natural repair, taking the minimum over the indices with $x_i > 0$ only, gives a function that need not be continuous. For $\mathbf{D} = \text{diag}(2, 1)$ that function takes the value 2 at the corner $\mathbf{e}_1$, where only the first ratio counts, but the value $\min(2, 1) = 1$ at every point $(1 - t, t)$ with $0 < t < 1$. So the extreme value theorem cannot be applied to it as it stands. (For positive $\mathbf{A}$ the repaired function is in fact continuous, but proving that costs more than the route below; for the nonnegative matrices of §04 it can genuinely fail.) We avoid the quotient: instead of maximizing a function on Δ_n , we maximize the coordinate t over the set of pairs $(\mathbf{x}, t)$ with $\mathbf{A}\mathbf{x} \geq t\mathbf{x}$, which is closed, and on which t is continuous.

Idea. Step roadmap. ① The set $S = \{(\mathbf{x}, t) : \mathbf{x} \in \Delta_n, t \geq 0, \mathbf{A}\mathbf{x} \geq t\mathbf{x}\}$ is compact, so t attains a maximum r on it: facts (A3) and (A4) of Chapter 16's introduction. ② **Push-up claim:** if $\mathbf{u} \geq \mathbf{0}$, $\mathbf{u} \neq \mathbf{0}$, and $\mathbf{A}\mathbf{u} \geq r\mathbf{u}$, then $\mathbf{A}\mathbf{u} = r\mathbf{u}$ and $\mathbf{u} > \mathbf{0}$. Otherwise one more application of the positive matrix $\mathbf{A}$ turns the slack $\mathbf{A}\mathbf{u} - r\mathbf{u}$ into a **strict** inequality, and a strict inequality leaves room to increase r , contradicting maximality. The maximizer therefore gives the positive eigenvector. ③ For any eigenvalue λ with eigenvector $\mathbf{x}$, $|\lambda||\mathbf{x}| \leq \mathbf{A}|\mathbf{x}|$, so $|\lambda| \leq r$ by maximality; hence $r = \rho(\mathbf{A})$. ④ If $|\lambda| = r$, the push-up claim forces equality in the triangle inequality, which aligns all entries of $\mathbf{x}$ and gives $\lambda = r$. ⑤ Subtract a multiple of $\mathbf{v}$ until an entry vanishes: the push-up claim leaves nothing, so $E_r(\mathbf{A})$ is a line. ⑥ Steps ①–⑤ for $\mathbf{A}^\top$ give a positive $\mathbf{w}$ with $\mathbf{A}^\top \mathbf{w} = \rho(\mathbf{A}^\top)\mathbf{w}$, and pairing with $\mathbf{v}$ shows $\rho(\mathbf{A}^\top) = r$. A second vector in the generalized eigenspace would give $(\mathbf{A} - r\mathbf{I})\mathbf{y} = \mathbf{v}$, and pairing with $\mathbf{w}$ yields $0 = \mathbf{w}^\top \mathbf{v} > 0$. ⑦ A nonnegative eigenvector for $\lambda \neq r$ would be orthogonal to $\mathbf{w} > \mathbf{0}$, which is impossible.

Here is the proof, following the roadmap.

Proof. Let $\mathbf{A} > \mathbf{O}$ be $n \times n$. By Definition 18.1.2, Δ_n is the set of probability vectors in $\mathbb{R}^n$. Define

$$S = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{x} \in \Delta_n, t \geq 0, \mathbf{A}\mathbf{x} \geq t\mathbf{x}\}.$$

Step 1: a maximal t . S is nonempty: it contains $(\mathbf{x}, 0)$ for every $\mathbf{x} \in \Delta_n$, by [Lemma 19.1.1](#) (a).

S is bounded for the Euclidean norm on $\mathbb{R}^{n+1}$, the norm with which (A3) is stated. Let M be the largest entry of $\mathbf{A}$. If $(\mathbf{x}, t) \in S$, then $0 \leq x_i \leq 1$ for every i , and since the n entries of $\mathbf{x}$ add up to 1, some $x_k \geq 1/n$. Then

$$\frac{t}{n} \leq t x_k \leq (\mathbf{A}\mathbf{x})_k = \sum_j a_{kj} x_j \leq M \sum_j x_j = M,$$

so $0 \leq t \leq nM$. Hence $\|(\mathbf{x}, t)\|_2^2 = \sum_i x_i^2 + t^2 \leq n + (nM)^2$, and $\|(\mathbf{x}, t)\|_2 \leq \sqrt{n} + nM$.

S is closed in $\mathbb{R}^{n+1}$ (Definition 18.1.6). Let $(\mathbf{x}_k, t_k) \in S$ converge to $(\mathbf{x}, t)$ in the Euclidean norm. Then each coordinate converges, since $|u_i| \leq \|\mathbf{u}\|_2$ for every $\mathbf{u} \in \mathbb{R}^{n+1}$. Every condition defining S is a finite list of nonstrict inequalities and one equation between expressions built from the entries by sums and products: $x_i \geq 0$, $\sum_i x_i = 1$, $t \geq 0$, and $(\mathbf{A}\mathbf{x})_i - t x_i \geq 0$. Each holds for every $(\mathbf{x}_k, t_k)$, and by the algebra of limits of Chapter 16's introduction each survives the limit. So $(\mathbf{x}, t) \in S$.

