APPENDIX

1

Proof of Theorem 1

Here is a restatement of Theorem 1, which will be proven in this appendix:

THEOREM 1

If P is a regular $m \times m$ transition matrix with $m \geq 2$, then the following statements are all true.

a. There is a stochastic matrix Π such that $\lim_{n\to\infty} P^n = \Pi$.

b. Each column of Π is the same probability vector $\mathbf{q}$.

c. For any initial probability vector $\mathbf{x}_0$, $\lim_{n\to\infty} P^n\mathbf{x}_0 = \mathbf{q}$.

d. The vector $\mathbf{q}$ is the unique probability vector which is an eigenvector of P associated with the eigenvalue 1.

e. All eigenvalues λ of P other than 1 have $|\lambda| < 1$.

The proof of Theorem 1 requires creation of an order relation for vectors, and begins with the consideration of matrices whose entries are strictly positive or non-negative.

DEFINITION

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^m$, then

a. $\mathbf{x} > \mathbf{y}$ if $x_i > y_i$ for $i = 1, 2, \ldots, m$.

b. $\mathbf{x} < \mathbf{y}$ if $x_i < y_i$ for $i = 1, 2, \ldots, m$.

c. $\mathbf{x} \geq \mathbf{y}$ if $x_i \geq y_i$ for $i = 1, 2, \ldots, m$.

d. $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for $i = 1, 2, \ldots, m$.

DEFINITION

An $m \times n$ matrix A is **positive** if all its entries are positive. An $m \times n$ matrix A is **non-negative** if it has no negative entries.

Notice that all stochastic matrices are non-negative. The row-vector rule (Section 1.3) shows that multiplication of vectors by a positive matrix preserves order.

$$\text{If } A \text{ is a positive matrix and } \mathbf{x} > \mathbf{y}, \text{ then } A\mathbf{x} > A\mathbf{y}. \tag{1}$$

$$\text{If } A \text{ is a positive matrix and } \mathbf{x} \geq \mathbf{y}, \text{ then } A\mathbf{x} \geq A\mathbf{y}. \tag{2}$$

In addition, multiplication by non-negative matrices "almost" preserves order in the following sense.

$$\text{If } A \text{ is a non-negative matrix and } \mathbf{x} \geq \mathbf{y}, \text{ then } A\mathbf{x} \geq A\mathbf{y}. \tag{3}$$

The first step toward proving Theorem 1 is a lemma which shows how the transpose of a stochastic matrix acts on a vector.

LEMMA 1 Let P be an $m \times m$ stochastic matrix, and let ϵ be the smallest entry in P. Let $\mathbf{a}$ be in $\mathbb{R}^m$; let M_a be the largest entry in $\mathbf{a}$, and let m_a be the smallest entry in $\mathbf{a}$. Likewise, let $\mathbf{b} = P^T\mathbf{a}$, let M_b be the largest entry in $\mathbf{b}$, and let m_b be the smallest entry in $\mathbf{b}$. Then $m_a \leq m_b \leq M_b \leq M_a$ and

$$M_b - m_b \leq (1 - 2\epsilon)(M_a - m_a)$$

PROOF Create a new vector $\mathbf{c}$ from $\mathbf{a}$ by replacing every entry of $\mathbf{a}$ by M_a except for one occurrence of m_a. Suppose that this single m_a entry lies in the i^{th} row of $\mathbf{c}$. Then $\mathbf{c} \geq \mathbf{a}$. If the columns of P^T are labeled $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m$, we have

$$\begin{aligned} P^T\mathbf{c} &= \sum_{k=1}^{m} c_k\mathbf{q}_k \\ &= \sum_{k=1}^{m} M_a\mathbf{q}_k - M_a\mathbf{q}_i + m_a\mathbf{q}_i \end{aligned}$$

Since P is a stochastic matrix, each row of P^T sums to 1. If we let $\mathbf{u}$ be the vector in $\mathbb{R}^m$ consisting of all 1's, then $\displaystyle\sum_{k=1}^{m} M_a\mathbf{q}_k = M_a \sum_{k=1}^{m} \mathbf{q}_k = M_a\mathbf{u}$, and

$$\sum_{k=1}^{m} M_a\mathbf{q}_k - M_a\mathbf{q}_i + m_a\mathbf{q}_i = M_a\mathbf{u} - (M_a - m_a)\mathbf{q}_i$$

