Matrix Analysis
Chapter 8

Contents

8.1 Inequalities and Genralities

Notation and Definition

Let $A = [a_{ij}] \in M_{mn}$, and $B = [b_{ij}] \in M_{m,n}$ and define $|A| = [|a_{ij}|$. If A and B have real entries, we write
$A \geq 0$ if all $a_{ij} > 0$, and $A > 0$ if all $a_{ij} > 0$
$A \geq B$ if $A - B \geq 0$, and $A > B$ if $A - B > 0$. The reversed relations $\leq$ and $<$ are defined similarly. If $A \geq 0$, we say that A is a *nonnegative* matrix, and if $A > 0$, we say that A is a *positive* matrix.

Exercise

Let $A, B \in M_{m,n}$. Show that

(8.1.1) $|A| \geq 0$ and $|A| = 0$ if and only if $A = 0$

(8.1.2) $|aA| = |a||A|$ for all $a \in \boldsymbol{C}$

(8.1.3) $|A + B| \leq |A| + |B|$

(8.1.4) $A \geq 0$ and $A \neq 0 \Rightarrow A > 0$ only if $m = n = 1$

(8.1.5) if $A \geq 0, B \geq 0$, and $a, b \geq 0$, then $aA + bB \geq 0$

(8.1.6) if $A \geq B$ and $C \geq D$, then $A + C \geq B + D$

(8.1.7) if $A \geq B$ and $B \geq C$, then $A \geq C$

Proposition 8.1.8. Let $A = [a_{ij}] \in M_n$ and $x = [x_i] \in \boldsymbol{C}^n$ be given.

(a) $|Ax| \leq |A||x|$.

(b) Suppose that A is nonnegative and has a positive row. If $|Ax| = A|x|$, then there is a real $\theta \in [0, 2\pi])$ such that $e^{-i\theta}x = |x|$.

(c) Suppose that x is positive. If $Ax = |A|x$, then $A = |A|$, so A in nonnegative.

Proof

(a) The assertion follows from the triangle inequality:

$$|Ax|_k = |\sum_j a_{kj}x_j| \leq \sum_j |a_{kj}x_j| = \sum_j |a_{kj}||x_j| = (|A||x|)_k$$

for each k=1,...,n.

(b) The hypothesis is that $A \geq 0$, $a_{k1}, ..., a_{kn}$ are all positive, and $|Ax| = A|x|$. Then $|Ax|_k = |\Sigma_j a_{kj}x_j| = \Sigma_j a_{kj}|x_j| = (A|x|)_k$. This is a case of equality in the triangle inequality (8.1.8.1), so there is a $\theta \in \boldsymbol{R}$ such that $e^{-i\theta}a_{kj}x_k = a_{kj}|x_j|$ for each $j = 1, ..., n$; see Appendix A. Since each a_{kj} is positive, it follows that $e^{-i\theta}x_j = |x_j|$ for each $j = 1, ..., n$, that is, $e^{-i\theta}x = |x|$.

(c) We have $|A|x = Re(|A|x) = Re(Ax) - (Re\ A)x$ so $(|A| - Re\ A)x = 0$. But $|A| - Re\ A \geq 0$ and $x > 0$, so (8.1.1) ensures that $|A| = Re\ A$.Then $A = |A| \geq 0$.

Exercise

Let $A, B, C, D \in M_n$, let $x, y \in \boldsymbol{C}^n$, and $m \in \{1, 2, ...\}$. Show that

(8.1.9) $|AB| \leq |A||B|$

(8.1.10) $|A^m| \leq |A|^m$

(8.1.11) if $0 \leq A \leq B$ and $0 \leq C \leq D$, then $0 \leq AC \leq BD$

(8.1.12) if $0 \leq A \leq B$, then $0 \leq A^m \leq B^m$

(8.1.13) if $A \geq 0$, then $A^m \geq 0$; if $A > 0$, then $A^m > 0$

(8.1.14) if $A > 0, x \geq 0$, and $x \neq 0$, then $Ax > 0$

(8.1.15) if $A \geq 0, x > 0$, and $Ax = 0$, then $A = 0$

(8.1.16) if $|A| \leq |B|$, then $\|A\|_2 \leq \|B\|_2$

(8.1.17) $\|A\|_2 = \||A|\|_2$

Theorem 8.1.18. Let $A, B \in M_n$ and suppose that B is nonnegative. If $|A| \leq B$, then $\rho(A) \leq \rho(|A|) \leq \rho(B)$.

Proof. Invoking (8.1.1.10), we have $|A^m| \leq |A|^m \leq B^m$ for each $m = 1, 2, \ldots$. Thus, (8.1.16) and (8.1.17) ensure that
$\|A^m\|_2 \leq \||A|^m\|_2 \leq \|B^m\|$ and $\|A^m\|_2^{1/m} \leq \||A|^m\|_2^{1/m} \leq \|B^m\|_2^{1/m}$
for each $m = 1, 2, \ldots$. If we now let $m \to \infty$ and apply the Gelfand formula (5.6.14), we deduce that $\rho(A) \leq \rho(|A|) \leq \rho(B)$.

Corollary 8.1.19 Let $A, B \in M_n$ be nonnegative. If $0 \leq A \leq B$, then $\rho(A) \leq \rho(B)$.

Corollary 8.1.20. Let $A = [a_{ij}] \in M_n$ be nonnegative.
(a) If $\hat{A}$ is pricipal submatrix of A, then $\rho(\hat{A}) \leq \rho(A)$.
(b) $\max_{i=1,\ldots,n} a_{ii} \leq \rho(A)$
(c) $\rho(A) > 0$ if any main diagonal entry of A is positive

Proof

(a) If $r = n$, there is nothing to prove. Suppose that $1 \leq r < n$, let $\hat{A}$ be an r-by-r pricipal square submatrix of A, and let P be a permutation matrix such that $PAP^T = \begin{bmatrix} \hat{A} & B \\ C & D \end{bmatrix}$. The preceding theorem ensures that

$$\rho(\hat{A}) = \rho(\hat{A} \oplus 0_{n-r}) = \rho\left(\begin{bmatrix} \hat{A} & 0 \\ 0 & 0 \end{bmatrix}\right) \quad \leq \rho\left(\begin{bmatrix} \hat{A} & B \\ C & D \end{bmatrix}\right) = \rho(PAP^T) = \rho(PAP^{-1}) = \rho(A)$$

(b) Take $r = 1$ to see that $a_{ii} \leq \rho(A)$ for all $i = 1, ..., n$.

(c) $\rho(A) \geq \max_{i=1,\ldots,n} a_{ii} > 0$.

Exercise

- The hypothesis that A is nonnegative is essential for the inequalities in (8.1.20). Consider $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$. Is $1 \leq \rho(A)$?
- $\rho(A) = 0$.

Theorem 8.1.21. Let $A = [a_{ij}] \in M_n$ be nonnegative. Then $\rho(A) \leq |||A|||_\infty = \max_{1 \leq i \leq n} \Sigma_{j=1}^n a_{ij}$ and $\rho(A) \leq |||A|||_1 = \max_{1 \leq j \leq n} \Sigma_{i=1}^n a_{ij}$. If all the row sums of A are equal, then $\rho(A) = |||A|||_\infty$; if all the column sums of A are equal, then $\rho(A) = |||A|||_1$.

Proof. We know that $|\lambda| \leq \rho(A) \leq |||A|||$ for any eigenvalue λ of A and any matrix norm $|||\cdot|||$. If all the row sums of A are equal, then $e = [1...1]^T$ is an eigenvector of A with eigenvalue $\lambda = |||A|||_\infty$ and so $|||A|||_\infty = \lambda \leq \rho(A) = |||A|||_\infty$. The statement for column sums follows from applying the same argument to A^T.

Theorem 8.1.22. Let $A = [a_{ij}] \in M_n$ be nonnegative. Then

$$\min_{1 \le i \le n} \sum_{j=1}^{n} a_{ij} \le \rho(A) \le \max_{1 \le i \le n} \sum_{j=1}^{n} a_{ij}$$

and

$$\min_{1 \le j \le n} \sum_{i=1}^{n} a_{ij} \le \rho(A) \le \max_{1 \le j \le n} \sum_{i=1}^{n} a_{ij}$$

Train of Thought

- Convert the left side to the spectrum radius of a matrix B.
- Compare A and B and conduct the relationship of their spectrum radius.

Proof

Let $\alpha = \min_{1 \leq i \leq n} \Sigma_{j=1}^{n} a_{ij}$. If $\alpha = 0$, ket $B = 0$. If $\alpha > 0$, define $B = [b_{ij}]$ by letting each $b_{ij} = \alpha a_{ij}(\Sigma_{k=1}^{n} a_{ik})^{-1}$. Then $A \leq B \leq 0$ and $\Sigma_{j=1}^{n} b_{ij} = \alpha$ for all $i = 1, ..., n$. The preceding lemma ensures that $\rho(B) = \alpha$, and (8.1.19) tells us that $\rho(B) \leq \rho(A)$. The upper bound in (8.1.23) is the nor =m bound in (8.1.21). The column sum bounds floolw from applying the row sum bounds to A^T.