By fact (A3) of Chapter 16's introduction, the closed bounded set $S \subseteq \mathbb{R}^{n+1}$ is compact, and the function $(\mathbf{x}, t) \mapsto t$ is continuous on it. By fact (A4), the extreme value theorem, it attains a maximum: there are $\mathbf{z} \in \Delta_n$ and a largest number r such that

$$\mathbf{A}\mathbf{z} \geq r\mathbf{z}, \quad \text{and } t \leq r \text{ for every } (\mathbf{x}, t) \in S.$$

Moreover $r > 0$. Indeed, let s be the smallest row sum of $\mathbf{A}$, which is positive as $\mathbf{A} > \mathbf{0}$. The vector $\mathbf{x} = \frac{1}{n}\mathbf{1} \in \Delta_n$ satisfies $(\mathbf{A}\mathbf{x})_i = \frac{1}{n} \sum_j a_{ij} \geq \frac{s}{n} = s x_i$, so $(\frac{1}{n}\mathbf{1}, s) \in S$ and $r \geq s > 0$.

Step 2: the push-up claim.

Claim. If $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \geq \mathbf{0}$, $\mathbf{u} \neq \mathbf{0}$, and $\mathbf{A}\mathbf{u} \geq r\mathbf{u}$, then $\mathbf{A}\mathbf{u} = r\mathbf{u}$ and $\mathbf{u} > \mathbf{0}$.

Proof. Suppose $\mathbf{A}\mathbf{u} \neq r\mathbf{u}$. Then $\mathbf{y} = \mathbf{A}\mathbf{u} - r\mathbf{u}$ satisfies $\mathbf{y} \geq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$, so $\mathbf{A}\mathbf{y} > \mathbf{0}$ by [Lemma 19.1.1](#) (b). Put $\mathbf{q} = \mathbf{A}\mathbf{u}$; by the same rule $\mathbf{q} > \mathbf{0}$, and

$$\mathbf{A}\mathbf{q} - r\mathbf{q} = \mathbf{A}(\mathbf{A}\mathbf{u} - r\mathbf{u}) = \mathbf{A}\mathbf{y} > \mathbf{0}.$$

Every $q_i > 0$, so $\varepsilon = \min_i (\mathbf{A}\mathbf{y})_i / q_i$ is defined and positive, and $(\mathbf{A}\mathbf{q})_i - r q_i = (\mathbf{A}\mathbf{y})_i \geq \varepsilon q_i$ for every i , that is, $\mathbf{A}\mathbf{q} \geq (r + \varepsilon)\mathbf{q}$. Let $\sigma = q_1 + \cdots + q_n > 0$ and $\mathbf{x} = \mathbf{q}/\sigma$. Then $\mathbf{x} \in \Delta_n$, and dividing by $\sigma > 0$ preserves the inequality, so $\mathbf{A}\mathbf{x} \geq (r + \varepsilon)\mathbf{x}$. Hence $(\mathbf{x}, r + \varepsilon) \in S$, which contradicts the maximality of r . Therefore $\mathbf{A}\mathbf{u} = r\mathbf{u}$. Finally $\mathbf{A}\mathbf{u} > \mathbf{0}$ by [Lemma 19.1.1](#) (b), and $r > 0$, so $\mathbf{u} = r^{-1}\mathbf{A}\mathbf{u} > \mathbf{0}$. ■

Apply the claim to $\mathbf{u} = \mathbf{z}$, which is nonnegative and nonzero because its entries add up to 1. It gives $\mathbf{A}\mathbf{z} = r\mathbf{z}$ with $\mathbf{z} > \mathbf{0}$. Put $\mathbf{v} = \mathbf{z}$. So r is an eigenvalue of $\mathbf{A}$ with a positive eigenvector.

Step 3: $r = \rho(\mathbf{A})$. Let $\lambda \in \mathbb{C}$ be an eigenvalue of $\mathbf{A}$ and $\mathbf{x} \in \mathbb{C}^n$ an eigenvector for it. By Theorem 0.6.2 (d) and [Lemma 19.1.2](#) (a), with $|\mathbf{A}| = \mathbf{A}$,

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

The vector $|\mathbf{x}|$ is nonnegative and nonzero, since $\mathbf{x} \neq \mathbf{0}$. Dividing by the sum $\sigma > 0$ of its entries, $\mathbf{x}' = |\mathbf{x}|/\sigma \in \Delta_n$ satisfies $\mathbf{A}\mathbf{x}' \geq |\lambda|\mathbf{x}'$, so $(\mathbf{x}', |\lambda|) \in S$ and $|\lambda| \leq r$. Since r is itself an eigenvalue, $r = \max\{|\lambda| : \lambda \in \text{spec}(\mathbf{A})\} = \rho(\mathbf{A})$ (Definition 16.4.1). This proves (a).

Step 4: eigenvalues of modulus r . Let λ be an eigenvalue with $|\lambda| = r$, and $\mathbf{x}$ an eigenvector. The display of Step 3 reads $\mathbf{A}|\mathbf{x}| \geq r|\mathbf{x}|$, so the push-up claim applies to $|\mathbf{x}|$ and gives

$$\mathbf{A}|\mathbf{x}| = r|\mathbf{x}| \quad \text{and} \quad |\mathbf{x}| > \mathbf{0}.$$

