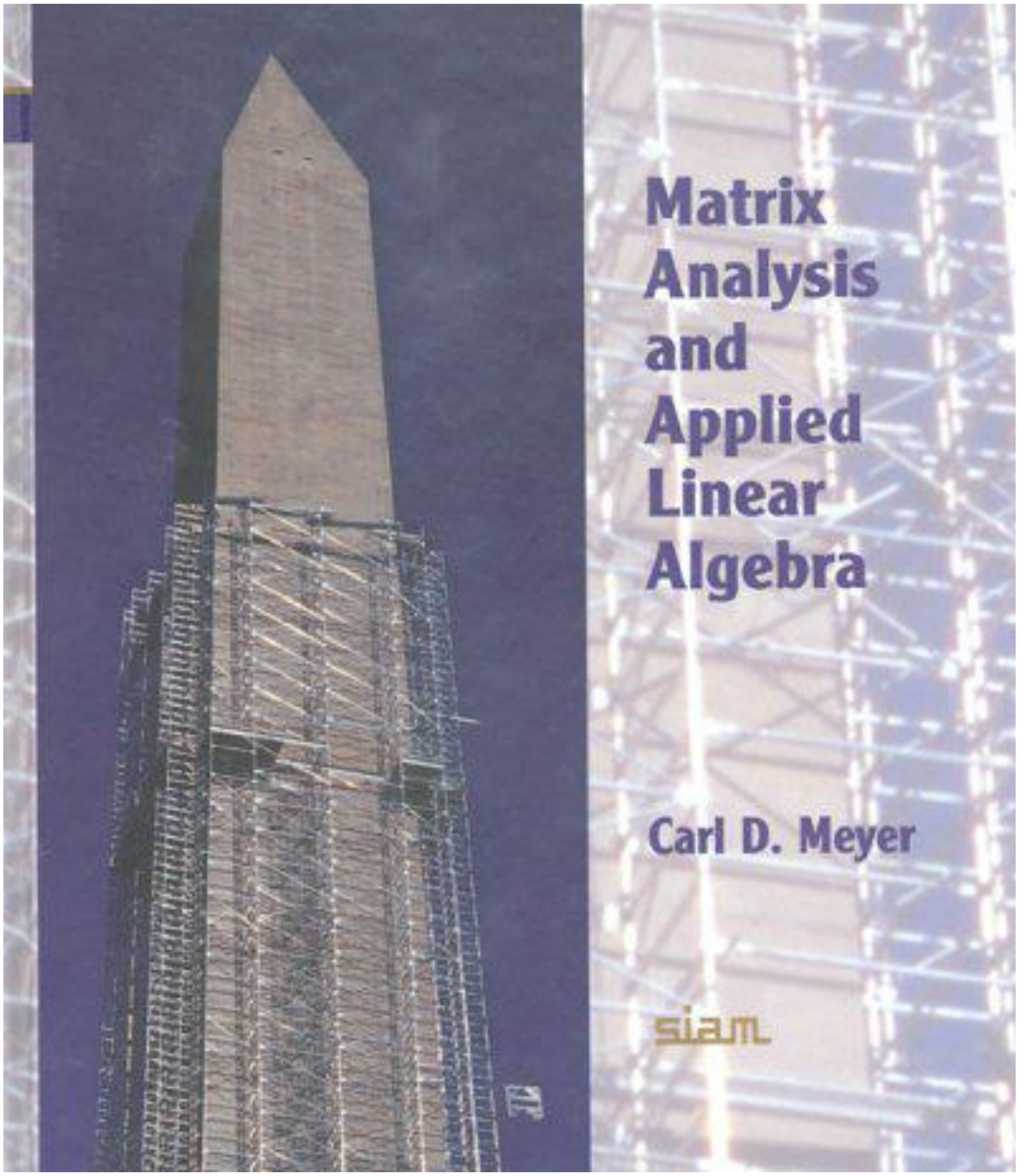
Matrix
Analysis
and
Applied
Linear
Algebra
Carl D. Meyer
siam

CHAPTER 8

Perron–Frobenius Theory of Nonnegative Matrices

8.1 INTRODUCTION

$\mathbf{A} \in \Re^{m\times n}$ is said to be a ***nonnegative matrix*** whenever each $a_{ij} \geq 0$, and this is denoted by writing $\mathbf{A} \geq \mathbf{0}$. In general, $\mathbf{A} \geq \mathbf{B}$ means that each $a_{ij} \geq b_{ij}$. Similarly, $\mathbf{A}$ is a ***positive matrix*** when each $a_{ij} > 0$, and this is denoted by writing $\mathbf{A} > \mathbf{0}$. More generally, $\mathbf{A} > \mathbf{B}$ means that each $a_{ij} > b_{ij}$.

