DEDICATION

To my parents

Georgia and Anastasio

Στούς γονεῖς μου

Γεωργία καί 'Αναστάσιο

ACKNOWLEDGEMENTS

The author would like to express his gratitude to his thesis supervisor, Dr. Abraham Boyarsky of Concordia University. Dr. Boyarsky's encouragement and guidance were most helpful.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

CHAPTER I

INTRODUCTION

1.1. MARKOV MAPS

A function f is said to be piecewise continuous on an interval $I = [a,b]$ if there exists a finite partition $\mathcal{J} = \{I_i\}_{i=1}^{n}$, $I_i = (a_{i-1}, a_i)$, $a_0 = a$, $a_n = b$ such that:

(i) f is continuous on each open subinterval I_i, $1 \leq i \leq n$;

(ii) f approaches a finite limit as the end points of each subinterval are approached from within the subinterval.

In other words, f is piecewise continuous on $[a,b]$ if it is continuous, except for a finite number of jump discontinuities. A simple example of a piecewise continuous function is the following:

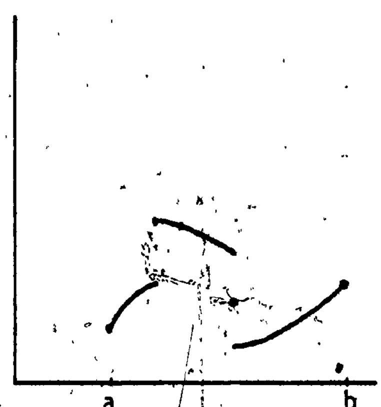


Fig. 1.1

Moreover, f is said to be piecewise C^k if for each i, $1 \leq i \leq n$, the restriction f_i of f on the open interval (a_{i-1}, a_i) is a C^k

function[1] which can be extended to the closed interval $[a_{i-1}, a_i]$, f need not be continuous at the partition points a_i.

In the following we shall use the terminology "map" or "transformation", instead of a function.

<u>Definition 1.1.1</u>: Let $I = [a,b]$ be an interval on the real line and let us consider a finite partition $\mathcal{J}$ of I. Let us denote by Q the partition points of $\mathcal{J}$. We say that a piecewise continuous transformation $\tau: I \to I$ takes partition points into partition points if $\tau(Q) \subset Q$. If τ is discontinuous at the partition points a_i, we require $\tau(a_i^-)$ and $\tau(a_i^+)$ to be in Q.

The above partition $\mathcal{J}$ is called a <u>Markov partition</u> for τ, and the transformation τ is called a <u>Markov map</u> with respect to the partition $\mathcal{J}$.

<u>Definition 1.1.2</u>: Let $I_i = (a_{i-1}, a_i)$, $1 \le i \le n$. We say that the partition $\mathcal{J}$ has the <u>communication property</u> ("irreducibility property" in Markov chain terminology), under the piecewise continuous transformation $\tau: I \to I$, if for any $I_i, I_j \in \mathcal{J}$ there exist integers p and q such that:

$$I_i \subset \tau^p(I_j) \quad \text{and} \quad I_j \subset \tau^q(I_i),$$

where τ^k is the kth iterate of τ.

Let $I = [a,b]$ and let $\tau: I \to I$ be a measurable non-singular transformation. By "non-singular" we mean that $m(\tau^{-1}(A)) = 0$ whenever $m(A) = 0$ for A a measurable set (m denotes the Lebesgue

---

[1] A continuous function with continuous derivatives up to kth order.

measure on I and $\tau^{-1}(A) = \{x \in I : \tau(x) \in A\}$). Also let μ be a measure defined on the σ-algebra of I. We say that μ is invariant under τ if for all measurable sets $A \subset I$ we have $\mu(A) = \mu(\tau^{-1}(A))$. μ is absolutely continuous if there exists a function $f: I \to [0,\infty)$, $f \in L_1 = L_1(I)$, the space of integrable functions on I, such that:

$$\mu(A) = \int_A f(x)dx \quad ,$$

for every Lebesque measurable set $A \subset I$. We refer to f as the invariant density (of μ) under τ.