Hence $|\mathbf{A}\mathbf{x}| = |\lambda||\mathbf{x}| = r|\mathbf{x}| = \mathbf{A}|\mathbf{x}|$. Read the first entry: with $z_j = a_{1j}x_j$,

$$\left| \sum_j z_j \right| = |(\mathbf{A}\mathbf{x})_1| = (\mathbf{A}|\mathbf{x}|)_1 = \sum_j a_{1j}|x_j| = \sum_j |z_j|,$$

and $\sum_j z_j = (\mathbf{A}\mathbf{x})_1 = \lambda x_1 \neq 0$, because $|\lambda x_1| = r|x_1| > 0$. By Lemma 19.1.3 there is $c \in \mathbb{C}$ with $|c| = 1$ and $a_{1j}x_j = a_{1j}|x_j|c$ for every j . Since $a_{1j} > 0$, we may divide: $x_j = c|x_j|$, that is, $\mathbf{x} = c|\mathbf{x}|$. Therefore

$$\lambda \mathbf{x} = \mathbf{A}\mathbf{x} = c \mathbf{A}|\mathbf{x}| = c r |\mathbf{x}| = r \mathbf{x},$$

and as $\mathbf{x} \neq \mathbf{0}$, $\lambda = r$. So every eigenvalue other than r has modulus $< r$, by Step 3. This proves (c). We record what the step showed for $\lambda = r$ itself: **every eigenvector $\mathbf{x}$ for r has the form $c|\mathbf{x}|$ with $|c| = 1$, $|\mathbf{x}| > \mathbf{0}$ and $\mathbf{A}|\mathbf{x}| = r|\mathbf{x}|$.**

Step 5: $E_r(\mathbf{A}) = \text{span}(\mathbf{v})$. Let $\mathbf{x}$ be an eigenvector for r , and write $\mathbf{x} = c\mathbf{u}$ with $\mathbf{u} = |\mathbf{x}| > \mathbf{0}$ and $\mathbf{A}\mathbf{u} = r\mathbf{u}$, as just recorded. Let $\theta = \min_i u_i/v_i$, which is positive, and put $\mathbf{y} = \mathbf{u} - \theta\mathbf{v}$. Then $\mathbf{y}$ is real, $\mathbf{y} \geq \mathbf{0}$ by the choice of θ , $y_i = 0$ for an index i attaining the minimum, and $\mathbf{A}\mathbf{y} = r\mathbf{y}$. If $\mathbf{y} \neq \mathbf{0}$, the push-up claim would give $\mathbf{y} > \mathbf{0}$, contradicting $y_i = 0$. So $\mathbf{y} = \mathbf{0}$, and $\mathbf{x} = c\theta\mathbf{v} \in \text{span}(\mathbf{v})$. Hence $E_r(\mathbf{A}) = \text{span}(\mathbf{v})$ and $g_{\mathbf{A}}(r) = 1$.

Step 6: $a_{\mathbf{A}}(r) = 1$. First a positive left eigenvector. $\mathbf{A}^\top$ is positive, since it has the same entries, and Steps 1–5 used nothing about $\mathbf{A}$ but its positivity. Applied to $\mathbf{A}^\top$, they give $\mathbf{w} > \mathbf{0}$ with $\mathbf{A}^\top \mathbf{w} = r'\mathbf{w}$, where $r' = \rho(\mathbf{A}^\top)$. By Theorem 0.9.5, $(\mathbf{A}^\top \mathbf{w})^\top = \mathbf{w}^\top \mathbf{A}$, so

$$r' \mathbf{w}^\top \mathbf{v} = (\mathbf{A}^\top \mathbf{w})^\top \mathbf{v} = \mathbf{w}^\top \mathbf{A} \mathbf{v} = r \mathbf{w}^\top \mathbf{v}.$$

Here $\mathbf{w}^\top \mathbf{v} = \sum_i w_i v_i > 0$, because every term is positive, so $r' = r$. Hence $\rho(\mathbf{A}^\top) = r$ and $\mathbf{w}^\top \mathbf{A} = r \mathbf{w}^\top$.

Now suppose $a_{\mathbf{A}}(r) \geq 2$. The characteristic polynomial splits over $\mathbb{C}$ (Corollary 6.5.5), so Theorem 10.1.3 (c) applies to $\mathbf{A} \in M_n(\mathbb{C})$ and gives $\dim G_r(\mathbf{A}) = a_{\mathbf{A}}(r) \geq 2 > 1 = \dim E_r(\mathbf{A})$. Choose $\mathbf{x} \in G_r(\mathbf{A})$ with $\mathbf{x} \notin E_r(\mathbf{A})$, put $\mathbf{N} = \mathbf{A} - r\mathbf{I}$, and let $k \geq 1$ be least with $\mathbf{N}^k \mathbf{x} = \mathbf{0}$, which exists by Definition 10.1.2 and Theorem 0.2.11. Since $\mathbf{N}\mathbf{x} \neq \mathbf{0}$, $k \geq 2$. Put $\mathbf{y} = \mathbf{N}^{k-2} \mathbf{x}$. Then $\mathbf{N}\mathbf{y} = \mathbf{N}^{k-1} \mathbf{x} \neq \mathbf{0}$ by the minimality of k , and $\mathbf{N}(\mathbf{N}\mathbf{y}) = \mathbf{N}^k \mathbf{x} = \mathbf{0}$. So $\mathbf{N}\mathbf{y}$ is an eigenvector for r , and by Step 5, $\mathbf{N}\mathbf{y} = b\mathbf{v}$ with $b \neq 0$. But $\mathbf{w}^\top \mathbf{N} = \mathbf{w}^\top \mathbf{A} - r \mathbf{w}^\top = \mathbf{0}^\top$, so

$$0 = (\mathbf{w}^\top \mathbf{N})\mathbf{y} = \mathbf{w}^\top (\mathbf{N}\mathbf{y}) = b \mathbf{w}^\top \mathbf{v}.$$

Since $\mathbf{w}^\top \mathbf{v} > 0$ and $b \neq 0$, this is a contradiction. Hence $a_{\mathbf{A}}(r) \leq 1$, and $a_{\mathbf{A}}(r) \geq 1$ because r is an eigenvalue. Together with Step 5 this proves (b).