Applications abound with nonnegative and positive matrices. In fact, many of the applications considered in this text involve nonnegative matrices. For example, the connectivity matrix $\mathbf{C}$ in Example 3.5.2 (p. 100) is nonnegative. The discrete Laplacian $\mathbf{L}$ from Example 7.6.2 (p. 563) leads to a nonnegative matrix because $(4\mathbf{I} - \mathbf{L}) \geq \mathbf{0}$. The matrix $e^{\mathbf{A}t}$ that defines the solution of the system of differential equations in the mixing problem of Example 7.9.7 (p. 610) is nonnegative for all $t \geq 0$. And the system of difference equations $\mathbf{p}(k) = \mathbf{A}\mathbf{p}(k-1)$ resulting from the shell game of Example 7.10.8 (p. 635) has a nonnegative coefficient matrix $\mathbf{A}$.

Since nonnegative matrices are pervasive, it's natural to investigate their properties, and that's the purpose of this chapter. A primary issue concerns the extent to which the properties $\mathbf{A} > \mathbf{0}$ or $\mathbf{A} \geq \mathbf{0}$ translate to spectral properties—e.g., to what extent does $\mathbf{A}$ have positive (or nonnegative) eigenvalues and eigenvectors?

The topic is called the "Perron–Frobenius theory" because it evolved from the contributions of the German mathematicians Oskar (or Oscar) Perron[89] and

[89] Oskar Perron (1880–1975) originally set out to fulfill his father's wishes to be in the family busi-

Ferdinand Georg Frobenius.[90] Perron published his treatment of positive matrices in 1907, and in 1912 Frobenius contributed substantial extensions of Perron's results to cover the case of nonnegative matrices.

In addition to saying something useful, the Perron–Frobenius theory is elegant. It is a testament to the fact that beautiful mathematics eventually tends to be useful, and useful mathematics eventually tends to be beautiful.

---

ness, so he only studied mathematics in his spare time. But he was eventually captured by the subject, and, after studying at Berlin, Tübingen, and Göttingen, he completed his doctorate, writing on geometry, at the University of Munich under the direction of Carl von Lindemann (1852–1939) (who first proved that π was transcendental). Upon graduation in 1906, Perron held positions at Munich, Tübingen, and Heidelberg. Perron's career was interrupted in 1915 by World War I in which he earned the Iron Cross. After the war he resumed work at Heidelberg, but in 1922 he returned to Munich to accept a chair in mathematics, a position he occupied for the rest of his career. In addition to his contributions to matrix theory, Perron's work covered a wide range of other topics in algebra, analysis, differential equations, continued fractions, geometry, and number theory. He was a man of extraordinary mental and physical energy. In addition to being able to climb mountains until he was in his midseventies, Perron continued to teach at Munich until he was 80 (although he formally retired at age 71), and he maintained a remarkably energetic research program into his nineties. He published 18 of his 218 papers *after* he was 84.

[90] Ferdinand Georg Frobenius (1849–1917) earned his doctorate under the supervision of Karl Weierstrass (p. 589) at the University of Berlin in 1870. As mentioned earlier, Frobenius was a mentor to and a collaborator with Issai Schur (p. 123), and, in addition to their joint work in group theory, they were among the first to study matrix theory as a discipline unto itself. Frobenius in particular must be considered along with Cayley and Sylvester when thinking of core developers of matrix theory. However, in the beginning, Frobenius's motivation came from Kronecker (p. 597) and Weierstrass, and he seemed oblivious to Cayley's work (p. 80). It was not until 1896 that Frobenius became aware of Cayley's 1857 work, *A Memoir on the Theory of Matrices,* and only then did the terminology "matrix" appear in Frobenius's work. Even though Frobenius was the first to give a rigorous proof of the Cayley–Hamilton theorem (p. 509), he generously attributed it to Cayley in spite of the fact that Cayley had only discussed the result for 2×2 and 3×3 matrices. But credit in this regard is not overly missed because Frobenius's extension of Perron's results are more substantial, and they alone may keep Frobenius's name alive forever.