Corollary 8.1.25. Let $A = [a_{ij}] \in M_n$. If A is nonnegative and either $\Sigma_{j=1}^{n} a_{ij} > 0$ for all $i = 1, ..., n$ or $\Sigma_{i=1}^{n} a_{ij} > 0$ for all $j = 1, ..., n$, then $\rho(A) > 0$. In pariticular, $\rho(A) > 0$ if $n \geq 2$ and A is irreducible and nonnegative.

We can generalize the preceding theorem by introducing some free parameters. If $A \leq 0$,$S = diag(x_1, ..., x_n)$, and all $x_i > 0$, then $S^{-1}AS = [a_{ij}x_i^{-1}x_j] \geq 0$ and $\rho(A) = \rho(S^{-1}AS)$. Applying (8.1.22) to $S^{-1}AS$ yields the following result.

Theorem 8.1.26. Let $A = [a_{ij}] \in M_n$ be nonnegative. Then for any postive vector $x = [x_i] \in \boldsymbol{R}^n$ we have

$$\min_{1\leq i\leq n} \frac{1}{x_i}\sum_{j=1}^{n} a_{ij}x_j \leq \rho(A) \leq \max_{1\leq i\leq n} \frac{1}{x_i}\sum_{j=1}^{n} a_{ij}x_j$$

and

$$\min_{1\leq j\leq n} x_j\sum_{i=1}^{n} \frac{a_{ij}}{x_i} \leq \rho(A) \leq \max_{1\leq j\leq n} x_j\sum_{i=1}^{n} \frac{a_{ij}}{x_i}$$

Corollary 8.1.30. Let $A = [a_{ij}] \in M_n$ be nonnegative and let x$x = [x_i] \in \boldsymbol{R}^n$ be a positive vector. If $\alpha, \beta \geq 0$ are such that $\alpha x \leq Ax \leq \beta x$, then $\alpha \leq \rho(A) \leq \beta$. If $\alpha x < Ax$, then $\alpha < \rho(A)$;; if $Ax < \beta x$, then $\rho(A) < \beta$.

Proof. If $\alpha x \leq Ax$, then $\alpha x_i \leq (Ax)_i$ and $\alpha \leq \min_{1 \leq i \leq n} x_i^{-1} \Sigma_{j=1}^{n} a_{ij} x_j$, so the preceding theorem ensures that $\alpha \leq \rho(A)$. If $\alpha x < Ax$, then there is some$\alpha' > \alpha$ such that $\alpha x < \alpha' x \leq Ax$. In this event, $\rho(A) \geq \alpha' > \alpha$. The upper bounds can be verified in a similar fashion.

Let $\alpha = \beta = \lambda$

Corollary 8.1.30. Let $A \in M_n$ be nonnegative. If x is a positive eigenvector of A, then $\rho(A)$,x is an eigenpair for A; that is, if $A \geq 0, x > 0$, and $Ax = \lambda x$, then $\lambda = \rho(A)$.

8.1.31. Let $A = [a_{ij}] \in M_n$ be nonnegative. If A has a positive eigenvector, then

$$\rho(A) = \max_{x>0} \min_{1\le i\le n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j = \min_{x>0} \max_{1\le i\le n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j$$

Proof. Let x be an positive eigenvector of A. Since $Ax = \lambda x$,we have $(Ax)_i = \lambda x_i$. That is, $\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i$. So $\frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j = \frac{1}{x_i} \cdot \lambda x_i = \lambda$. According to Corollary 8.1.30, $\lambda = \rho(A)$. Then $\rho(A) = \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j$ and the equality is Theorem 8.1.26 qualifies.

Corollary 8.1.33. Let $A = [a_{ij} \in M_n$ be nonnegative and write $A^m = [a_{ij}^{(m)}$. If A has a positive eigenvector $x = [x_i]$, then for all $m = 1, 2, ...$ and for all $i = 1, 2, ..., n$ we have

$$\sum_{j=1}^{n} a_{ij}^{(m)} \leq \left(\frac{\max_{1 \leq k \leq n} x_k}{\min_{1 \leq k \leq n} x_k} \right) \rho(A)^m$$

and

$$\left(\frac{\min_{1 \leq k \leq n} x_k}{\max_{1 \leq k \leq n} x_k} \right) \rho(A)^m \leq \sum_{j=1}^{n} a_{ij}^{(m)}$$

If $\rho(A) > 0$, then the entries of $[\rho(A)^{-1} A]^m$ are uniformly bounded for $m = 1, 2, ...$

Proof

Let $x = [x_i]$ be a positive eigenvector of A. Then (8.1.30) ensures that $Ax = \rho(A)x$, so $A^m x = \rho(A)^m x$ for each $m = 1, 2, ...$ Since $A^m \geq 0$, for any $i = 1, ..., n$ we have

$$\begin{aligned} \rho(A)^m \max_{1\leq k\leq n} x_k \geq \rho(A^m)x_i = (A^m x)_i &= \sum_{j=1}^{n} a_{ij}^{(m)} x_j \\ &\geq \left(\min_{1\leq k\leq n} x_k\right) \sum_{j=1}^{n} a_{ij}^{(m)} \end{aligned}$$

Since $\min_{1\leq k\leq n} x_k > 0$, the asserted upper bound is proved. The asserted lower bound follows in a similar fashion.

8.2 Positive Matrices

Lemma 8.2.1. Let $A \in M_n$ be positive. If λ, x is an eigenpair of A and $|\lambda| = \rho(A)$, then $|x| > 0$ and $A|x| = \rho(A)|x|$.

Proof. The hypotheses ensure that $z = A|x| > 0$ (8.1.14). We have $z = A|x| \geq |Ax| = |\lambda x| = |\lambda||x| = \rho(A)|X|$, so $y = z - \rho(A)|x| \geq 0$. If $y = 0$, then $\rho(A)|x| = A|x| > 0$, so $\rho(A) > 0$ and $|x| > 0$. If, however, $y \neq 0$, (8.1.14) again ensures that $0 < Ay = Az - \rho(A)A|x| = Az - \rho(A)z$, in which case $Az > \rho(A)z$. It follows from (8.1.29) that $\rho(A) > \rho(A)$, which is not possible. We conclude that $y = 0$.

Theorem 8.2.2. If $A \in M_n$ is positive, there are positive vectors x and y such that $Ax = \rho(A)x$ and $y^T A = \rho(A)y^T$.

Proof. There is an eigenpair λ, x of A with $|\lambda| = \rho(A)$. The preceding lemma ensures that $\rho(A), |x|$ is also an eiganpair of A and $|x| > 0$. The assertion about y follows from considersing A^T.

Exercise

- If $A \in M_n$ and $A > \mathbf{0}$, use (8.1.31) and the preceding theorem to explain why $\rho(A) = \underset{x>0}{max} \; \underset{i}{min} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j = \underset{x>0}{min} \; \underset{i}{max} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j$.
- A>0. Exists positive eigenvector.

Lemma 8.2.3. Let $A \in M_n$ be positive. If λ, x is an eigenpair of A and $|\lambda| = \rho(A)$, then there is a $\theta \in \boldsymbol{R}$ such that $e^{-i\theta}x = |x| > 0$.

Proof. The hypothesis is that $x \in \boldsymbol{C}^n$ is nonzero and $|Ax| = |\lambda x| = \rho(A)|x|$; (8.2.1) ebsures that $A|x| = \rho(A)|x|$ and $|x| > 0$. Since $|Ax| = \rho(A)x = A|x|$ and some (in fact, every) row of A is positive, (8.1.8b) ensures that there is a $\theta \in \boldsymbol{R}$ such that $e^{-i\theta}x = |x|$.

Theorem 8.2.4. Let $A \in M_n$ be positive. If λ is an eigenvalue of A and $\lambda \neq \rho(A)$, then $|\lambda| < \rho(A)$.

Proof. Let λ, x be an eigenpair of A, so $|\lambda| \leq \rho(A)$, (8.2.3) ensures that $w = e_{-i\theta}x > 0$ for some $\theta \in \boldsymbol{R}$. Since $Aw = \lambda w$ and $w > 0$, it follows from (8.1.30) that $|\lambda| = \rho(A)$.

Theorem 8.2.5. If $A \in M_n$ is positive, then the geometric multiplicity of $\rho(A)$ as an eigenvalue of A is 1.