Step 7: nonnegative eigenvectors. Let $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} \neq \mathbf{0}$, be an eigenvector for an eigenvalue λ . If $\lambda \neq r$, then Theorem 9.10.6, applied over $\mathbb{C}$ to the left eigenvector $\mathbf{w}$ for r and the right eigenvector $\mathbf{x}$ for λ , gives $\mathbf{w}^\top \mathbf{x} = 0$. But $\mathbf{w}^\top \mathbf{x} > 0$ by Lemma 19.1.1 (b), applied to the positive $1 \times n$ matrix $\mathbf{w}^\top$. So $\lambda = r$, and by Step 5, $\mathbf{x} = \theta \mathbf{v}$ for a scalar θ . Comparing any entry, $\theta = x_i/v_i$ is real and ≥ 0 , and it is not 0 because $\mathbf{x} \neq \mathbf{0}$. This proves (d).

Finally, (e): $\mathbf{A}^\top$ is positive, and Step 6 showed $\rho(\mathbf{A}^\top) = r$ by pairing with $\mathbf{v}$, so (a)–(d) hold for $\mathbf{A}^\top$ by what has been proved, with the positive eigenvector $\mathbf{w}$ of Step 6. This completes the proof. ■

Remark. Which analysis was used. Exactly two facts from Chapter 16's introduction, both in Step 1: (A3), that the closed bounded set S is compact, and (A4), that the continuous function t attains its maximum on it, together with the algebra of limits from the same introduction, which showed that S is closed. Everything else is algebra: the order rules, the triangle inequality with its equality case, and the generalized eigenspaces of Chapter 10.

Two examples show what the theorem says.

Example. Perron's theorem for Chapter 9's first example

Check each part of [Theorem 19.1.5](#) for $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$.

Solution. Example (Eigenvalues and Eigenspaces of a 2×2 Matrix) found $p_{\mathbf{A}}(x) = (x - 4)(x + 1)$, $E_4(\mathbf{A}) = \text{span}((2, 3))$ and $E_{-1}(\mathbf{A}) = \text{span}((1, -1))$. So $r = \rho(\mathbf{A}) = 4 > 0$ is an eigenvalue, with the positive eigenvector $\mathbf{v} = (2, 3)$: part (a). It is a simple root of $p_{\mathbf{A}}$: part (b). The other eigenvalue has $|-1| = 1 < 4$: part (c). Every eigenvector for -1 is $t(1, -1)$ with $t \neq 0$, which has entries of opposite signs, so none is nonnegative: part (d). For (e), the columns of $\mathbf{A}$ both sum to 4, so $\mathbf{1}^\top \mathbf{A} = 4 \mathbf{1}^\top$ and $\mathbf{w} = (1, 1)$ is a positive left eigenvector for 4, as the text after Definition 9.10.1 already observed.

Example. A positive 3×3 matrix

Let $\mathbf{A} = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 4 & 2 & 3 \end{pmatrix}$. Find its eigenvalues and eigenvectors, and its positive left eigenvector.

Solution. Try the three vectors $(1, 1, 2)$, $(-1, -1, 3)$ and $(-5, 2, 4)$:

$$\mathbf{A} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 12 \end{pmatrix}, \quad \mathbf{A} \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix}, \quad \mathbf{A} \begin{pmatrix} -5 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ -4 \end{pmatrix}.$$

So 6, 1 and -1 are eigenvalues. They are distinct, and a 3×3 matrix has at most three eigenvalues (Corollary 9.2.3), so these are all, each with algebraic multiplicity 1 because the multiplicities add up to 3 (Corollary 6.5.5 (a)). As checks, $\text{tr } \mathbf{A} = 6 = 6 + 1 - 1$ and $\det \mathbf{A} = -6 = 6 \cdot 1 \cdot (-1)$. Hence $\rho(\mathbf{A}) = 6$, with the positive eigenvector $\mathbf{v} = (1, 1, 2)$, and the other two eigenvalues have modulus $1 < 6$. The eigenvectors for 1 and -1 are the nonzero multiples of $(-1, -1, 3)$ and $(-5, 2, 4)$, and each has entries of both signs, as [Theorem 19.1.5](#)

(d) requires.

For the left eigenvector, solve $(\mathbf{A}^\top - 6\mathbf{I})\mathbf{w} = \mathbf{0}$. Row reduction gives $\mathbf{w} = (10, 11, 7)$, and indeed

$$\mathbf{w}^\top \mathbf{A} = \begin{pmatrix} 10 + 22 + 28 & 30 + 22 + 14 & 10 + 11 + 21 \end{pmatrix} = \begin{pmatrix} 60 & 66 & 42 \end{pmatrix} = 6 \mathbf{w}^\top.$$

The left and right eigenvectors for $\rho(\mathbf{A})$ are unrelated in general; here $\mathbf{w}^\top \mathbf{v} = 10 + 11 + 14 = 35$ for $\mathbf{v} = (1, 1, 2)$. [Definition 19.1.2](#) below scales them so that this pairing is 1.

Quick check. The swap $\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is nonnegative. Which conclusions of [Theorem 19.1.5](#) hold for $\mathbf{S}$, and which fail?

Solution. $\mathbf{S}$ has eigenvalues 1 and -1 , with eigenvectors $(1, 1)$ and $(1, -1)$. So $\rho(\mathbf{S}) = 1$ is an eigenvalue with a positive eigenvector, it is simple, and the only nonnegative eigenvectors are the positive multiples of $(1, 1)$: (a), (b), (d) and (e) hold. Part (c) fails, since $-1 \neq 1$ and $|-1| = 1 = \rho(\mathbf{S})$. The hypothesis that fails is $\mathbf{S} > \mathbf{0}$: two entries are 0. This is the matrix whose powers do not converge in the warning of Chapter 9 §11, and (c) is exactly what was missing there.

The positive eigenvector is so useful that it gets a name.