8.2 POSITIVE MATRICES

The purpose of this section is to focus on matrices $\mathbf{A}_{n\times n} > \mathbf{0}$ with positive entries, and the aim is to investigate the extent to which this positivity is inherited by the eigenvalues and eigenvectors of $\mathbf{A}$.

There are a few elementary observations that will help along the way, so let's begin with them. First, notice that

$$\mathbf{A} > \mathbf{0} \implies \rho(\mathbf{A}) > 0 \tag{8.2.1}$$

because if $\sigma(\mathbf{A}) = \{0\}$, then the Jordan form for $\mathbf{A}$, and hence $\mathbf{A}$ itself, is nilpotent, which is impossible when each $a_{ij} > 0$. This means that our discussions can be limited to positive matrices having spectral radius 1 because $\mathbf{A}$ can always be normalized by its spectral radius—i.e., $\mathbf{A} > \mathbf{0} \Longleftrightarrow \mathbf{A}/\rho(\mathbf{A}) > \mathbf{0}$, and $\rho(\mathbf{A}) = r \Longleftrightarrow \rho(\mathbf{A}/r) = 1$. Other easily verified observations are

$$\begin{aligned}
&\mathbf{P} > \mathbf{0},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{x} \neq \mathbf{0} &&\implies \mathbf{P}\mathbf{x} > \mathbf{0}, && (8.2.2)\\
&\mathbf{N} \geq \mathbf{0},\ \mathbf{u} \geq \mathbf{v} \geq \mathbf{0} &&\implies \mathbf{N}\mathbf{u} \geq \mathbf{N}\mathbf{v}, && (8.2.3)\\
&\mathbf{N} \geq \mathbf{0},\ \mathbf{z} > \mathbf{0},\ \mathbf{N}\mathbf{z} = \mathbf{0} &&\implies \mathbf{N} = \mathbf{0}, && (8.2.4)\\
&\mathbf{N} \geq \mathbf{0},\ \mathbf{N} \neq \mathbf{0},\ \mathbf{u} > \mathbf{v} > \mathbf{0} &&\implies \mathbf{N}\mathbf{u} > \mathbf{N}\mathbf{v}. && (8.2.5)
\end{aligned}$$

In all that follows, the bar notation $|\star|$ is used to denote a matrix of absolute values—i.e., $|\mathbf{M}|$ is the matrix having entries $|m_{ij}|$. The bar notation will *never* denote a determinant in the sequel. Finally, notice that as a simple consequence of the triangle inequality, it's always true that $|\mathbf{A}\mathbf{x}| \leq |\mathbf{A}|\,|\mathbf{x}|$.

Positive Eigenpair

If $\mathbf{A}_{n\times n} > \mathbf{0}$, then the following statements are true.

- $\rho(\mathbf{A}) \in \sigma(\mathbf{A})$. (8.2.6)
- If $\mathbf{A}\mathbf{x} = \rho(\mathbf{A})\,\mathbf{x}$, then $\mathbf{A}|\mathbf{x}| = \rho(\mathbf{A})\,|\mathbf{x}|$ and $|\mathbf{x}| > \mathbf{0}$. (8.2.7)

In other words, $\mathbf{A}$ has an eigenpair of the form $(\rho(\mathbf{A}), \mathbf{v})$ with $\mathbf{v} > 0$.