Proof

Suppose that $w, z \in \boldsymbol{C}^n$ are nonzero vectors such that $Aw = \rho(A)w$ and $Az = \rho(A)z$. We need to prove that $w = \alpha z$ for some $\alpha \in \boldsymbol{C}$.

To better compare the two vectors, we first transform them into real vectors. Lemma 8.2.3 ensures that there are real numbers θ_1 and θ_2 such that $p = [p_j] = e^{-i\theta_1} z >=$ and $q = [q_j] = e^{-i\theta_2} w > 0$.

Then we should make a difference between p and q to verify if p is a multiplication of q. Let $\beta = \min_{1 \leq i \leq n} q_i p_i^{-1}$ and let $r - q - \beta p$. Notice that $r \geq 0$ and at least one entry of r is zero. If $r \neq 0$, then $0 < Ar = Aq - \beta Ap = \rho(A)q = \beta\rho(A)p = \rho(A)(q - \beta p) = \rho(A)r$, so $\rho(A)r > 0$ and r¿0, which is a contradiction. We conclude that $r = 0$, $q = \beta p$, and $w = \beta e^{i(\theta_2 - \theta_1)} z$.

Corollary 8.2.6. Let $A \in M_n$ be positive. There is a unique vector $x = [x_i] \in \boldsymbol{C}^n$ such that $Ax = \rho(A)x$ and $\Sigma_i x_i = 1$. Such a vector must be positive.

Proof. Suppose $Ax_1 = \rho(A)x_1$ and $Ax_2 = \rho(A)x_2$, then 8.2.5 tells that $x_1 = \beta x_2$. Add entries together. So$1 = \beta \Rightarrow x_1 = x_2$.

Notation

- The unique normalized eigenvector characterized in (8.2.6) is the **Perron** vector of A, sometimes called **right Perron vector**; $\rho(A)$ is the **Perron root** of A.
- Of course, the matrix A^T is positive if A is positive, so all the preceding results about eigenvalues of A apply to A^T as well. An eigenvector $y = [y_i]$ of A^T corresponding to the eigenvalue of $\rho(A)$ and normalized so that $\Sigma_i x_i y_i = 1$ is positive and unique; it is the **left Perron vector** of A.

Theorem 8.2.7. Let $A \in M_n$ be positive. The algebraic multiplicity of $\rho(A)$ as an eigenvalue of A is 1. If x and y are the right and left Perron vectors of A, then $\lim_{m\to\infty}(\rho(A)^{-1}A)^m = xy^T$, which is a positive rank-one matrix.

Proof

We know that $\rho(A) > 0$, and that x and y are positive vectors such that $Ax = \rho(A)x$, $y^T A = \rho(A)y^T$, and $y * x = y^T x = 1$. Theorem 1.4.12b ensures that $\rho(A)$ has algebraic multiplicity 1, and (1.4.7b) tells us that there is a nonsingular $S = [x\ S_1]$ such that $S^{-*} = [y\ Z_1]$ and $A = S([\rho(A)] \oplus B)S^{-1}$. Since there is a simple eigenvalue of A that is its eigenvalue of strictly largest modulus, $\rho(B) < \rho(A)$, that is, $\rho(\rho(A)^{-1}B) < 1$.Theorem 5.6.12 ensures that

$$\left(\frac{1}{\rho(A)}\right)^m = \begin{bmatrix} 1 & 0 \\ 0 & (\rho(A)^{-1}B)^m \end{bmatrix}$$

$$\rightarrow [x \quad S_1] \begin{bmatrix} 1 & 0 \\ 0 & 0_{n-1} \end{bmatrix} \begin{bmatrix} y^T \\ X_1^T \end{bmatrix} = xy^T as\ m \rightarrow \infty$$

Summarize

Theorem 8.2.8 (Perron). Let $A \in M_n$ be positive. Then

(a) $\rho(A)0$

(b) $\rho(A)$ is an algebraically simple eigenvalue of A

(c) there is a unique real vector $x = [x_i]$ such that $Ax = \rho(A)x$ and $x_1 + ... + x_n = 1$; this vector is positive

(d) there is a unique real vector $y = [y_i]$ such that $y^T A = \rho(A)y^T$ and $x_1y_1 + ... + x_ny_n = 1$; this vector is positive

(e) $|\lambda| < \rho(A)$ for every eigenvalue λ of A such that $\lambda \neq \rho(A)$

(f) $(\rho(A)^{-1}A)^m \to xy^T$ as $m \to \infty$

Theorem 8.2.9 (Fan). Let $A = [a_{ij}] \in M_n$. Suppose that $B = [b_{ij}] \in M_n$ is nonnegative and $b_{ij} \geq |a_{ij}|$ for all $i \neq j$. Then every eigenvalue of A is an union of n discs

$$\bigcup_{i=1}^{n} \{z \in \boldsymbol{C} : |z - a_{ii}| \leq \rho(B) - b_{ii}\}$$

In particular, A is nonsingular if $|a_{ii} > \rho(B) - b_{ii}$ for all $i = 1, ..., n$.

Proof

First, assume that $B > 0$. Theorem 8.2.8 ensures that there is a positive vector x such that $Bx = \rho(B)x$, and hence

$$\sum_{j \neq i} |a_{ij}| x_j \leq \sum_{j \neq i} b_{ij} x_j = \rho(B) x_i - b_{ii} x_i$$

for each i=1,...,n.Thus, we have

$$\frac{1}{x_i} \sum_{j \neq i} |a_{ij}| x_j \leq \rho(B) - b_{ii}$$

for each i=1,...,n.The result follows from (6.1.6) with $p_1 = x_i$.

If some entry of B is zero, consider $B_\epsilon = B + \epsilon J_n$ for $\epsilon > 0$. Then $b_{ij} + \epsilon > |a_{ij}|$ for all $i \neq j$, so Ky Fan's eigenvalue inclusion set with respect to B_ϵ is a union of n disks of the form $\{z \in \boldsymbol{C} : |z - a_{ii}| \leq \rho(B_\epsilon) - (b_{ii} + \epsilon)\}$. The assertion for a nonnegative B now follows from observing that $\rho(B_\epsilon) - (b_{ii} + \epsilon) \to \rho(B) - b_{ii}$ as $\epsilon \to 0$.

If $|a_{ii}| > \rho(B) - b_{ii}$ for all i=1,...,n, then $z = 0$ is not in the set (8.2.9a).

8.3 Nonnegative Matrices

Theorem 8.3.1. If $A \in M_n$ is nonnegative, then $\rho(A)$ is an eigenvalue of A and there is a nonnegative nonzero vector x such that $Ax = \rho(A)x$.

Proof

For any $\epsilon > 0$, define $A(\epsilon) = A + \epsilon J_n$. Let $x(\epsilon) = [x(\epsilon)_i]$ be the Perron vector of $A(\epsilon)$, so $x(\epsilon) > 0$ and $\Sigma_{i=1}^{n} x(\epsilon)_i = 1$. Since the set of vectors $\{x(\epsilon) : \epsilon > 0\}$ is contained in the compact set $\{x : x \in \boldsymbol{C}^n, \|x\|_1 \leq 1\}$, there is a monotone decreasing sequence $\epsilon_1 \geq \epsilon_2 \geq ...$ with $\lim_{k\to\infty} \epsilon_k = 0$ such that $\lim_{k\to\infty} x(\epsilon_k) = x$ exists. Since $x(\epsilon_k) > 0$ and $\|x(\epsilon_k)\|_1 = 1$ for all $k = 1, 2, ...$, the limit vector $x = \lim_{k\to\infty} x(\epsilon_k)$ must be nonnegative and nonzero (indeed, $\|x\|_1 = 1$). Theorem 8.1.18 ensures that $\rho(A(\epsilon_k)) \geq \rho(A(\epsilon_{k+1})) \geq ... \geq \rho(A)$ for all $k = 1, 2, ...$, so $\rho = \lim_{k\to\infty} \rho(A(\epsilon_k))$ exists and $\rho \geq \rho(A)$. However, $x \neq 0$ and

$$Ax = \lim_{k\to\infty} A(\epsilon_k)x(\epsilon_k) = \lim_{k\to\infty} \rho(A(\epsilon_k))x(\epsilon_k) \quad = \lim_{k\to\infty} \rho(A(\epsilon_k)) \lim_{k\to\infty} x(\epsilon) = \rho x$$

so ρ is an eigenvalue of A. It follows that $\rho \leq \rho(A)$, so $\rho = \rho(A)$

Theorem 8.3.2. Let $A \in M_n$ be nonnegative, and let $x \in \boldsymbol{R}^n$ be nonnegative and nonzero. If $\alpha \in \boldsymbol{R}$ and $Ax \geq \alpha x$, then $\rho(A) \geq \alpha$.