Definition 19.1.2. Perron Root, Perron Vector

Let $\mathbf{A} \in M_n(\mathbb{R})$ be positive. The number $\rho(\mathbf{A})$ is the **Perron root** of $\mathbf{A}$. The **Perron vector** of $\mathbf{A}$ is the unique $\mathbf{v} > \mathbf{0}$ with $\mathbf{A}\mathbf{v} = \rho(\mathbf{A})\mathbf{v}$ and $v_1 + \cdots + v_n = 1$. The **left Perron vector** of $\mathbf{A}$ is the unique $\mathbf{w} > \mathbf{0}$ with $\mathbf{w}^\top \mathbf{A} = \rho(\mathbf{A})\mathbf{w}^\top$ and $\mathbf{w}^\top \mathbf{v} = 1$, where $\mathbf{v}$ is the Perron vector.

Existence and uniqueness are [Theorem 19.1.5](#) (a) and (b): the eigenspace for $\rho(\mathbf{A})$ is a line through a positive vector, and it meets the hyperplane $\{\mathbf{1}^\top \mathbf{x} = 1\}$ in exactly one point, the vector $\mathbf{v}/(\mathbf{1}^\top \mathbf{v})$. By (d), the Perron vector is also the **only** eigenvector of $\mathbf{A}$ in Δ_n . Likewise, by (e) and (b) applied to $\mathbf{A}^\top$, the positive left eigenvectors for $\rho(\mathbf{A})$ are the positive multiples of one $\mathbf{w}_0 > \mathbf{0}$, and since $\mathbf{w}_0^\top \mathbf{v} > 0$, exactly one of them, $\mathbf{w}_0/(\mathbf{w}_0^\top \mathbf{v})$, pairs with the Perron vector to 1. The two normalizations differ on purpose: $\mathbf{v}$ is a probability vector, and $\mathbf{w}^\top \mathbf{v} = 1$ makes $\mathbf{v}\mathbf{w}^\top$ a projection, the limit of $(\mathbf{A}/\rho(\mathbf{A}))^k$ that §05 finds. For [Example \(Perron's theorem for Chapter 9's first example\)](#) the Perron vector is $\frac{1}{5}(2, 3)$ and the left Perron vector is $(1, 1)$, since $\mathbf{1}^\top \mathbf{v} = 1$. For [Example \(A positive \$3 \times 3\$ matrix\)](#) the Perron vector is $\frac{1}{4}(1, 1, 2)$, and the left Perron vector is $\frac{4}{35}(10, 11, 7)$, since $(10, 11, 7)$ pairs with $\frac{1}{4}(1, 1, 2)$ to $\frac{35}{4}$.

Warning. **Perron's theorem says nothing about the other eigenvalues except their size.** They may be negative, as -1 is in [Example \(Perron's theorem for Chapter 9's first example\)](#), or nonreal: the positive matrix $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$ has row sums 6, hence the eigenvector $\mathbf{1}$ for 6, and its other two eigenvalues are the roots of $x^2 + 3x + 3$, namely $\frac{-3 \pm i\sqrt{3}}{2}$, of modulus

$\sqrt{3} < 6$. They may also be repeated, as the eigenvalue 1 of $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ is. Simplicity is claimed for $\rho(\mathbf{A})$ alone.

19.1.4 What positivity buys, and what it costs

Each hypothesis of [Theorem 19.1.5](#) has a job, and it is instructive to see where it went. **Positivity** of $\mathbf{A}$ was used in the push-up claim, to turn a nonnegative, nonzero slack into a strictly positive one; in Step 4, to divide by a_{1j} ; and in Step 1, to make $r > 0$. For a matrix that is only nonnegative, each of these uses fails, and so can each conclusion: Exercise C2 gives a nonnegative 2×2 matrix for each of (a)–(d), and the swap of the Quick check is the one for (c).

§03 shows that one conclusion survives in general, that $\rho(\mathbf{A})$ is an eigenvalue with a nonnegative eigenvector, and that most of the rest returns under a hypothesis much weaker than positivity. That hypothesis, **irreducibility**, is a condition on where the zero entries are, and §02 introduces it.

Here is the payoff for Chapter 9. A stochastic matrix has $\mathbf{1}^\top \mathbf{A} = \mathbf{1}^\top$, so $\mathbf{1}$ is a positive left eigenvector for 1. If $\mathbf{A}$ is stochastic and positive, [Theorem 19.1.5](#) (d) applied to $\mathbf{A}^\top$ says that $\mathbf{1}$ is a left eigenvector for $\rho(\mathbf{A})$ and for no other eigenvalue, so $\rho(\mathbf{A}) = 1$. Then (b) and (c) say that 1 is simple and every other eigenvalue has modulus less than 1: two of the three hypotheses of Theorem 9.11.4 hold automatically. The third, diagonalizability, is the one §06 removes.

19.1.5 Exercises

19.1.5.1 A. Check your understanding

Exercise. A1

- Define what it means for a real matrix to be **nonnegative** and to be **positive**, and define $|\mathbf{M}|$ for a complex matrix $\mathbf{M}$.
- State [Theorem 19.1.5](#).
- True or false: a positive symmetric matrix is positive definite. Justify your answer.
- True or false: a positive matrix has no negative eigenvalue. Justify your answer.
- True or false: if $\mathbf{O} \leq \mathbf{A} \leq \mathbf{B}$ and $\mathbf{A} \neq \mathbf{B}$, then $\rho(\mathbf{A}) < \rho(\mathbf{B})$. Justify your answer.
- In the proof of [Theorem 19.1.5](#), which step uses compactness, and why is the quotient $\min_i (\mathbf{A}\mathbf{x})_i / x_i$ not maximized directly?

Solution.