Proof. As mentioned earlier, it can be assumed that $\rho(\mathbf{A}) = 1$ without any loss of generality. If $(\lambda, \mathbf{x})$ is any eigenpair for $\mathbf{A}$ such that $|\lambda| = 1$, then

$$|\mathbf{x}| = |\lambda|\,|\mathbf{x}| = |\lambda\mathbf{x}| = |\mathbf{A}\mathbf{x}| \leq |\mathbf{A}|\,|\mathbf{x}| = \mathbf{A}\,|\mathbf{x}| \implies |\mathbf{x}| \leq \mathbf{A}\,|\mathbf{x}|. \tag{8.2.8}$$

The goal is to show that equality holds. For convenience, let $\mathbf{z} = \mathbf{A}\,|\mathbf{x}|$ and $\mathbf{y} = \mathbf{z} - |\mathbf{x}|$, and notice that (8.2.8) implies $\mathbf{y} \geq \mathbf{0}$. Suppose that $\mathbf{y} \neq \mathbf{0}$—i.e., suppose that some $y_i > 0$. In this case, it follows from (8.2.2) that $\mathbf{Ay} > \mathbf{0}$ and $\mathbf{z} > \mathbf{0}$, so there must exist a number $\epsilon > 0$ such that $\mathbf{Ay} > \epsilon\,\mathbf{z}$ or, equivalently,

$$\frac{\mathbf{A}}{1+\epsilon}\mathbf{z} > \mathbf{z}.$$

Writing this inequality as $\mathbf{Bz} > \mathbf{z}$, where $\mathbf{B} = \mathbf{A}/(1+\epsilon)$, and successively multiplying both sides by $\mathbf{B}$ while using (8.2.5) produces

$$\mathbf{B}^2\mathbf{z} > \mathbf{Bz} > \mathbf{z}, \quad \mathbf{B}^3\mathbf{z} > \mathbf{B}^2\mathbf{z} > \mathbf{z}, \quad \ldots \quad \Longrightarrow \quad \mathbf{B}^k\mathbf{z} > \mathbf{z} \quad \text{for all } k = 1, 2, \ldots.$$

But $\lim_{k\to\infty} \mathbf{B}^k = \mathbf{0}$ because $\rho(\mathbf{B}) = \sigma\big(\mathbf{A}/(1+\epsilon)\big) = 1/(1+\epsilon) < 1$ (recall (7.10.5) on p. 617), so, in the limit, we have $\mathbf{0} > \mathbf{z}$, which contradicts the fact that $\mathbf{z} > \mathbf{0}$. Since the supposition that $\mathbf{y} \neq \mathbf{0}$ led to this contradiction, the supposition must be false and, consequently, $\mathbf{0} = \mathbf{y} = \mathbf{A}\,|\mathbf{x}| - |\mathbf{x}|$. Thus $|\mathbf{x}|$ is an eigenvector for $\mathbf{A}$ associated with the eigenvalue $1 = \rho(\mathbf{A})$. The proof is completed by observing that $|\mathbf{x}| = \mathbf{A}\,|\mathbf{x}| = \mathbf{z} > \mathbf{0}$. ■

Now that it's been established that $\rho(\mathbf{A}) > 0$ is in fact an eigenvalue for $\mathbf{A} > \mathbf{0}$, the next step is to investigate the index of this special eigenvalue.

Index of $\rho(\mathbf{A})$

If $\mathbf{A}_{n\times n} > \mathbf{0}$, then the following statements are true.

- $\rho(\mathbf{A})$ is the only eigenvalue of $\mathbf{A}$ on the spectral circle.
- *index*$(\rho(\mathbf{A})) = 1$. In other words, $\rho(\mathbf{A})$ is a *semisimple* eigenvalue. Recall Exercise 7.8.4 (p. 596).

Proof. Again, assume without loss of generality that $\rho(\mathbf{A}) = 1$. We know from (8.2.7) on p. 663 that if $(\lambda, \mathbf{x})$ is an eigenpair for $\mathbf{A}$ such that $|\lambda| = 1$, then $\mathbf{0} < |\mathbf{x}| = \mathbf{A}\,|\mathbf{x}|$, so $0 < |x_k| = \big(\mathbf{A}\,|\mathbf{x}|\big)_k = \sum_{j=1}^n a_{kj}|x_j|$. But it's also true that $|x_k| = |\lambda|\,|x_k| = |(\lambda\mathbf{x})_k| = |(\mathbf{Ax})_k| = \big|\sum_{j=1}^n a_{kj}x_j\big|$, and thus

$$\Big|\sum_j a_{kj}x_j\Big| = \sum_j a_{kj}|x_j| = \sum_j |a_{kj}x_j|. \tag{8.2.9}$$