Proof. Let $A = [a_{ij}]$, let $\epsilon > 0$, and define $A(\epsilon) = A + \epsilon J_n > 0$. Then $A(\epsilon)$ has a positive left Perron vector $y(\epsilon)$: $y(\epsilon)^T A(\epsilon) = \rho(A(\epsilon))y(\epsilon)^T$. We are given that $Ax = \alpha x \geq 0$, so $A(\epsilon)x - \alpha x > Ax - \alpha x \geq 0$ and hence $y(\epsilon)^T(A(\epsilon))x - \alpha x) = (\rho(A(\epsilon)) - \alpha)t(\epsilon)^T x > 0$. Since $y(\epsilon)^T x > 0$, we have $\rho(A(\epsilon)) - \alpha > 0$ for all $\epsilon > 0$. But $\rho(A(\epsilon)) \to \rho(A)$ as $\epsilon \to 0$, so we conclude that $\rho(A) \geq \alpha$.

Corollary 8.3.3. If $A \in M_n$ is nonnegative, then

$$\rho(A) = \max_{\substack{x \geq 0 \\ x \neq 0}} \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{1}{x_i} \Sigma_{j=1}^{n} a_{ij} x_j$$

Main Steps

- $\rho(A) \geq \min\limits_{\substack{1 \geq i \geq n \\ x_i \neq 0}} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j$
- The equality can be attained.

Proof

Let x be any nonzero nonnegative vector and let $\alpha = min_{x_i \neq 0} \Sigma_j a_{ij} x_j / x_i$. Then $Ax \geq \alpha x$, so the preceding theorem ensures that $\rho(A) \geq \alpha$, and hence

$$\rho(A) \geq \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{1}{x_i} \Sigma_{j=1}^{n} a_{ij} x_j$$

Now use (8.3.1) to choose a nonzero nonnegative x such that $Ax = \rho(A)x$, which shows that equality can be attained with $\alpha = \rho(A)$.

Note

- Consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Explain why the implication $Ax \leq \alpha x \Rightarrow \rho(A) \leq \alpha$ in (8.1.29) need not be correct if the nonnegative vector x is not positive. Show that the "min max" characterization in (8.1.32) need not be correct for nonnegative matrices.
- $\rho(A) = 2$

Theorem 8.3.4. Let $A \in M_n$ be nonnegative. Suppose that there is a positive vector x and a nonnegative real number λ such that either $Ax = \lambda x$ or $x^T A = \lambda x^T$. then $\lambda = \rho(A)$.

Proof. Suppose that $x = [x_i] \in \boldsymbol{R}^n$ and $Ax = \lambda x$. Let $D = diag(x_1, ..., x_n)$ and define $B = D^{-1}AD$, which has the same eigenvalues as A. Then $Be = D^{-1}ADe = D^{-1}Ax = \lambda D^{-1}x = \lambda e$, so every row sum of the nonnegative matrix B is equal to λ. It follows from (8.1.21) that $\rho(B) = \lambda$. If $x^T A = \lambda x^T$, apply this argument to A^T.

Theorem 8.3.5. Suppose that $A \in M_n$ is nonnegative and has a positive left eigenvector.
(a) If $x \in M_n$ is nonnegative and $Ax \geq \rho(A)x$, then x is an eigenvector of A corresponding to the eigenvalue $\rho(A)$.
(b) If $A \neq 0$, then $\rho(A) > 0$ and every eigenvalue λ of A such that $|\lambda| = \rho(A)$ is semisimple, that is, every Jordan block of A corresponding to a maximum-modulus eigenvalue is one-by-one.

Proof

Let y be a positive left eigenvector of A. The preceding theorem ensures that $A^T y = \rho(A)y$.
(a) We know that $x \neq 0$ and $Ax - \rho(A)x \geq 0$. We need to show that $Ax - \rho(A)x = 0$. If $Ax - \rho(A)x \neq 0$, then $y^T(Ax - \rho(A)x) > 0$. However, $y^T(Ax - \rho(A)x) = \rho(A)y^T x - \rho(A)y^T x = 0$, which is a contradiction.
(b) Since y is positive and A is nonzero and nonnegative, some entry of $y^T A$ is positive. Consequently, the identity $y^T A = \rho(A)y^T$ ensures that $\rho(A) > 0$. Let $D = diag(y_1, ..., y_n)$ and let $B = \rho(A)^{-1}DAD^{-1}$. It suffices to show that every eigenvalue of B with unit modulus is semisimple. Compute $e^T B = \rho(A)^{-1}e^T DAD^{-1} = \rho(A)^{-1}y^T AD^{-1} = \rho(A)^{-1}\rho(A)yD^{-1} = e^T$. So every column sum of the nonnegative matrix B is one, so B is power bounded and the assertion follows from (3.2.5.2).

Note

- If $A \in M_n$ is nonnegative, its eigenvalue $\rho(A)$ is called the **Perron root** of A.
- Because **an eigenvector (even if normalized) associated with the Perron root of a nonnegative matrix need not be uniquely determined**, there is no well-determined notion of the "Perron vector" for a nonnegative matrix.

8.4 Irreducible Nonnegative Matrices

Lemma 8.4.1. Let $A \in M_n$ be nonnegative. Then A is irreducible if and only if $(I + A)^{n-1} > 0$

Lemma 8.4.2. Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of $A \in M_n$. Then $\lambda_1 + 1, ..., \lambda_n + 1$ are the eigenvalues of I+A and $\rho(I + A) \leq \rho(A) + 1$. If A is nonnegative, then $\rho(I + A) = \rho(A) + 1$.

Proof. The first assertion is easy to prove. To prove the second assertion, we need to prove that $\rho(A) + 1$ is the maximum eigenvalue of $I + A$. The first assertion implies that $\rho(A) + 1$ is an eigenvalue of $I + A$ if $A \geq 0$ since $\rho(A)$ is an eigenvalue of A. We have $\rho(I + A) = max_{1 \leq i \leq n} |\lambda_i + 1| \leq max_{1 \leq i \leq n} |lambda_i| + 1 = \rho(A) + 1$, so $\rho(I + A) = \rho(A) + 1$ in this case.

Lemma 8.4.3. If $A \in M_n$ is nonnegative and A^m is positive for some $m \geq 1$, then $\rho(A)$ is the only maximum-modulus eigenvalue of A; it is positive and algebraically simple.

Proof. Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of A. Then $\lambda_1^m, ..., \lambda_n^m$ are the eigenvalues of A^m. Theorem 8.2.8 ensures that exactly one of $\lambda_1^m, ..., \lambda_n^m$ is equal to $\rho(A^m) = \rho(A)^m$, which is positive; all the rest have modulus strictly less than $\rho(A^m)$. Consequently, $n - 1$ of $\lambda_1, ..., \lambda_n$ are strictly less than $\rho(A)$ in modulus; (8.3.1) ensures that $\rho(A)$ is the remaining eigenvalue.

Theorem 8.4.4.(Perron-Frobenius). Let $A \in M_n$ be irreducible and nonnegative, and suppose that $n \geq 2$. Then

(a)$\rho(A) > 0$ (b)$\rho(A)$ is an algebraically simple eigenvalue of A (c)there is a unique real vector $x = [x_i]$ such that $Ax = \rho(A)x$ and $x_1 + ... + x_n = 1$; this vector is positive (d)there is a unique real vector $y = [y_i]$ such that $y^T A = \rho(A)y^T$ and $x_1y_1 + ... + x_ny_n = 1$; this vector is positive

Proof

(a) Corollary 8.1.25 shows that $\rho(A) > 0$ under conditions even weaker than irreducibility
(b) $\rho(A)$ is an algebraically simple eigenvalue of A, then (8.4.2) ensures that $\rho(A) + I = \rho(I + A)$ is a multiple eigenvalue of $I + A$ and hence $(I + rho(A))^{n-1} = \rho((I + A)^{n-1})$ is a multiple eigenvalue of the positive matrix $(I + A)^{n-1}$, which contradicts ((8.2.8(b)).
(c) Theorem 8.3.1 ensures that there is a nonnegative nonzero vector x such that $Ax = \rho(A)x$. Then $(I + A)^{n-1}x = (\rho(A) + 1)^{n-1}x$, and since $(I + A)^{n-1}$ is positive (8.4.1), it follows from (8.1.14) that $(I + A)^{n-1}x$, and hence also $x = (\rho(A) + 1)^{1-n}(I + A)^{n-1}x$, is positive. If we impose the normalization $e^T x = 1$, then (b) ensures that x is unique.
(d) This follows by applying (c) to A^T

Theorem 8.4.5. Let $A, B \in M_n$. Suppose that A is nonnegative and irreducible, and $A \geq |B|$. Let $\lambda = e^{i\phi}\rho(B)$ be a given maximum-modulus eigenvalue of B. If $\rho(A) = \rho(B)$, then there is a diagonal unitary matrix $D \in M_n$ such that $B = e^{i\phi}DAD^{-1}$

A=|B|

Let x be a nonzero vector such that $Bx = \lambda x$, and let $\rho = \rho(A) = \rho(B)$. Then

$$\rho|x| = |\lambda x| = |Bx| \leq |B||x| \leq A|x|$$

Theorem 8.3.5 and the inequality $A|x| \geq \rho(|x|)$ imply that $A|x| = \rho|x|$, and (8.4.4) ensures that $|X|$ is positive. Equality tells us that $(A - |B|)x = 0$; since x is positive and $A - |B| \geq 0$, (8.1.1) ensures that $A = |B|$.