Since each entry in P (and thus P^T) is greater than or equal to ϵ, $\mathbf{q}_i \geq \epsilon\mathbf{u}$, and

$$M_a\mathbf{u} - (M_a - m_a)\mathbf{q}_i \leq M_a\mathbf{u} - \epsilon(M_a - m_a)\mathbf{u} = (M_a - \epsilon(M_a - m_a))\mathbf{u}$$

So

$$P^T\mathbf{c} \leq (M_a - \epsilon(M_a - m_a))\mathbf{u}$$

But since $\mathbf{a} \geq \mathbf{c}$ and P^T is non-negative, Equation (3) gives

$$\mathbf{b} = P^T\mathbf{a} \leq P^T\mathbf{c} \leq (M_a - \epsilon(M_a - m_a))\mathbf{u}$$

Thus each entry in $\mathbf{b}$ is less than or equal to $M_a - \epsilon(M_a - m_a)$. In particular,

$$M_b \leq M_a - \epsilon(M_a - m_a) \tag{4}$$

So $M_b \leq M_a$. If we now examine the vector $-\mathbf{a}$, we find that the largest entry in $-\mathbf{a}$ is $-m_a$, the smallest is $-M_a$, and similar results hold for $-\mathbf{b} = P^T(-\mathbf{a})$. Applying Equation (4) to this situation gives

$$-m_b \leq -m_a - \epsilon(-m_a + M_a) \tag{5}$$

so $m_b \geq m_a$. Adding Equations (4) and (5) together gives

$$\begin{aligned} M_b - m_b &\leq M_a - m_a - 2\epsilon(M_a - m_a) \\ &= (1 - 2\epsilon)(M_a - m_a) \end{aligned}$$ ■

Proof of Theorem 1 First assume that the transition matrix P is a **positive** stochastic matrix. As above, let $\epsilon > 0$ be the smallest entry in P. Consider the vector $\mathbf{e}_j$ where $1 \leq j \leq m$. Let M_n and m_n be the largest and smallest entries in the vector $(P^T)^n\mathbf{e}_j$. Since $(P^T)^n\mathbf{e}_j = P^T(P^T)^{n-1}\mathbf{e}_j$, Lemma 1 gives

$$M_n - m_n \leq (1-2\epsilon)(M_{n-1} - m_{n-1}) \tag{6}$$

Hence, by induction, it may be shown that

$$M_n - m_n \leq (1-2\epsilon)^n(M_0 - m_0) = (1-2\epsilon)^n$$

Since $m \geq 2$, $0 < \epsilon \leq 1/2$. Thus $0 \leq 1 - 2\epsilon < 1$, and $\lim_{n\to\infty}(M_n - m_n) = 0$. Therefore the entries in the vector $(P^T)^n\mathbf{e}_j$ approach the same value, say q_j, as n increases. Notice that since the entries in P^T are between 0 and 1, the entries in $(P^T)^n\mathbf{e}_j$ must also be between 0 and 1, and so q_j must also lie between 0 and 1. Now $(P^T)^n\mathbf{e}_j$ is the j^{th} column of $(P^T)^n$, which is the j^{th} row of P^n. Therefore P^n approaches a matrix all of whose rows are constant vectors, which is another way of saying the columns of P^n approach the same vector $\mathbf{q}$:

$$\lim_{n\to\infty} P^n = \Pi = \begin{bmatrix} \mathbf{q} & \mathbf{q} & \cdots & \mathbf{q} \end{bmatrix} = \begin{bmatrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & \ddots & \vdots \\ q_m & q_m & \cdots & q_m \end{bmatrix}$$

So Theorem 1(a) is true if P is a positive matrix. Suppose now that P is regular but not positive; since P is regular, there is a power P^k of P that is positive. We need to show that $\lim_{n\to\infty}(M_n - m_n) = 0$; the remainder of the proof follows exactly as above. No matter the value of n, there is always a multiple of k, say rk, with $rk < n \leq r(k+1)$. By the proof above, $\lim_{r\to\infty}(M_{rk} - m_{rk}) = 0$. But Equation (6) applies equally well to non-negative matrices, so $0 \leq M_n - m_n \leq M_{rk} - m_{rk}$, and $\lim_{n\to\infty} M_n - m_n = 0$, proving part (a) of Theorem 1.

To prove part (b), it suffices to show that $\mathbf{q}$ is a probability vector. To see this, note that since $(P^T)^n$ has row sums equal to 1 for any n, $(P^T)^n\mathbf{u} = \mathbf{u}$. Since $\lim_{n\to\infty}(P^T)^n = \Pi^T$, it must be the case that $\Pi^T\mathbf{u} = \mathbf{u}$. Thus the rows of Π^T, and so also the columns of Π, must sum to 1 and $\mathbf{q}$ is a probability vector.