- $\mathbf{A}$ is nonnegative if every entry is ≥ 0 , and positive if every entry is > 0 ([Definition 19.1.1](#)). $|\mathbf{M}|$ is the real matrix whose (i, j) entry is $|m_{ij}|$.
- If $\mathbf{A} \in M_n(\mathbb{R})$ and $\mathbf{A} > \mathbf{O}$, then $r = \rho(\mathbf{A}) > 0$ is an eigenvalue of $\mathbf{A}$ with a positive eigenvector $\mathbf{v}$; it is algebraically simple; every other eigenvalue has modulus less than r ; every nonnegative eigenvector is a positive multiple of $\mathbf{v}$; and the same holds for $\mathbf{A}^\top$, which gives a positive left eigenvector.
- False. $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is positive and symmetric with the eigenvalue -1 ; the warning after [Definition 19.1.1](#) computes $\mathbf{x}^\top \mathbf{A}\mathbf{x} = -2$ for $\mathbf{x} = (1, -1)$.

- (d) False. $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ has the eigenvalue -1 (Example (Perron's theorem for Chapter 9's first example)). Perron's theorem controls only the modulus of the other eigenvalues.
- (e) False. $\text{diag}(1, 0) \leq \mathbf{I}_2$, the two differ, and both have spectral radius 1.
- (f) Step 1: the set S is closed and bounded, so by (A3) and (A4) the coordinate t attains a maximum on it. The quotient is undefined where some $x_i = 0$. For a nonnegative matrix its natural repair can be discontinuous on the boundary of Δ_n (for $\text{diag}(2, 1)$ it jumps at $\mathbf{e}_1$), so the extreme value theorem does not apply to it directly. For positive $\mathbf{A}$ the repair is in fact continuous, but proving that costs more than maximizing t over S .

19.1.5.2 B. Practice

Exercise. B1: A Perron vector by hand

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 4 & 3 \\ 3 & 1 & 4 \end{pmatrix}$. Find the eigenvalues of $\mathbf{A}$, its Perron root, Perron vector, and left Perron vector. Check that no eigenvector for the other eigenvalues is nonnegative.

Solution. Expanding $\det(x\mathbf{I} - \mathbf{A})$ along the first row,

$$\begin{aligned} p_{\mathbf{A}}(x) &= (x - 2)((x - 4)^2 - 3) + (-3(x - 4) - 9) - (3 + 3(x - 4)) \\ &= x^3 - 10x^2 + 23x - 14 = (x - 1)(x - 2)(x - 7). \end{aligned}$$

(Check: $\text{tr } \mathbf{A} = 10$ and $\det \mathbf{A} = 14$.) So the eigenvalues are 7, 2, 1, and $\rho(\mathbf{A}) = 7$. Solving $(\mathbf{A} - 7\mathbf{I})\mathbf{x} = \mathbf{0}$, whose first two rows are $-5x_1 + x_2 + x_3 = 0$ and $3x_1 - 3x_2 + 3x_3 = 0$, gives $\mathbf{x} = t(1, 3, 2)$; indeed $\mathbf{A}(1, 3, 2) = (7, 21, 14)$. The entries of $(1, 3, 2)$ add up to 6, so the Perron vector is $\frac{1}{6}(1, 3, 2)$.

For the left eigenvector, $(\mathbf{A}^T - 7\mathbf{I})\mathbf{w} = \mathbf{0}$ gives the multiples of $(3, 2, 3)$, and $(3, 2, 3)\mathbf{A} = (21, 14, 21) = 7(3, 2, 3)$. Against the Perron vector, $(3, 2, 3) \cdot \frac{1}{6}(1, 3, 2) = \frac{15}{6} = \frac{5}{2}$, so by Definition 19.1.2 the left Perron vector is $\frac{2}{5}(3, 2, 3) = (\frac{6}{5}, \frac{4}{5}, \frac{6}{5})$.

For 2: $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$ gives $\mathbf{x} = t(1, 3, -3)$, since $\mathbf{A}(1, 3, -3) = (2, 6, -6)$. For 1: $\mathbf{x} = t(1, 0, -1)$, since $\mathbf{A}(1, 0, -1) = (1, 0, -1)$. For $t \neq 0$ each has entries of both signs, so neither is nonnegative, as Theorem 19.1.5 (d) predicts.

Exercise. B2: Which order?

For each matrix, decide whether it is nonnegative, whether it is positive, and whether it is positive definite. Justify your answers, and say whether Theorem 19.1.5 applies.

- (a) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$
- (b) $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$

- (c) $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$
 (d) $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$
 (e) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

Solution. For a real symmetric 2×2 matrix, positive definite means both eigenvalues are positive (Definition 13.1.1, with the spectral theorem).

- (a) Positive, hence nonnegative. Eigenvalues 3, 1, both positive, so positive definite. Perron applies: $\rho = 3$ with eigenvector $(1, 1)$.
 (b) Positive, hence nonnegative. Eigenvalues 4, -2 , so **not** positive definite. Perron applies: $\rho = 4$ with eigenvector $(1, 1)$, and $|-2| < 4$.
 (c) Not nonnegative, since the off-diagonal entries are -1 ; so not positive, and Perron does not apply. It is positive definite, with eigenvalues 1, 3.
 (d) Nonnegative, not positive (the $(1, 2)$ entry is 0), and not symmetric, so not positive definite by clause (P1). Perron does not apply, although here $\rho = 3$ is an eigenvalue with the eigenvector $\mathbf{e}_2 \geq \mathbf{0}$, which is not positive.
 (e) Positive. Eigenvalues 2 and 0, so positive semidefinite but **not** positive definite, as $\mathbf{x} = (1, -1)$ gives $\mathbf{x}^\top \mathbf{A} \mathbf{x} = 0$. Perron applies: $\rho = 2$, eigenvector $(1, 1)$, and the eigenvalue 0 has modulus less than 2.