It is known [14] that if $\inf_{x \in I} |\tau'(x)| > 1$, and τ is a piecewise C^2 transformation, then it admits an absolutely continuous invariant measure. It is also known [3] that a certain subclass of the aforementioned class possesses a unique absolutely continuous invariant measure. Moreover, [14], the invariant density f (of μ) under τ is a fixed point of the Frobenius-Perron operator $P_\tau : L_1 \to L_1$ defined by:

$$P_\tau f(x) = \frac{d}{dx} \int_{\tau^{-1}([0,x])} f(s)ds.$$

On the other hand, it is known [10] that the fixed points of the Frobenius-Perron operator P_τ are piecewise constant functions. Therefore, the unique invariant density under τ, for τ belonging to that particular subclass, is a piecewise constant function.

Generally, it is very difficult to find the unique invariant density for τ. In this thesis, we shall show that in the case where τ is piecewise linear, the unique invariant density under τ, is found by simply solving a system of linear equations. Moreover, we shall find the invariant densities of some particular piecewise linear transformations and we shall characterize some classes of piecewise constant functions, which can serve as the unique invariant densities for constructible linear transformations.

Remark: Since the results of the theorems remain valid when we use linear transformations (i) from [a,b] into [a,b], [a,b] any interval on the real line and (ii) from [0,1] into [0,1], we shall use, mostly, linear transformations from [0,1] into [0,1].

Let us now give some matrix theoretic definitions.

Definition 1.1.3: A square matrix $A = (a_{ij})$, $(i,j = 1,2,\ldots,n)$, is called non-negative, $A \geq 0$ (or positive, $A > 0$) if all the entries of A are non-negative (or positive).

Definition 1.1.4: A square matrix $A = (a_{ij})$, $(i,j = 1,2,\ldots,n, n \geq 2)$ is called reducible if there is a permutation[1] matrix P such that

$$P^T A P = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}.$$

where A_{11}, A_{22} are square matrices of order less than n, and T denotes transpose. Otherwise, A is called irreducible.

Definition 1.1.5: A $n \times n$ square non-negative matrix A is primitive if there exists an integer $k > 0$, such that $A^k > 0$, where $A^k = (a_{ij}^{(k)})$ denotes the kth power of A.

[1] Permutation matrix is a square matrix where the entries are either zero or one and which has precisely a single 1 in each row and in each column.

Definition 1.1.6: A $n \times n$ square non-negative matrix A is said to be row stochastic if each row sum is equal to 1, i.e.,

$$\sum_{j=1}^{n} a_{ij} = 1 \ , \quad i = 1,2,\ldots,n.$$

Definition 1.1.7: A non-negative matrix A is contiguous if all the positive entries in each row are equal and there are no zero entries between them.

1.2. OUTLINE OF THE THESIS

In Chapter II, we explore the property that the Frobenius-Perron operator P_τ, when restricted to piecewise constant functions, can be represented as a product of a vector and of an $n \times n$ matrix M_τ. The matrix M_τ is referent as the matrix induced by the transformation τ. We define a certain class of transformations $\mathcal{T}$, and study some properties of M_τ for $\tau \in \mathcal{T}$. Finally we prove a theorem (Theorem 2.2.1) which gives necessary and sufficient conditions for an irreducible matrix to be primitive. This result will be used frequently in Chapter IV.

We start Chapter III by proving the famous Perron-Frobenius Theorem (Theorem 3.1.1) and using that we show (Theorem 3.2.1) that the matrix M_τ, induced by a piecewise linear transformation $\tau \in \mathcal{T}$ has 1 as the eigenvalue of maximum absolute value. This permits us to show (Theorem 3.2.2) that it is possible to construct the unique invariant density for a piecewise linear transformation $\tau \in \mathcal{T}$ when τ satisfies two simple conditions. Moreover, we shall show that it is possible to construct a transformation τ under which a given function is invariant.

