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 $\Pi$ such that $\lim_{n \to \infty} P^n = \Pi$.
b. Each column of $\Pi$ is the same probability vector $q$.
c. For any initial probability vector $x_0$, $\lim_{n \to \infty} P^n x_0 = q$.
d. The vector $q$ is the unique probability vector which is an eigenvector of $P$ associated with the eigenvalue 1.
e. All eigenvalues $\lambda$ 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 $x$ and $y$ are in $\mathbb{R}^m$, then

a. $x > y$ if $x_i > y_i$ for $i = 1, 2, \ldots, m$.
b. $x < y$ if $x_i < y_i$ for $i = 1, 2, \ldots, m$.
c. $x \geq y$ if $x_i \geq y_i$ for $i = 1, 2, \ldots, m$.
d. $x \leq y$ 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:

If $A$ is a positive matrix and $x > y$, then $Ax > Ay$. \hfill (1)

If $A$ is a positive matrix and $x \geq y$, then $Ax \geq Ay$. \hfill (2)
In addition, multiplication by non-negative matrices “almost” preserves order in the following sense.

\[
\text{If } A \text{ is a non-negative matrix and } x \succeq y, \text{ then } Ax \succeq Ay. \tag{3}
\]

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 \( \epsilon \) 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 \( \mathbf{c} \succeq \mathbf{a} \). If the columns of \( P^T \) are labeled \( q_1, q_2, \ldots, q_m \), we have

\[
P^T \mathbf{c} = \sum_{k=1}^{m} c_k q_k
\]

\[
= \sum_{k=1}^{m} M_a q_k - M_a q_i + m_a q_i
\]

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

\[
\sum_{k=1}^{m} M_a q_k = M_a \sum_{k=1}^{m} q_k = M_a \mathbf{u}, \text{ and}
\]

\[
\sum_{k=1}^{m} M_a q_k - M_a q_i + m_a q_i = M_a \mathbf{u} - (M_a - m_a)q_i
\]

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

\[
M_a \mathbf{u} - (M_a - m_a)q_i \leq M_a \mathbf{u} - \epsilon(M_a - m_a)u = (M_a - \epsilon(M_a - m_a))\mathbf{u}
\]

So

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

But since \( \mathbf{a} \succeq \mathbf{c} \) 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

\[
-m_b \leq -m_a - \epsilon(-m_a + M_a)
\]

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

\[
M_b - m_b \leq M_a - m_a - 2\epsilon(M_a - m_a)
\]

\[
= (1 - 2\epsilon)(M_a - m_a)
\]

\[\square\]
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 \( e_j \) where \( 1 \leq j \leq m \). Let \( M_n \) and \( m_n \) be the largest and smallest entries in the vector \((P^T)^ne_j\). Since \((P^T)^ne_j = P^T((P^T)^{n-1}e_j)\), Lemma 1 gives

\[
M_n - m_n \leq (1 - 2\epsilon)(M_{n-1} - m_{n-1})
\]

(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 \((P^T)^ne_j\) 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 \((P^T)^ne_j\) must also be between 0 and 1, and so \( q_j \) must also lie between 0 and 1. Now \((P^T)^ne_j\) 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 \( q \):

\[
\lim_{n \to \infty} P^n = \Pi = \begin{bmatrix} q & q & \cdots & q \\ q & q & \cdots & q \\ \vdots & \vdots & \ddots & \vdots \\ q & q & \cdots & q \end{bmatrix}
\]

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, \( \lim_{n \to \infty} (M_{rk} - m_{rk}) = 0 \). 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 \( q \) is a probability vector. To see this, note that since \((P^T)^ne_j\) has row sums equal to 1 for any \( n \), \((P^T)^nu = u \). Since \( \lim_{n \to \infty} (P^T)^nu = \Pi^T \), it must be the case that \( \Pi^Tu = u \). Thus the rows of \( \Pi^T \), and so also the columns of \( \Pi \), must sum to 1 and \( q \) is a probability vector.

The proof of part (c) follows from the definition of matrix multiplication and the fact that \( P^n \) approaches \( \Pi \) by part (a). Let \( x_0 \) be any probability vector. Then

\[
\lim_{n \to \infty} P^n x_0 = \lim_{n \to \infty} P^n(x_1e_1 + \ldots + x_me_m) \\
= x_1(\lim_{n \to \infty} P^n e_1) + \ldots + x_m(\lim_{n \to \infty} P^n e_m) \\
= x_1(\Pi e_1) + \ldots + x_m(\Pi e_m) = x_1q + \ldots + x_mq \\
= (x_1 + \ldots + x_m)q = q
\]

since the entries in \( x_0 \) 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 \( Pq = q \). Thus \( q \) 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 \( v \) be any eigenvector for \( P \) associated with the eigenvalue \( \lambda = 1 \), which is also a probability vector. Then \( Pv = v \), and \( P^n v = v \) for any \( n \). But by part (c), 
\[
\lim_{n \to \infty} P^n v = q,
\]
which can happen only if \( v = q \). Thus \( q \) 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 \( x \) be an associated eigenvector. Assume that \( \sum_{k=1}^{m} x_k \neq 0 \). Since any nonzero scalar multiple of \( x \) will also be an eigenvector associated with \( \lambda \), we may scale the eigenvector \( x \) by the reciprocal of \( \sum_{k=1}^{m} x_k \) to form the eigenvector \( w \). Notice that the sum of the entries in \( w \) is 1. Then \( Pw = \lambda w \), so \( P^n w = \lambda^n w \) for any \( n \). By the proof of part (c), 
\[
\lim_{n \to \infty} P^n w = q
\]
since the entries in \( w \) sum to 1. Thus
\[
\lim_{n \to \infty} \lambda^n w = q
\]
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 \( 0 \neq q \). This contradicts our assumption, so it must be the case that \( \sum_{k=1}^{m} w_k = 0 \). By part (a), 
\[
\lim_{n \to \infty} P^n w = \Pi w.
\]
Since
\[
\Pi w = [q \quad q \quad \cdots \quad q]w
\]
\[
= w_1 q + w_2 q + \cdots + w_m q
\]
\[
= (w_1 + w_2 + \cdots + w_m)q = 0q = 0
\]
then \( \lim_{n \to \infty} P^n w = 0 \). Since \( P^n w = \lambda^n w \) and \( w \neq 0 \), \( \lim_{n \to \infty} \lambda^n = 0 \), and \( |\lambda| < 1 \). ■