Construct D

Let D be the unique diagonal unitary matrix such that $x = D|x|$. The idendity $Bx = \lambda x = e^{i\phi}\rho x$ is equivalent to the identity $BD|x| = e^{i\phi}\rho D|x|$, or $e^{-i\phi}D^{-1}BDx = \rho|x| = A|x| = |B||x|$. If we let $C = e^{-i\phi}D^{-1}BD$, we have $C|x| = |C||x|$, so (8.1.8(c)) ensures that $C = |C| = |B| = A$. Thus, $B = e^{i\phi}DAD^{-1}$

$$B = e^{i\phi} DAD^{-1}$$

If we let $C = e^{-i\phi} D^{-1} BD$, we have $C|x| = |C||x|$, so (8.1.8(c)) ensures that $C = |C| = |B| = A$. Thus, $B = e^{i\phi} DAD^{-1}$.

Corollary 8.4.6. Let $A \in M_n$ be irreducible and nonnegative,and suppose that it has exactly k distinct eigenvalues of maximum modulus.Then

(a) A is similar to $e^{2\pi ip/k}$ for each p=0,1,...,k-1

(b) if $J_{m_1}(\lambda) \oplus ... \oplus J_{m_l}(\lambda)$ is a direct summand of the Jordan canonical form of A, and if $p \in \{1, ..., k-1\}$, then $J_{m_1}(e^{2\pi ip/k}\lambda) \oplus ... \oplus J_{m_l}(e^{2\pi ip/k}\lambda)$ is also a direct summand of the Jordan canonical form of A

(c) the maximum-modulus eigenvalues of A are $e^{2\pi ip/k}\rho(A), p = 0, 1, ..., k-1$, and each has algebraic multiplicity 1

Train of Thought

- Since we are not clear about the distribution of the k eigenvalues, just denote them as $\lambda_p = e^{i\phi_p}\rho(A)$.
- Replace each $e^{2\pi ip/k}$ by $e^{i\phi_p}$ then prove (a') and (b') respectively.
- Prove the distribution of ϕ_p.So (c) can be easily illustrated.
- (a) and (b) are also completely proved.

Part 1:Obscure the exponent

- If k=1,there is nothing to prove, so assume that $k \geq 2$.
- Let $\lambda_p = e^{i\phi_p}\rho(A), p = 0,1, \ldots, k-1$,be the distinct maximum-modulus eigenvalues of A, in which $0 = \phi_0 < \phi_1 < \phi_2 < \cdots < \phi_{k-1} < 2\pi$.
- Since A is real, it's easy to observe that $\phi_{k-p} + \phi_p \equiv 0(mod\ 2\pi)$.

Part 2.1 Replace

(a') A is similar to $e^{2\pi ip/k}$ for each p=0,1,...,k-1

(b') if $J_{m_1}(\lambda) \oplus ... \oplus J_{m_l}(\lambda)$ is a direct summand of the Jordan canonical form of A, and if $p \in \{1, ..., k-1\}$, then $J_{m_1}(e^{2\pi ip/k}\lambda) \oplus ... \oplus J_{m_l}(e^{2\pi ip/k}\lambda)$ is also a direct summand of the Jordan canonical form of A

Part 2.2 Prove (a’) and (b’)

- The preceding theorem also shows that B is similar to $e^{i\phi}A$.Apply the theorem with B=A.So A is similar to $e^{i\phi_p}A$ for each p=1,2,⋯,k-1.
- Therefore, if $J_{m_1}(\lambda) \oplus \cdots \oplus J_{m_l}(\lambda)$ is a direct summand of the Jordan canonical form of A, then $J_{m_1}\left(e^{i\phi_p}\lambda\right) \oplus \cdots \oplus J_{m_l}(e^{i\phi_p}\lambda)$ is also a direct summand.

Part 3 Main Steps

- There exists positive integer p such that $e^{ip\phi_1} = 1$.
- Suppose p be the smallest number satisfying the preceding condition. Prove that each element ϕ_m is some positive integer multiple of ϕ_1.
- Prove that p=k. So the distribution is clearfied.

Part3 Prove

Denote $S = \{\phi_0 = 0, \phi_1, \phi_2, ..., \phi_{k-1}\}$.Since A is similar to $e^{i\phi_p}A$ as well as to $e^{i\phi_q}A$, it follows that A is similar to $e^{i(\phi_p+\phi_q)}A$ for any $p, q \in \{0, 1, ..., k-1\}$. That is, for each pair of elements $\phi_p, \phi_q \in S, \phi_p + \phi_p(mod 2\pi)$ is also in S.By induction, we can conclude that $r\phi_1 = \phi_1 + ... + \phi_1(mod 2\pi)$ is in the finite set S for all r=1,2,....
The k+1 elements $\phi_1, 2\phi_1, ..., k\phi_1, (k+1)\phi_1$ of S cannot all be distinct, so there are positive integers $r > s \geq 1$ such that $r\phi_1 = s\phi_1(mod 2\pi)$, in which case $1 < (r-s) \leq k$.It follows that $(r-s)\phi_1 = 0(mod 2\pi)$, that is, $e^{i(r-s)\phi_1} = 1$, so $e^{i\phi_1}$ is a root of unity.

Part3 Prove

Let p be the smallest positive integer such that $e^{ip\phi_1} = 1$.Choose any $\phi_m \in S$.Divide the interval $[0, 2\pi)$ into p half-open subintervals $[0, \phi_1), [\phi_1, 2\phi_1), ..., [(p-1)\phi_1, 2\pi)$. Since ϕ_m is in one of these subintervals, there is some integer q with $0 \leq q \leq p-1$ such that $q\phi_1 \leq \phi_m \leq (q+1)\phi_1$;that is, $0 \leq \phi_m - q\phi_1 < \phi_1$.So each element ϕ_m is some integer multiple of ϕ_1.

Part3 Prove

Now we can see that p=k.Because if $p < k$, there will be fewer than k distinct elements is S.$k\phi_1 = 2\pi$, so $\phi_1 = 2\pi/k$.And $\phi_m = 2\pi m/k$ for each m=0,1,...,k-1.The original theorem is then be proved.

Exercise

- Can an irreducible nonnegative matrix $A \in M_3$ with spectral redius 1 have eigenvalues 1, i, and –i?
- No.
- $tr(A^2) = e^{\frac{4\pi ip}{3}} tr(A^2)$.So $tr(A^2) = \mathbf{0}$.The eigenvalues of A^2 are 1,-1 and 0.So det(A)=0, which contradicts to det(A)=1*i*(-i)=1.

Corollary 8.4.7. Suppose that $A \in M_n$ is irreducible and nonnegative. If A has $k > 1$ eigenvalues of maximum modulus, then every diagonal entry of A is zero.More over, every main diagonal entry of A^m is zero for each positive integer m that is not divisible by k.

Proof

- Let $\phi = 2\pi/k$. Corollary 8.4.6a ensures that A is similar to $e^{i\phi}A$, so A^m is similar to $e^{im\phi}A^m$ for each m=1, 2, 3, ⋯ and $tr(A^m) = e^{im\phi}tr(A^m)$. Since $e^{im\phi}$ is real and positive only if m is an integer multiple of k, this is impossible if A^m has any positive main diagonal entry and m is not divisible by k.

8.5 Primitive Matrices

Definition 8.5.0. A nonnegative mareix $A \in M_n$ is promitive if it is irreducible and has only one nonzero eigenvalue of maximum modulus.

Theorem 8.5.1. If $A \in M_n$ is nonnegative and primitive, and if x and y are, respectively, the right and left Perron vectors of A, then $\lim_{m\to\infty}(\rho(A)^{-1}A)^m = xy^T$, which is a positive rank-one matrix.

It can be proved in the same way as in 8.2.7.

- In practice can one test a given irreducible nonnegative matrix for primitivity without computing its maximum-modulus values?
- The following characterization of primitivity, while not itself a computationally effective test, leads to several useful criteria.

Theorem 8.5.2. If $A \in M_n$ is nonnegative, then A is primitive if and only if $A^m > 0$ for some $m \geq 1$.