Exercise. B3: Bounds from the smallest and largest entry

- (a) Show that $\rho(c\mathbf{J}) = nc$ for every real $c \geq 0$, where $\mathbf{J}$ is the $n \times n$ all-ones matrix.
 (b) Let $\mathbf{A} \in M_n(\mathbb{R})$ with $a \leq a_{ij} \leq b$ for all i, j , where $0 \leq a \leq b$. Prove that $na \leq \rho(\mathbf{A}) \leq nb$.
 (c) Hence bound $\rho(\mathbf{A})$ for $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$, and compare with its exact value.

Solution.

- (a) $\mathbf{J}\mathbf{1} = n\mathbf{1}$, so n is an eigenvalue of $\mathbf{J}$. Also $\mathbf{J}^2 = n\mathbf{J}$, since every entry of $\mathbf{J}^2$ is a sum of n ones. If $\mathbf{J}\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$, then $\lambda^2\mathbf{x} = \mathbf{J}^2\mathbf{x} = n\mathbf{J}\mathbf{x} = n\lambda\mathbf{x}$, so $\lambda^2 = n\lambda$ and $\lambda \in \{0, n\}$. Hence $\rho(\mathbf{J}) = n$, and $\rho(c\mathbf{J}) = c\rho(\mathbf{J}) = nc$ because the eigenvalues of $c\mathbf{J}$ are c times those of $\mathbf{J}$.
 (b) The hypothesis says $\mathbf{0} \leq a\mathbf{J} \leq \mathbf{A} \leq b\mathbf{J}$. By Proposition 19.1.4 (a), applied twice, $\rho(a\mathbf{J}) \leq \rho(\mathbf{A}) \leq \rho(b\mathbf{J})$, and by (a) this reads $na \leq \rho(\mathbf{A}) \leq nb$.
 (c) The entries lie between 1 and 3 and $n = 3$, so $3 \leq \rho(\mathbf{A}) \leq 9$. In fact every row sums to 6, so $\mathbf{A}\mathbf{1} = 6\mathbf{1}$, and Theorem 19.1.5 (d) says the positive eigenvector $\mathbf{1}$ belongs to $\rho(\mathbf{A})$. Hence $\rho(\mathbf{A}) = 6$.

19.1.5.3 C. Going deeper

Exercise. C1: Symmetry passes to the Perron vector

Let $\mathbf{A} > \mathbf{O}$ be $n \times n$ with Perron vector $\mathbf{v}$, and let $\mathbf{P}$ be a permutation matrix with $\mathbf{P}^\top \mathbf{A} \mathbf{P} = \mathbf{A}$.

(a) Prove that $\mathbf{P}\mathbf{v} = \mathbf{v}$.

(b) Deduce that the Perron vector of $\mathbf{A} = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix}$ has equal first and second entries, and find it.

(c) Deduce that if $\mathbf{A} > \mathbf{O}$ commutes with the cyclic permutation matrix of $1 \rightarrow 2 \rightarrow \cdots \rightarrow n \rightarrow 1$, then its Perron vector is $\frac{1}{n}\mathbf{1}$.

Hint: for (a), $\mathbf{P}\mathbf{v}$ is a positive vector in Δ_n .

Solution.

(a) By Lemma 3.4.7 (b), $\mathbf{P}\mathbf{P}^\top = \mathbf{I}$, so $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{P}^\top \mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{A}$. Hence $\mathbf{A}(\mathbf{P}\mathbf{v}) = \mathbf{P}\mathbf{A}\mathbf{v} = \rho(\mathbf{A})\mathbf{P}\mathbf{v}$. The entries of $\mathbf{P}\mathbf{v}$ are those of $\mathbf{v}$ in another order ($\mathbf{P} = \mathbf{P}_\sigma$ sends $\mathbf{e}_j$ to $\mathbf{e}_{\sigma(j)}$), so $\mathbf{P}\mathbf{v} > \mathbf{0}$ and its entries add up to 1. By Definition 19.1.2 and the uniqueness proved after it, $\mathbf{P}\mathbf{v} = \mathbf{v}$.

(b) Let $\mathbf{P}$ be the permutation matrix that swaps $\mathbf{e}_1$ and $\mathbf{e}_2$. Then $\mathbf{P}^\top \mathbf{A} \mathbf{P}$ is $\mathbf{A}$ with rows 1, 2 swapped and columns 1, 2 swapped, which is $\mathbf{A}$ again, because rows 1 and 2 of $\mathbf{A}$ are equal and so are columns 1 and 2. By (a), $\mathbf{P}\mathbf{v} = \mathbf{v}$, that is, $v_1 = v_2$. Now try $\mathbf{v} = (1, 1, c)$: $\mathbf{A}\mathbf{v} = (2 + 3c, 2 + 3c, 2 + c)$, and $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ requires $\lambda = 2 + 3c$ and $2 + c = (2 + 3c)c$, that is, $3c^2 + c - 2 = 0$, so $c = \frac{2}{3}$ or $c = -1$. The positive choice $c = \frac{2}{3}$ gives the eigenvector $(3, 3, 2)$ with eigenvalue 4, and by Theorem 19.1.5 (d) this is the Perron eigenvector, so $\rho(\mathbf{A}) = 4$ and the Perron vector is $\frac{1}{8}(3, 3, 2)$.