For nonzero vectors $\{\mathbf{z}_1, \ldots, \mathbf{z}_n\} \subset \mathcal{C}^n$, it's a fact that $\|\sum_j \mathbf{z}_j\|_2 = \sum_j \|\mathbf{z}_j\|_2$ (equality in the triangle inequality) if and only if each $\mathbf{z}_j = \alpha_j\mathbf{z}_1$ for some $\alpha_j > 0$ (Exercise 5.1.10, p. 277). In particular, this holds for scalars, so (8.2.9) insures the existence of numbers $\alpha_j > 0$ such that

$$a_{kj}x_j = \alpha_j(a_{k1}x_1) \quad \text{or, equivalently,} \quad x_j = \pi_j x_1 \quad \text{with } \pi_j = \frac{\alpha_j a_{k1}}{a_{kj}} > 0.$$

In other words, if $|\lambda| = 1$, then $\mathbf{x} = x_1\mathbf{p}$, where $\mathbf{p} = (1, \pi_2, \ldots, \pi_n)^T > \mathbf{0}$, so

$$\lambda\mathbf{x} = \mathbf{A}\mathbf{x} \implies \lambda\mathbf{p} = \mathbf{A}\mathbf{p} = |\mathbf{A}\mathbf{p}| = |\lambda\mathbf{p}| = |\lambda|\mathbf{p} = \mathbf{p} \implies \lambda = 1,$$

and thus 1 is the only eigenvalue of $\mathbf{A}$ on the spectral circle. Now suppose that $index(1) = m > 1$. It follows that $\left\|\mathbf{A}^k\right\|_\infty \to \infty$ as $k \to \infty$ because there is an $m \times m$ Jordan block $\mathbf{J}_\star$ in the Jordan form $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ that looks like (7.10.30) on p. 629, so $\left\|\mathbf{J}_\star^k\right\|_\infty \to \infty$, which in turn means that $\left\|\mathbf{J}^k\right\|_\infty \to \infty$ and, consequently, $\left\|\mathbf{J}^k\right\|_\infty = \left\|\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}\right\|_\infty \le \left\|\mathbf{P}^{-1}\right\|_\infty \left\|\mathbf{A}^k\right\|_\infty \left\|\mathbf{P}\right\|_\infty$ implies

$$\left\|\mathbf{A}^k\right\|_\infty \ge \frac{\left\|\mathbf{J}^k\right\|_\infty}{\left\|\mathbf{P}^{-1}\right\|_\infty \left\|\mathbf{P}\right\|_\infty} \to \infty.$$

Let $\mathbf{A}^k = \left[a_{ij}^{(k)}\right]$, and let i_k denote the row index for which $\left\|\mathbf{A}^k\right\|_\infty = \sum_j a_{i_k j}^{(k)}$. We know that there exists a vector $\mathbf{p} > \mathbf{0}$ such that $\mathbf{p} = \mathbf{A}\mathbf{p}$, so for such an eigenvector,

$$\|\mathbf{p}\|_\infty \ge p_{i_k} = \sum_j a_{i_k j}^{(k)} p_j \ge \left(\sum_j a_{i_k j}^{(k)}\right)(\min_i p_i) = \left\|\mathbf{A}^k\right\|_\infty (\min_i p_i) \to \infty.$$

But this is impossible because $\mathbf{p}$ is a constant vector, so the supposition that $index(1) > 1$ must be false, and thus $index(1) = 1$. ■

Establishing that $\rho(\mathbf{A})$ is a semisimple eigenvalue of $\mathbf{A} > \mathbf{0}$ was just a steppingstone (but an important one) to get to the following theorem concerning the multiplicities of $\rho(\mathbf{A})$.

Multiplicities of $\rho(\mathbf{A})$

If $\mathbf{A}_{n\times n} > \mathbf{0}$, then $alg\ mult_{\mathbf{A}}(\rho(\mathbf{A})) = 1$. In other words, the spectral radius of $\mathbf{A}$ is a *simple* eigenvalue of $\mathbf{A}$.