Pre-required Knowledge

- Let A be an adjacence matrix of a directed graph $\Gamma(A)$, then $(A^m)_{ij}$ represents the way of moving from node i to j through m steps.
- If $\Gamma(A)$ is strongly connected, then A is irreducible(See theorem 6.2.24 for proof).

Proof

If A^m is positive, there is a directed path of length m between every pair of nodes of the directed graph $\Gamma(A)$ of A, so $\Gamma(A)$ is strongly connected and A is irreducible.In addition, (8.4.3) ensures that there are no maximum-modulus eigenvalues of A other than $\rho(A)$, which is algebraically simple.
Conversely, if A is primitive then $\lim_{m\to\infty}(\rho(A)^{-1}A)^m = xy^T > 0$, so there is some m such that $(\rho(A)^{-1}A)^m > 0$, which means $A^m > 0$.

Note

- From the proof of the preceding theorem, we can also conduct that if $A \in M_n$ is nonnegative and irreducible, and if $A^m > \mathbf{0}$, then $A^p > \mathbf{0}$ for all p=m+1, m+2,⋯.
- The following theorem provides a graph=theoretical criterion for primitivity.

Theorem 8.5.3. Let $A \in M_n$ be irreducible and nonnegative, and let $P_1, ..., P_n$ be the nodes of the directed graph $\Gamma(A)$. Let $L_i = \{l_1^{(i)}, k_2^{(i)}, ...\}$ be the set of lengths of all directed paths in $\Gamma(A)$ that both start and end at the node $P_i, i = 1, 2, ..., n$.Let g_i be the greatest commom divisor of all the lengths in L_i. Then A is primitive if and only if $g_1 = g_2 = ... = g_n = 1$.

Proof

- Irreducibility of A implies that no set L_i is empty.
- If A is primitive, then (8.5.2) ensures that there is some $m \geq 1$ such that $A^m > 0$, and hence $A^k > 0$ for all $k \geq m$. But then $m + p \in L_i$ for each integer $p \geq 1$ and each$i = 1, \ldots, n$, so $g_i = 1$ for all $i = 1, \ldots, n$.
- Suppose that $A = \left[a_{ij}\right]$ is not primitive and has exactly $k > 1$ eigenvalues of maximum-modulus. Corollary 8.4.8 $diag(A^m) = 0$ for all m such that $\mathrm{k} \nmid m$.For each such m, there is no directed path in $\Gamma(A)$ that both starts and ends at any node of $\Gamma(A)$ of length m. Thus, $L_i \subset \{k, 2k, 3k, \ldots\}$, and hence $g_i \geq k > 1$ for each $i = 1, \ldots, n$.

Note

- A theorem of Romanovsky provides additional insight into the preceding result: If $A \in M_n$ is irreducible and nonnegative, then $g_1 = g_2 = \cdots = g_n = k$ is the number of maximum-modulus eigenvalues of A.

Lemma 8.5.4. If $A \in M_n$ is irreducible and nonnegative, and if all its main diagonal entries are positive, then $A^{n-1} > 0$, so A is primitive.

Train of Thought

- When I saw the exponent n-1, I found it similar to lemma 8.4.1. So I guessed that A can be lowered or equaled to something like I+B.
- If render A=I+B, then B is not certainly nonnegative, so there must be a coefficient α on I, such that A=$\alpha(I+B)$.Since the equality is not always able to attain, we can replace the formular to $A \geq \alpha(I+B)$, or to make B simpler to represent,$A \geq \alpha\left(I+\frac{1}{\alpha}B\right)=\alpha I+B$. To ensure that $A-\alpha I$ is positive, we can let $\alpha=\min\{a_{11},\ldots,a_{nn}\}$, then B can be defined that $B=A-diag(a_{11},\ldots,a_{nn})$. Then the lemma can be proved by applying 8.4.1.

Proof

If every main diagonal entry of A is positive, let $\alpha = min\{a_{11}, ..., a_{nn}\} > 0$ and define $B = A - diag(a_{11}, ..., a_{nn})$. Then B is nonnegative and irreducible(because A is irreducible), and $A \geq \alpha I + B = \alpha(I + (1/\alpha)B)$. then (8.4.1) ensures that $A^{n-1} \geq \alpha_{n-1}(I + (1/\alpha)B)^{n-1} > 0$.

Exercise

- If $A \in M_n$ is nonnegative has positive diagonal entries, and if the i, j entry of A^m is positive, explain why the I, j entry of A^{m+p} is positive for each integer $p \geq 1$.
- Consider $(A^m)_{ij}A_{jj}$.

Lemma 8.5.5. Let $A \in M_n$ be nonnegative and primitive. Then A^m is nonnegative and primitive for every integer $m > 1$.

Proof. Since all sufficiently large powers of A are positive, the same is true for A^m for any m. If A^m were reducible, then A^{mp} would be reducible for all $p = 2, 3, ...$, and hence these matrices cannot be positive. This contradiction shows that no power of A can be reducible.

Theorem 8.5.6. Let $A \in M_n$ be nonnegative. If A is primitive, then $A^k > 0$ for some positive integer $k \leq (n-1)n^n$.

Train of Thought

- Since lemma 8.5.4. gives an upper bound of k in case of A with diagonal entries all positive, we can construct a power of A whose diagonal entries are all positive.
- We can consider the diagonal entries one by one. That is: construct a power of A with a positive 1,1 entry, then construct a power of the preceding matrix with a positive 2,2 entry. Continue this process until all the diagonal entries are positive.
- To make the result more accurate, we should make each exponent as small as possible.

Proof

Because A is irreducible, there is a directed path from the node P_1 in $\Gamma(A)$ back to itself;let k_1 be the **shortest** such path, so that $k_1 \le n$. The matrix A^{k_1} has a positive entry in its 1,1 position, and any power of A^{k_1} also has a positive 1,1 entry.Primitivity of A and (8.5.5) ensure that A^{k_1} is irreducible, so there is a directed path from the node P_2 in $\Gamma(Ak_1)$ back to itself;let $k_2 \le n$ be the length of the shortest such path. The matrix $(A^{k_1})^{k_2} = A^{k_1 k_2}$ has positive 1,1 and 2,2 entries.Continue this process down the main diagonal to obtain a matrix $A^{k_1 \cdots k_n}$ (with each $k_1 \le n$) that is irreducible and has positive diagonal entries.Lemma 8.5.4 ensures that $(A^{k_1 \cdots k_n})^{n-1} > 0$. Finally, observe that $k_1 ... k_n(n-1) \le n^n(n-1)$.

Notation

- If $A \in M_n$ is nonnegative and primitive, the least k such that $A^k > 0$ is the index of primitivity of A, which we denote by $\gamma(A)$.
- If there is at least one cycle in $\Gamma(A)$ has length less than s, we say that the shortest cycle in $\Gamma(A)$ has length s.

Theorem 8.5.7. Let $A \in M_n$ be nonnegative and primitive, and suppose that the shortest cycle in $\Gamma(A)$ has length s. Then $\gamma(A) \leq n + s(n-2)$, that is, $A^{n+s(n-2)} > 0$.

Train of Thought

- We rewrite $A^{n+s(n-2)}$ as $A^{n-s}(A^s)^{n-1}$, since we can then use the condition sufficiently.eg: each node contained in the circle has a loop(of length 1) in $\Gamma(A^s)$.
- Partition $A^s = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$ with $X_{11} \in M_s$ and $X_{22} \in M_{n-s}$.Partition $(A^s)^{n-1} = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}$ in the same way. We compute

$$A^{n-s}(A^s)^{n-1} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} \geq \begin{pmatrix} X_{11}Y_{11} & X_{11}Y_{12} \\ X_{21}Y_{11} & X_{21}Y_{12} \end{pmatrix}.$$

- Then we can explore the properties of X_{11}, X_{21}, Y_{11}, Y_{12}. For example, if we prove that each row of X_{11} and X_{21} has at least one nonzero entry and Y_{11} and Y_{12} are positive, then the theorem is proved.

Some Details

- Take the first row of X_{11} for example. If the ij entry is positive, then there exists a path of length n-s going from node i to node j in $\Gamma(A)$. To ensure that there exists a positive entry in this row, we just need to prove that there is a directed path in $\Gamma(A)$ of length n-s from the first node to an arbitrary one contained in X_{11}.
- In case of X_{12}, we can prove that there is a directed path in $\Gamma(A)$ of length n-s from a node out of X_{11} to an arbitrary one contained in X_{11}.
- To prove that Y_{11} and Y_{12} are positive, we can prove that for any $i, j \in \{1, \dots, n\}$ there is a directed path in $\Gamma(A^s)$ of length n-s from P_i to P_j. This can be proved in the similar way as shown above.