In Chapter IV, we define a class (denoted A_n) of non-negative square matrices A, which are primitive, row stochastic and contiguous, and we define a 1-1 correspondence between A_n and $g(A_n)$, where $g(A_n)$ is a subspace of the space of 0-1 matrices. Since each $A \in g(A_n)$ induces an equivalent class of transformation $\tau \in \mathcal{T}$ (Chapter II), we are able to find A-invariant vectors which are invariant densities under τ.

First, we consider the class P_n of all the permutation invariant matrices A, $A \in A_n$, (section 4.3) and we associate to each $A \in P_n$ a digraph G (Lemma 4.3.4). Then, following a special labelling procedure, we construct (Theorem 4.3.1) the A-invariant vector.

In section 4.4, we extend the P_n-invariant vectors in a number of directions.

In section 4.5, starting with some vectors π, we construct matrices A, which are in A_n but not in P_n, such that $\pi P = \pi$.

In section 4.6, we construct new A_n-invariant vectors from those known to be A_n-invariant.

Finally, in section 4.7, we investigate some properties of A_n-invariant vectors when $A \notin P_n$.

In Chapter V, we use Markov maps to check the irreducibility or primitivity for constructible matrices. More explicitly, we give a new matrix representation for a Markov map, and we show that when a Markov map τ satisfies a simple condition, then a given large, non-negative, square matrix, induced by τ, can be reduced to a smaller one preserving irreducibility or primitivity.

CHAPTER II

FROBENIUS-PERRON OPERATOR AND NON-NEGATIVE MATRICES

2.1. REPRESENTATION OF THE FROBENIUS-PERRON OPERATOR FOR MARKOV MAPS

Let $\tau: [0,1] \to [0,1]$ be a measurable, non-singular transformation. We define the Frobenius-Perron operator $P_\tau: L_1 \to L_1$ by the formula:

$$P_\tau f(x) = \frac{d}{dx} \int_{\tau^{-1}([0,x])} f(s)ds ,$$

where L_1 is the space of integrable functions on $[0,1]$.

It is well-known [14] that the operator P_τ is linear and continuous. Moreover, it satisfies the following conditions:

a) P_τ is positive : $f \geq 0 \Rightarrow P_\tau f \geq 0$.

b) P_τ preserves integrals:

$$\int_0^1 P_\tau f dm = \int_0^1 f dm , \quad f \in L_1 .$$

c) $P_{\tau^n} = P_\tau^n$.

d) $P_\tau f = f$, if and only if the measure $d\mu = f dm$ is invariant under τ .

In this section, working with piecewise constant functions f, we shall show that the Frobenius-Perron operator can be represented as a product of a vector f and of a $n \times n$ matrix M_τ, depending on τ. To show that we need the following result.

Theorem 2.1.1: Let $0 = b_0 < b_1 < \dots < b_n = 1$ be a partition of $[0,1]$. Let $\varphi_i \in C^1(b_{i-1}, b_i)$ and monotone. Assume also that each φ_i can be extended as a monotone C^1 function on $[0,1]$. Then for

$\phi = \sum_{i=1}^{n} \varphi_i \chi_{B_i}$, where $B_i = (b_{i-1}, b_i)$,

$$P_\phi f(x) = \sum_{i=1}^{n} f(\psi_i(x))\, \sigma_i(x) \chi_{J_i}(x) ,$$

where $\psi_i = \varphi_i^{-1}$, $\sigma_i = |\psi_i'|$, $J_i = \varphi_i(B_i)$ and χ_A is the characteristic function of the set A.