(c) Let $\mathbf{C}$ be the cyclic permutation matrix, $\mathbf{C}\mathbf{e}_j = \mathbf{e}_{j+1}$ for $j < n$ and $\mathbf{C}\mathbf{e}_n = \mathbf{e}_1$. If $\mathbf{A}\mathbf{C} = \mathbf{C}\mathbf{A}$, then $\mathbf{C}^\top \mathbf{A} \mathbf{C} = \mathbf{C}^\top \mathbf{C} \mathbf{A} = \mathbf{A}$ by Lemma 3.4.7 (b). By (a), $\mathbf{C}\mathbf{v} = \mathbf{v}$, which says $v_n = v_1$ and $v_j = v_{j+1}$ for $j < n$, so all entries are equal. As they add up to 1, $\mathbf{v} = \frac{1}{n}\mathbf{1}$.

Exercise. C2: Every conclusion fails for some nonnegative matrix

For each part of Theorem 19.1.5, (a) to (d), give a nonnegative 2×2 matrix for which that part fails, and name the step of the proof that breaks. Then give a nonnegative matrix for which (a), (b) and (c) fail at the same time.

Solution.

(a) $\mathbf{J}_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has $\rho = 0$, and its eigenvectors are the nonzero multiples of $\mathbf{e}_1$, none positive. In the proof, $r > 0$ came from the smallest row sum, which here is 0; and the push-up claim fails, since $\mathbf{u} = \mathbf{e}_1$ satisfies $\mathbf{A}\mathbf{u} = 0 \cdot \mathbf{u}$ without being positive.

(b) $\mathbf{I}_2$: $\rho = 1$ is a double eigenvalue and $E_1 = \mathbb{R}^2$. Step 5 fails: $\mathbf{u} - \theta\mathbf{v}$ can be nonzero and

nonnegative with a zero entry, since $\mathbf{A}\mathbf{y} > \mathbf{0}$ is no longer forced.

- (c) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$: the eigenvalue -1 has modulus $\rho = 1$. Step 4 fails: equality in the triangle inequality (Lemma 19.1.3) is automatic when each row has a single nonzero term, so it aligns nothing.
- (d) $\text{diag}(2, 1)$: the eigenvector $\mathbf{e}_2 \geq \mathbf{0}$ belongs to $1 \neq 2 = \rho$. Step 7 fails, because the left eigenvector for 2 is $\mathbf{e}_1$, which is not positive, and $\mathbf{e}_1^\top \mathbf{e}_2 = 0$ is no contradiction.

For all three at once, take

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

It is block upper triangular with diagonal blocks $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and (1) , so $p_{\mathbf{A}}(x) = (x^2 - 1)(x - 1) = (x - 1)^2(x + 1)$ by Theorem 7.8.4, applied to $x\mathbf{I} - \mathbf{A}$ over $F[x]$, as the paragraph after it notes. Hence $\rho(\mathbf{A}) = 1$ and $a_{\mathbf{A}}(1) = 2$, so (b) fails, and $|-1| = \rho(\mathbf{A})$, so (c) fails. The system $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$ reads $-x_1 + x_2 + x_3 = 0$ and $x_1 - x_2 = 0$, which forces $x_3 = 0$; so every eigenvector for 1 is a multiple of $(1, 1, 0)$, none of them positive, and (a) fails.

Exercise. C3: The Perron root is strictly monotone

Let $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{R})$ with $\mathbf{0} < \mathbf{A} \leq \mathbf{B}$ and $\mathbf{A} \neq \mathbf{B}$.

- (a) Prove that $\rho(\mathbf{A}) < \rho(\mathbf{B})$.
- (b) Show by an example that the conclusion fails if $\mathbf{A} > \mathbf{0}$ is weakened to $\mathbf{A} \geq \mathbf{0}$.

Hint: pair $\mathbf{B}\mathbf{v}$ with a positive left eigenvector of $\mathbf{B}$, where $\mathbf{v}$ is the Perron vector of $\mathbf{A}$.

Solution.

- (a) $\mathbf{B} \geq \mathbf{A} > \mathbf{0}$, so $\mathbf{B}$ is positive too. Let $\mathbf{v} > \mathbf{0}$ be the Perron vector of $\mathbf{A}$, and by Theorem 19.1.5 (e) let $\mathbf{w} > \mathbf{0}$ satisfy $\mathbf{w}^\top \mathbf{B} = \rho(\mathbf{B})\mathbf{w}^\top$. Then

$$\rho(\mathbf{B})\mathbf{w}^\top \mathbf{v} = \mathbf{w}^\top \mathbf{B}\mathbf{v} = \mathbf{w}^\top \mathbf{A}\mathbf{v} + \mathbf{w}^\top (\mathbf{B} - \mathbf{A})\mathbf{v} = \rho(\mathbf{A})\mathbf{w}^\top \mathbf{v} + \mathbf{w}^\top (\mathbf{B} - \mathbf{A})\mathbf{v}.$$

The matrix $\mathbf{B} - \mathbf{A}$ is nonnegative and nonzero, say $b_{kl} - a_{kl} > 0$, so $\mathbf{w}^\top (\mathbf{B} - \mathbf{A})\mathbf{v} \geq w_k(b_{kl} - a_{kl})v_l > 0$, every other term being ≥ 0 . Since $\mathbf{w}^\top \mathbf{v} > 0$, dividing gives $\rho(\mathbf{B}) > \rho(\mathbf{A})$.

- (b) $\mathbf{A} = \text{diag}(1, 0)$ and $\mathbf{B} = \mathbf{I}_2$ satisfy $\mathbf{0} \leq \mathbf{A} \leq \mathbf{B}$ and $\mathbf{A} \neq \mathbf{B}$, with $\rho(\mathbf{A}) = \rho(\mathbf{B}) = 1$. In the argument of (a), the Perron vector of $\mathbf{A}$ is replaced by $\mathbf{e}_1$, and $(\mathbf{B} - \mathbf{A})\mathbf{e}_1 = \mathbf{0}$: the extra mass sits where the eigenvector does not look.