Proof: the Properties of X

Because A is irreducible, every node in $\Gamma(A)$ is contained in a cycle, and any shortest cycle has length at most n. We may assume that the distinct nodes inn a shortest cycle are $P_1, p_2, ..., P_s$. Notice that $n+s(n-2) = n-s+s(n-1)$ and consider $A^{n-s+s(n-1)} = A^{n-s}(A^s)^{n-1}$. Partition $A^{n-s} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ with $X_{11} \in M_s$ and $X_{22} \in M_{n-s}$.

Because the nodes $P_1, ..., P_s$ comprise a cycle in $\Gamma(A)$, for each positive integer m and any $i \in \{1, ..., s\}$, there is a directed path in $\Gamma(A)$ of length m from P_i to some P_j with $j \in \{1, ..., s\}$. In particluar, taking m=n-s, each row of $X_1 1$ must contain at least one positive entry.

For each $i \in \{s+1, ..., n\}$ there is a directed path in $\Gamma(A)$ of length $r \leq n-s$ (the number of nodes not in the cycle) from P_i (not in the cycle) to some node in the cycle. If $r < n-s$, one can go an additional n-s-r steps around the cycle to obtain a directed path in $\Gamma(A)$ of length exactly n-s from P_i to some node in the cycle. It follows that there is at least one nonzero entry in each row of X_{21}

Proof: the Properties of Y

Now partition $(A^s)^{n-1} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$ with $Y_{11} \in M_s$ and $Y_{22} \in M_{n-s}$. Because $P_1, ..., P_s$ comprise a cycle in $\Gamma(A)$, there is a loop at each node $P_1, ..., P_s$ in $\Gamma(A^s)$. Since A is primitive, A^s is also primitive, and hence it is irreducible. Therefore, for any $i, j \in \{1, ..., n\}$ there is a directed path in $\Gamma(A^s)$ of length at most n-1 from P_i to P_j. By first going a sufficient number of times around the loop at P_i, we can always construct such a path that has length exactly n-1. It follows that $Y_{11} > 0$ and $Y_{12} > 0$.

Proof: Final Steps

- To complete the argument, we compute

$$A^{n-s}(A^s)^{n-1} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} \geq \begin{pmatrix} X_{11}Y_{11} & X_{11}Y_{12} \\ X_{21}Y_{11} & X_{21}Y_{12} \end{pmatrix}$$

And use (8.1.14) to conclude that $A^{n-s}(A^s)^{n-1} > \mathbf{0}$.

Estimate s in the Preceding Theorem

Corollary 8.5.8(Wielandt). Let $A \in M_n$ be nonnegative. Then A is primitive if and only if $A^{n^2-2n+2} > 0$.

Proof

If some power of A is positive, then A is primitive, so only the converse implication is of interest.
If n=1, the result is trival, so assume that $n > 1$.
If A is primitive, then it is reducible and there are cycles in $\Gamma(A)$. If the shortest cycle in $\Gamma(A)$ has length n, then the length of every cycle in $\Gamma(A)$ is a multiple of n and (8.5.3) tells us that A cannot be primitive.
Thus, the length of the shortest cycle in $\Gamma(A)$ in n-1 or less, so (8.5.7) tells us that $\gamma(A) \leq n + s(n-2) \leq n + (n-1)(n-2) = n^2 - 2n + 2$.

Theorem 8.5.9. Let $A \in M_n$ be irreducible and nonnegative, and suppose that A has d positive main diagonal entries, $1 \leq d \leq n$. Then $A^{2n-d-1} > 0$; that is, $\gamma(A) \leq 2n - d - 1$.

Proof

Under the stated hypotheses, A must be primitive (strongly connected+loop+8.5.3), and $\Gamma(A)$ has d cycles with (maximum) length one.We may assume that $P_1, ..., P_d$ are the nodes in $\Gamma(A)$ that have loops. Consider $A^{2n-d-1} = A^{n-d}(A^1)^{n-1}$ and partition $A^{n-d} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ and $A^{n-1} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$, in which $X_{11}, Y_{11} \in M_d$ and $X_{22}, Y_{22} \in M_{n-d}$. The argument in the proof of (8.5.7) shows that each row of the blocks X_{11} and X_{21} contains at least one nonzero entry, the blocks Y_{11} and Y_{12} are positive, and $A^{n-d}A^{n-1}$ is positive.

8.6 A general Limit Theorem

Two Pre-required Exercises

Exercise 1. Let $\theta \in (0, 2\pi)$. Show that $(1 - e^{i\theta})\Sigma_{m=1}^{N} e^{im\theta} = e^{i\theta} - e^{i(N+1)\theta}$ and conclude that

$$\frac{1}{N}\Sigma_{m=1}^{N} e^{im\theta} = \frac{e^{i\theta} - e^{i(N+1)\theta}}{N(1 - e^{i\theta})} \to 0$$

as $N \to \infty$

Exercise 2. Let $B \in M_n$ and suppose that $\rho(B) < 1$. Show that $(I - B)\Sigma_{m=1}^{N} B^m = B - B^{N+1}$ and conclude that

$$\frac{1}{N}\Sigma_{m=1}^{N} B^m = \frac{1}{N}(B - B^{N+1})(I - B)^{-1} \to 0$$

as $N \to \infty$

Theorem 8.6.1. Let $A \in M_n$ be irreducible and nonnegative, let $n \geq 2$, and let x and y, respectively, be the right and left Perron vectors of A. Then

$$\lim_{N \to \infty} \frac{1}{N} \Sigma_{m=1}^{N} (\rho(A)^{-1} A)^m = xy^T$$

Moreover, there exists a finite positive constant $C = C(A)$ such that

$$\left\| \lim_{N \to \infty} \frac{1}{N} \Sigma_{m=1}^{N} (\rho(A)^{-1} A)^m - xy^T \right\|_{\infty} \leq \frac{C}{N}$$

for all $N = 1, 2, \ldots.$

Proof

If A is primitive, $\rho(A)^{-1}A$ can be factored as in (8.2.7a), in which x is the first column of S and y is the first column of S^{-1}. We have

$$\lim_{N\to\infty}\frac{1}{N}\Sigma_{m=1}^{N}(\rho(A)^{-1}A)^m = S\begin{bmatrix}1 & 0\\ 0 & \frac{1}{N}\Sigma_{m=1}^{N}B^m\end{bmatrix}S^{-1}$$

in which $\rho(B) < 1$, so the preceding exercise ensures that

$$\lim_{N\to\infty}\frac{1}{N}\Sigma_{m=1}^{N}(\rho(A)^{-1}A)^m \to S\begin{bmatrix}1 & 0\\ 0 & 0_{n-1}\end{bmatrix}S^{-1} = xy^T$$

Now suppose that A has exactly $k > 1$ eigenvalues of maximum modulus and let $\theta = 2\pi/k$. Corollary 8.4.6(c) ensures that the maximum-modulus eigenvalues of $\rho(A)^{-1}A$ are $1, e^{i\theta}, e^{2i\theta}, ..., e^{(k-1)i\theta}$, and each is s a simple eigenvalue. Thus, there is a nonsingular $S \in M_n$ such that x is its first column, y is its first column of S^{-1}, and

$$\rho(A)^{-1}A = S([1] \oplus [e^{i\theta}] \oplus ... \oplus [e^{(k-1)i\theta}] \oplus B)S^{-1}$$

in ehich $B \in M_{n-k}$ and $\rho(B) < 1$. The preceding exercises ensure that

$$\frac{1}{N}\Sigma_{m=1}N(\rho(A)^{-1}A)^m = S([1] \oplus [\lambda_{1,N}] \oplus ... \oplus [\lambda_{k-1,N}] \oplus B_N)S^{-1}$$

in which

$$\lambda_{1,N} = \frac{e^{i\theta} - e^{i(N+1)\theta}}{N(1-e^{i\theta})} \to 0$$

$$\vdots$$

$$\lambda_{k-1,N} = \frac{e^{i(k-1)\theta} - e^{i(N+1)(k-1)\theta}}{N(1-e^{i(k-1)\theta})} \to 0$$

$$B_N = \frac{1}{N}(B - B^{N+1})(I-B)^{-1} \to 0$$

as $N \to \infty$
Therefore,

$$\frac{1}{N}\Sigma_{m=1}N(\rho(A)^{-1}A)^m \to S([1] \oplus [0] \oplus ... \oplus [0] \oplus 0_{n-k})S^{-1} = xy^T$$

The bound in (8.6.3) is revealed in the representation

$$\frac{1}{N}\Sigma_{m=1}^{N}(\rho(A)^{-1}A)^m - xy^T$$
$$= S([1] \oplus [\lambda_{1,N}] \oplus ... \oplus [\lambda_{k-1,N}] \oplus B_N)S^{-1} - S([1] \oplus 0^{n-1})S^{-1}$$
$$= S([0] \oplus [\lambda_{1,N}] \oplus ... \oplus [\lambda_{k-1,N}] \oplus B_N)S^{-1}$$
$$= \frac{1}{N}S([0] \oplus [N\lambda_{1,N}] \oplus ... \oplus [N\lambda_{k-1,N}] \oplus NB_N)S^{-1}$$

The preceding identities ensure that the matrix factor is bounded as $N \to \infty$.