So $\dim N(\mathbf{A} - \rho(\mathbf{A})\mathbf{I}) = geo\ mult_{\mathbf{A}}(\rho(\mathbf{A})) = alg\ mult_{\mathbf{A}}(\rho(\mathbf{A})) = 1$.

Proof. As before, assume without loss of generality that $\rho(\mathbf{A}) = 1$, and suppose that $alg\ mult_{\mathbf{A}}(\lambda = 1) = m > 1$. We already know that $\lambda = 1$ is a semisimple eigenvalue, which means that $alg\ mult_{\mathbf{A}}(1) = geo\ mult_{\mathbf{A}}(1)$ (p. 510), so there are m linearly independent eigenvectors associated with $\lambda = 1$. If $\mathbf{x}$ and $\mathbf{y}$ are a pair of independent eigenvectors associated with $\lambda = 1$, then $\mathbf{x} \ne \alpha\mathbf{y}$ for all $\alpha \in \mathcal{C}$. Select a nonzero component from $\mathbf{y}$, say $y_i \ne 0$, and set $\mathbf{z} = \mathbf{x} - (x_i/y_i)\mathbf{y}$. Since $\mathbf{A}\mathbf{z} = \mathbf{z}$, we know from (8.2.7) on p. 663 that $\mathbf{A}|\mathbf{z}| = |\mathbf{z}| > \mathbf{0}$. But this contradicts the fact that $z_i = x_i - (x_i/y_i)y_i = 0$. Therefore, the supposition that $m > 1$ must be false, and thus $m = 1$. ■

Since $N(\mathbf{A} - \rho(\mathbf{A})\mathbf{I})$ is a one-dimensional space that can be spanned by some $\mathbf{v} > 0$, there is a *unique* eigenvector $\mathbf{p} \in N(\mathbf{A} - \rho(\mathbf{A})\mathbf{I})$ such that $\mathbf{p} > \mathbf{0}$ and $\sum_j p_j = 1$ (it's obtained by the normalization $\mathbf{p} = \mathbf{v}/\|\mathbf{v}\|_1$—see Exercise 8.2.3). This special eigenvector $\mathbf{p}$ is called the ***Perron vector*** for $\mathbf{A} > \mathbf{0}$, and the associated eigenvalue $r = \rho(\mathbf{A})$ is called the ***Perron root*** of $\mathbf{A}$.

Since $\mathbf{A} > \mathbf{0} \iff \mathbf{A}^T > \mathbf{0}$, and since $\rho(\mathbf{A}) = \rho(\mathbf{A}^T)$, it's clear that if $\mathbf{A} > \mathbf{0}$, then in addition to the Perron eigenpair $(r, \mathbf{p})$ for $\mathbf{A}$ there is a corresponding Perron eigenpair $(r, \mathbf{q})$ for $\mathbf{A}^T$. Because $\mathbf{q}^T\mathbf{A} = r\mathbf{q}^T$, the vector $\mathbf{q}^T > \mathbf{0}$ is called the ***left-hand Perron vector*** for $\mathbf{A}$.

While eigenvalues of $\mathbf{A} > \mathbf{0}$ other than $\rho(\mathbf{A})$ may or may not be positive, it turns out that no eigenvectors other than positive multiples of the Perron vector can be positive—or even nonnegative.

No Other Positive Eigenvectors

There are no nonnegative eigenvectors for $\mathbf{A}_{n\times n} > \mathbf{0}$ other than the Perron vector $\mathbf{p}$ and its positive multiples. (8.2.10)

Proof. If $(\lambda, \mathbf{y})$ is an eigenpair for $\mathbf{A}$ such that $\mathbf{y} \geq \mathbf{0}$, and if $\mathbf{x} > \mathbf{0}$ is the Perron vector for $\mathbf{A}^T$, then $\mathbf{x}^T\mathbf{y} > 0$ by (8.2.2), so

$$\rho(\mathbf{A})\,\mathbf{x}^T = \mathbf{x}^T\mathbf{A} \quad\Longrightarrow\quad \rho(\mathbf{A})\,\mathbf{x}^T\mathbf{y} = \mathbf{x}^T\mathbf{A}\mathbf{y} = \lambda\mathbf{x}^T\mathbf{y} \quad\Longrightarrow\quad \rho(\mathbf{A}) = \lambda. \quad \blacksquare$$

In 1942 the German mathematician Lothar Collatz (1910–1990) discovered the following formula for the Perron root, and in 1950 Helmut Wielandt (p. 534) used it to develop the Perron–Frobenius theory.