The proof of part (c) follows from the definition of matrix multiplication and the fact that P^n approaches Π by part (a). Let $\mathbf{x}_0$ be any probability vector. Then

$$\begin{aligned}
\lim_{n\to\infty} P^n\mathbf{x}_0 &= \lim_{n\to\infty} P^n(x_1\mathbf{e}_1 + \ldots + x_m\mathbf{e}_m) \\
&= x_1(\lim_{n\to\infty} P^n\mathbf{e}_1) + \ldots + x_m(\lim_{n\to\infty} P^n\mathbf{e}_m) \\
&= x_1(\Pi\mathbf{e}_1) + \ldots + x_m(\Pi\mathbf{e}_m) = x_1\mathbf{q} + \ldots + x_m\mathbf{q} \\
&= (x_1 + \ldots + x_m)\mathbf{q} = \mathbf{q}
\end{aligned}$$

since the entries in $\mathbf{x}_0$ sum to 1.

To prove part (d), we calculate $P\Pi$. First note that $\lim_{n\to\infty} P^{n+1} = \Pi$. But since $P^{n+1} = PP^n$, and $\lim_{n\to\infty} P^n = \Pi$, $\lim_{n\to\infty} P^{n+1} = P\Pi$. Thus $P\Pi = \Pi$, and any column of this matrix equation gives $P\mathbf{q} = \mathbf{q}$. Thus $\mathbf{q}$ is a probability vector that is also an eigenvector for P associated with the eigenvalue $\lambda = 1$. To show that this vector is unique, let $\mathbf{v}$ be any eigenvector for P associated with the eigenvalue $\lambda = 1$, which is also a probability vector. Then $P\mathbf{v} = \mathbf{v}$, and $P^n\mathbf{v} = \mathbf{v}$ for any n. But by part (c), $\lim_{n\to\infty} P^n\mathbf{v} = \mathbf{q}$, which can happen only if $\mathbf{v} = \mathbf{q}$. Thus $\mathbf{q}$ is unique. Note that this part of the proof has also shown that the eigenspace associated with the eigenvalue $\lambda = 1$ has dimension 1.

To prove part (e), let $\lambda \neq 1$ be an eigenvalue of P, and let $\mathbf{x}$ be an associated eigenvector. Assume that $\sum_{k=1}^{m} x_k \neq 0$. Since any nonzero scalar multiple of $\mathbf{x}$ will also be an eigenvector associated with λ, we may scale the eigenvector $\mathbf{x}$ by the reciprocal of $\sum_{k=1}^{m} x_k$ to form the eigenvector $\mathbf{w}$. Notice that the sum of the entries in $\mathbf{w}$ is 1. Then $P\mathbf{w} = \lambda\mathbf{w}$, so $P^n\mathbf{w} = \lambda^n\mathbf{w}$ for any n. By the proof of part (c), $\lim_{n\to\infty} P^n\mathbf{w} = \mathbf{q}$ since the entries in $\mathbf{w}$ sum to 1. Thus

$$\lim_{n\to\infty} \lambda^n\mathbf{w} = \mathbf{q} \tag{7}$$

Notice that Equation (7) can be true only if $\lambda = 1$. If $|\lambda| \geq 1$ and $\lambda \neq 1$, the left side of Equation (7) diverges; if $|\lambda| < 1$, the left side of Equation (7) must converge to $\mathbf{0} \neq \mathbf{q}$. This contradicts our assumption, so it must be the case that $\sum_{k=1}^{m} w_k = 0$. By part (a), $\lim_{n\to\infty} P^n\mathbf{w} = \Pi\mathbf{w}$. Since

$$\begin{aligned} \Pi\mathbf{w} &= \begin{bmatrix} \mathbf{q} & \mathbf{q} & \cdots & \mathbf{q} \end{bmatrix}\mathbf{w} \\ &= w_1\mathbf{q} + w_2\mathbf{q} + \cdots + w_m\mathbf{q} \\ &= (w_1 + w_2 + \cdots + w_m)\mathbf{q} = 0\mathbf{q} = \mathbf{0} \end{aligned}$$

then $\lim_{n\to\infty} P^n\mathbf{w} = \mathbf{0}$. Since $P^n\mathbf{w} = \lambda^n\mathbf{w}$ and $\mathbf{w} \neq \mathbf{0}$, $\lim_{n\to\infty} \lambda^n = 0$, and $|\lambda| < 1$. ■