8.7 Stochastic and Doubly Stochastic Matrices

Definition

- Row stochastic matrix: each row sums up to 1
- Column stochastic matrix: each column sums up to 1
- Doubly stochastic matrix: both row and column stochastic matrix
- The set of stochastic matrices in M_n is a compact set that is also convex.

Exercise

Suppose that $n \geq 2, A = [a_{ij}] \in M_n$ is doubly stochastic, and some $a_{ii} = 1$. Explain why (a) $a_{ki} = a_{ik} = 0$ for all $k \in \{1, 2, ..., n\}$ such that $k \neq i$; (b) A is permutation similar to $[1] \oplus B$, in which B is doubly stochastic; (c) the main diagonal entries of B are obtained from the main diagonal entries of A by removing one entry equal to $+1$; and (d) the characteristic polynomials of A and B are related by the identity $p_A(t) = (t-1)p_B(t)$.

Lemma 8.7.1. Let $A = [a_{ij}] \in M_n$ be a doubly stochastic matrix that is not the identity matrix. There is a permutation of $\{1, ..., n\}$ that is not the identity permutation and is such that $a_{1\sigma(1)}...a_{n\sigma(n)} > 0$.

Proof

Proof. Suppose that every permutation σ of $\{1, ..., n\}$ that is not the identity permutation σ_0 has the property that $a_{1\sigma(1)}...a_{n\sigma(n)} = 0$. This assumption and (0.3.2.1) permit us to compute the characteristic polynomial of A:

$$\begin{aligned} p_A(t) &= det(tI - A) = \Pi_{i=1}^{n}(t - a_{ii}) + \Sigma_{\sigma \neq \sigma_0}(sgn\sigma \Pi_{i=1}^{n}(-a_{i\sigma(i)})) \\ &= \Pi_{i=1}^{n}(t - a_{ii}) \end{aligned}$$

It follows that the main diagonal entries of A are its eigenvalues. Since +1 is an eigenvalue of A, at least one of its main diagonal entries is +1. The preceding exercise ensures that A is permutation similar to $[1] \oplus B$, in which $B = [b_{ij} \in M_{n-1}$ is doubly stochastic; its main diagonal entries are obtained from the main diagonal entries of A by omitting one +1 entry; +1 is an eigenvalue of B; and $p_B(t) = p_A(t)/(t-1) = \Pi_{i=1}^{n-1}(t - b_{ii})$. Applying the preceding argument to B shows that some $b_{ii} = 1$, so at least two main diagonal entries of A are +1. Continuing in this way, after at most n-1 steps, we comclude that every main diagonal entry of A is +1, so A=I. This contradiction shows that some product $a_{1\sigma(1)}...a_{n\sigma(n)}$ must be positive.

Theorem 8.7.2(Birkhoff). A matrix $A \in M_n$ is doubly stochastic if and only if there are permutation matrices $P_1, ..., P_N \in M_n$ and positive scalars $t_1, ..., t_N \in R$ such that $t_1 + ... + t_N = 1$ and

$$A = t_1 P_1 + ... + t_N P_N$$

Moreover, $N \leq n^2 - n + 1$.

Proof

The sufficiency is clear.

Choose $a_{1\sigma(1)}...a_{n\sigma(n)} > 0$.Let α_1 be their minimum.The corresponding permutation matrix is P_1.

$\alpha_1 = 1$:trival.

Suppose $\alpha_1 < 1$.Let $A_1 = (1-\alpha_1)^{-1}(A_1 - \alpha_1 P_1)$.Then A_1 has at least one zero entry than A.

$A = (1-\alpha_1)A_1 + \alpha_1 P_1$. Continue to iterate, until A_k is a permutation matrix. Since A_k has at least k zero entries, and since a doubly stochastic matrix has at most $n^2 - n$ zero entries, we can have at most $n^2 - n$ iterations and at most $n^2 - n + 1$ summands.

Corollary 8.7.4. The maximum (respectively, minimum) of a convex (respectively, concave) real-valued function on the set of doubly stochastic n-by-n matrices is attained at a permutation matrix.

Proof

In case of convex function:

$$
\begin{aligned}
f(A) &= f(t_1 P_1 + ... + t_N P_N) \leq t_1 f(P_1) + ... + t_N f(P_N) \\
&\leq t_1 f(P_k) + ... + t_N f(P_k) = (t_1 + ... + t_N) f(P_k) = f(P_k)
\end{aligned}
$$

Since f achieves its maximum at A, we have $f(A) = f(P_k)$.

A nonnegative matrix $A \in M_n$ is doubly substochastic if $Ae \leq e$ and $e^T A \leq e^T$, that is, all row and column sums are at most one.

Lemma 8.7.5. Let $A \in M_n$ be doubly substochastic. There is a doubly stochastic matrix $S \in M_n$ such that $A \leq S$.

Proof

- Let N(S) denote the number of row sums and column sums of S that are less than one.
- Use **the method of finite descent** : construct a greater matrix with lower N value to accomplish the proof.
- Consider a column and a row with their sums less than 1. Enlarge the crossing entry until at least one sum becomes 1. Then N is decreased and the matrix is greater than the original one.
- Repeat until N(S)=0, which means that S is a doubly stochastic matrix.

Exercise

- Let $U \in [u_{ij}], V = [v_{ij}] \in M_n$ be unitary and let $S = [|u_{ij}v_{ji}|]$. Show that S is doubly substochastic.
- Use the Cauchy-Schwarz inequality.

Theorem 8.7.6 (von Neumann). Let the ordered singular values of $A, B \in M_n$ be $\sigma_1(A) \geq ... \geq \sigma_n(A)$ and $\sigma_1(B) \geq ... \geq \sigma_n(B)$. Then

$$Re\ tr(AB) \leq \Sigma_{i=1}^{n} \sigma_i(A)\sigma_i(B)$$

Proof

Let $A = V_1 \Sigma_A W_1*$ and $B = V_2 \Sigma_B W_2*$ be singular value deconpositions, in which $V_!, W_1, V_2, W_2 \in M_n$ are unitary, $\Sigma_A = diag(\sigma_1(A), ..., \sigma_n(A))$, and $\Sigma_B = diag(\sigma_1(B), ..., \sigma_n(B))$. Let $U = W_1 * V_2 = [u_{ij}]$ and $V = W_2 * V_1 = [v_{ij}]$. Then

$$\begin{aligned} Re\ tr(AB) &= Re\ tr(V_1 \Sigma_A W_1 * V_2 \Sigma_B W_2 *) \\ &= Re\ tr(\Sigma_A W_1 * V_2 \Sigma_B W_2 * V_1) = Re\ tr(\Sigma_A U \Sigma_B V) \\ &= Re \Sigma_{i,j=1}^n \sigma_i(A) \sigma_j(B) u_{ij} v_{ji} = \Sigma_{i,j=1}^n \sigma_i(A) \sigma_j(B) Re(u_{ij} v_{ji}) \\ &\leq \Sigma_{i,j=1}^n \sigma_i(A) \sigma_j(B) |u_{ij} v_{ji}| \end{aligned}$$

Proof

The preceding exercise tells us that the matrix $[|u_{ij}v_{ji}|$ is doubly substochastic, and (8.7.5) ensures that there is a doubly stochastic matrix C such that $[|u_{ij}v_{ji}| \leq C = [c_{ij}]$. Therefore,

$$\begin{aligned} Re\ tr(AB) &\leq \Sigma_{i,j=1}^{n} \sigma_i(A)\sigma_j(B)c_{ij} \\ &\leq max\{\Sigma_{i,j=1}^{n} \sigma_i(A)\sigma_j(B)s_{ij} : S = [s_{ij}]\ is\ doubly\ stochastic\} \end{aligned}$$

Proof

The function $f(S) = \Sigma^n_{i,j=1}\sigma_i(A)\sigma_j(B)s_{ij}$ is a linear (and therefore convex) function on the set of doubly stochastic matrices, so (8.1.4) tells us that it attains its maximum at a permutation matrix $P = [p_{ij}]$. If π is the permutation of $\{1, ..., n\}$ such that $p_{ij} = 1$ if and only if $j = \pi(i)$, then

$$\begin{aligned} Re\ tr(AB) &\leq \Sigma^n_{i,j=1}\sigma_i(A)\sigma_j(B)p_{ij} = \Sigma^n_{i=1}\sigma_i(A)\sigma_{\pi(i)}(B) \\ &\leq \Sigma^n_{i=1}\sigma_i(A)\sigma_i(B) \end{aligned}$$

Thanks