Collatz–Wielandt Formula

The Perron root of $\mathbf{A}_{n\times n} > \mathbf{0}$ is given by $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$, where

$$f(\mathbf{x}) = \min_{\substack{1\leq i\leq n\\ x_i\neq 0}} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i} \qquad\text{and}\qquad \mathcal{N} = \{\mathbf{x} \mid \mathbf{x} \geq \mathbf{0} \text{ with } \mathbf{x} \neq \mathbf{0}\}.$$

Proof. If $\xi = f(\mathbf{x})$ for $\mathbf{x} \in \mathcal{N}$, then $\mathbf{0} \leq \xi\mathbf{x} \leq \mathbf{A}\mathbf{x}$. Let $\mathbf{p}$ and $\mathbf{q}^T$ be the respective the right-hand and left-hand Perron vectors for $\mathbf{A}$ associated with the Perron root r, and use (8.2.3) along with $\mathbf{q}^T\mathbf{x} > 0$ (by (8.2.2)) to write

$$\xi\mathbf{x} \leq \mathbf{A}\mathbf{x} \quad\Longrightarrow\quad \xi\mathbf{q}^T\mathbf{x} \leq \mathbf{q}^T\mathbf{A}\mathbf{x} = r\mathbf{q}^T\mathbf{x} \quad\Longrightarrow\quad \xi \leq r \quad\Longrightarrow\quad f(\mathbf{x}) \leq r \ \ \forall\, \mathbf{x} \in \mathcal{N}.$$

Since $f(\mathbf{p}) = r$ and $\mathbf{p} \in \mathcal{N}$, it follows that $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$. $\blacksquare$

Below is a summary of the results obtained in this section.

Perron's Theorem

If $\mathbf{A}_{n\times n} > \mathbf{0}$ with $r = \rho(\mathbf{A})$, then the following statements are true.

- $r > 0$. (8.2.11)
- $r \in \sigma(\mathbf{A})$ (r is called the ***Perron root***). (8.2.12)
- $alg\ mult_{\mathbf{A}}(r) = 1$. (8.2.13)
- There exists an eigenvector $\mathbf{x} > \mathbf{0}$ such that $\mathbf{Ax} = r\mathbf{x}$. (8.2.14)
- The ***Perron vector*** is the unique vector defined by

$$\mathbf{Ap} = r\mathbf{p}, \qquad \mathbf{p} > \mathbf{0}, \qquad \text{and} \qquad \|\mathbf{p}\|_1 = 1,$$

  and, except for positive multiples of $\mathbf{p}$, there are no other nonnegative eigenvectors for $\mathbf{A}$, regardless of the eigenvalue.
- r is the only eigenvalue on the spectral circle of $\mathbf{A}$. (8.2.15)
- $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$ (***the Collatz–Wielandt formula***),

$$\text{where} \quad f(\mathbf{x}) = \min_{\substack{1\le i\le n \\ x_i \neq 0}} \frac{[\mathbf{Ax}]_i}{x_i} \quad \text{and} \quad \mathcal{N} = \{\mathbf{x} \mid \mathbf{x} \geq \mathbf{0} \text{ with } \mathbf{x} \neq \mathbf{0}\}.$$

Note: Our development is the reverse of that of Wielandt and others in the sense that we first proved the existence of the Perron eigenpair $(r, \mathbf{p})$ without reference to $f(\mathbf{x})$, and then we used the Perron eigenpair to established the Collatz-Wielandt formula. Wielandt's approach is to do things the other way around—first prove that $f(\mathbf{x})$ attains a maximum value on $\mathcal{N}$, and then establish existence of the Perron eigenpair by proving that $\max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x}) = \rho(\mathbf{A})$ with the maximum value being attained at a positive eigenvector $\mathbf{p}$.

Exercises for section 8.2

8.2.1. Verify Perron's theorem by by computing the eigenvalues and eigenvectors for

$$\mathbf{A} = \begin{pmatrix} 7 & 2 & 3 \\ 1 & 8 & 3 \\ 1 & 2 & 9 \end{pmatrix}.$$

Find the right-hand Perron vector $\mathbf{p}$ as well as the left-hand Perron vector $\mathbf{q}^T$.

8.2.2. Convince yourself that (8.2.2)–(8.2.5) are indeed true.

8.2.3. Provide the details that explain why the Perron vector is uniquely defined.

8.2.4. Find the Perron root and the Perron vector for

$$\mathbf{A} = \begin{pmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{pmatrix},$$

where $\alpha + \beta = 1$ with $\alpha, \beta > 0$.

8.2.5. Suppose that $\mathbf{A}_{n\times n} > \mathbf{0}$ has $\rho(\mathbf{A}) = r$.

(a) Explain why $\lim_{k\to\infty}(\mathbf{A}/r)^k$ exists.

(b) Explain why $\lim_{k\to\infty}(\mathbf{A}/r)^k = \mathbf{G} > \mathbf{0}$ is the projector onto $N(\mathbf{A} - r\mathbf{I})$ along $R(\mathbf{A} - r\mathbf{I})$.

(c) Explain why $rank\,(\mathbf{G}) = 1$.

8.2.6. Prove that if every row (or column) sum of $\mathbf{A}_{n\times n} > \mathbf{0}$ is equal to ρ, then $\rho\,(\mathbf{A}) = \rho$.

8.2.7. Prove that if $\mathbf{A}_{n\times n} > \mathbf{0}$, then

$$\min_i \sum_{j=1}^{n} a_{ij} \le \rho\,(\mathbf{A}) \le \max_i \sum_{j=1}^{n} a_{ij}.$$

Hint: Recall Example 7.10.2 (p. 619).

8.2.8. To show the extent to which the hypothesis of positivity cannot be relaxed in Perron's theorem, construct examples of square matrices $\mathbf{A}$ such that $\mathbf{A} \ge \mathbf{0}$, but $\mathbf{A} \not> \mathbf{0}$ (i.e., $\mathbf{A}$ has at least one zero entry), with $r = \rho\,(\mathbf{A}) \in \sigma\,(\mathbf{A})$ that demonstrate the validity of the following statements. Different examples may be used for the different statements.

(a) r can be 0.

(b) $alg\ mult_{\mathbf{A}}\,(r)$ can be greater than 1.

(c) $index(r)$ can be greater than 1.

(d) $N(\mathbf{A} - r\mathbf{I})$ need not contain a positive eigenvector.

(e) r need not be the only eigenvalue on the spectral circle.

8.2.9. Establish the min-max version of the Collatz–Wielandt formula that says the Perron root for $\mathbf{A} > \mathbf{0}$ is given by $r = \min_{\mathbf{x}\in\mathcal{P}} g(\mathbf{x})$, where

$$g(\mathbf{x}) = \max_{1\le i\le n} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i} \quad \text{and} \quad \mathcal{P} = \{\mathbf{x} \,|\, \mathbf{x} > \mathbf{0}\}.$$

8.2.10. Notice that $\mathcal{N} = \{\mathbf{x} \,|\, \mathbf{x} \ge \mathbf{0} \text{ with } \mathbf{x} \ne \mathbf{0}\}$ is used in the max-min version of the Collatz–Wielandt formula on p. 666, but $\mathcal{P} = \{\mathbf{x} \,|\, \mathbf{x} > \mathbf{0}\}$ is used in the min-max version in Exercise 8.2.9. Give an example of a matrix $\mathbf{A} > \mathbf{0}$ that shows $r \ne \min_{\mathbf{x}\in\mathcal{N}} g(\mathbf{x})$ when $g(\mathbf{x})$ is defined as

$$g(\mathbf{x}) = \max_{\substack{1\le i\le n \\ x_i\ne 0}} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i}.$$