Proof: Let $A_i(x) = \varphi_i^{-1}([0,x]) \cap B_i = \psi_i([0,x]) \cap B_i$. Then since $\phi = \sum_{i=1}^{n} \varphi_i \chi_{B_i}$, we get

$$\phi^{-1}([0,x]) = \bigcup_{i=1}^{n} A_i(x) ,$$

where $A_i(x)$'s are disjoint, since B_i's are disjoint. Thus

$$P_\phi f(x) = \frac{d}{dx} \int_{\phi^{-1}([0,x])} f(s)ds$$

$$= \sum_{i=1}^{n} \frac{d}{dx} \int_{A_i(x)} f(s)ds. \tag{2.1}$$

Considering now the integral $\int_{A_i(x)} f(s)ds$, we have:

$$\int_{A_i(x)} f(s)ds = \pm \int_{\psi_i(o)}^{\psi_i(x)} f(s)\, \chi_{B_i}(s)ds .$$

By condition (a), above, we want $\int_{A_i(x)} f \geq 0$ when $f \geq 0$. Now, since φ_i is monotone, ψ_i is monotone and φ_i and ψ_i are either both increasing or both decreasing. Therefore,

$$\frac{\psi_i'(x)}{|\psi_i'(x)|} = \frac{\psi_i'(y)}{|\psi_i'(y)|} \quad \text{for each } x,y \in [0,1].$$

We use this to establish the sign. Thus,

$$\int_{A_i(x)} f(s)ds = \frac{\psi_i'(x)}{|\psi_i'(x)|} \int_{\psi_i(o)}^{\psi_i(x)} f(s)\, \chi_{B_i}(s)ds$$

and

$$\frac{d}{dx}\int_{A_i(x)} f(s)ds = \frac{\psi_i'(x)}{|\psi_i'(x)|} \frac{d}{dx} \int_{\psi_i(o)}^{\psi_i(x)} f(s)\, \chi_{B_i}(s)ds$$

$$= \frac{\psi_i'(x)}{|\psi_i'(x)|} f(\psi_i(x))\, \chi_{B_i}(\psi_i(x))\, \psi_i'(x)$$

$$= \frac{|\psi_i'(x)|^2}{|\psi_i'(x)|} f(\psi_i(x))\, \chi_{B_i}(\psi_i(x))$$

$$= f(\psi_i(x))\, \sigma_i(x)\, \chi_{B_i}(\psi_i(x)) .$$

We note that

$$\chi_{B_i}(\psi_i(x)) = 1 \longleftrightarrow \psi_i(x) \in B_i$$

$$\longleftrightarrow x \in \varphi_i(B_i) = J_i$$

$$\longleftrightarrow \chi_{J_i}(x) = 1$$

Therefore, $\chi_{J_i}(x) = \chi_{B_i}(\psi_i(x))$ and we obtain

$$\frac{d}{dx}\int_{A_i(x)} f(s)ds = f(\psi_i(x))\, \sigma_i(x)\, \chi_{J_i}(x) . \tag{2.2}$$

Now, by substituting (2.2) into (2.1) we get

$$P_\phi f(x) = \sum_{i=1}^{n} f(\psi_i(x))\, \sigma_i(x)\, \chi_{J_i}(x) .$$

Q.E.D.

<u>Lemma 2.1.1</u>: If $f \varepsilon C^1[a,b]$ with $|f'| > 0$ then f is monotone on $[a,b]$.

<u>Proof</u>: $f \varepsilon C^1[a,b] \Rightarrow f' \varepsilon C[a,b]$

and

$$|f'| > 0 \text{ implies } -\infty < f' < 0 \text{ or } 0 < f' < \infty ,$$

since f' is continuous, it is only possible to have $f' < 0$ for each $x \varepsilon [a,b]$ or $f' > 0$ for each $x \varepsilon [a,b]$. In either case, f is monotone on $[a,b]$.

Q.E.D.

Let now I be the interval $[0,1]$ with a fixed partition J given by:

$$J = \left\{0 = a_0 < a_1 < \ldots < a_n = 1\right\} .$$

For each i, $1 \leq i \leq n$, let $I_i = (a_{i-1}, a_i)$ and let $\tau: I \to I$ be a piecewise linear C^1 transformation with respect to J, with $\inf_x |\tau'(x)| > 1$. Let F be the class of all functions which are piecewise constant on the above partition, that is,

$$f \varepsilon F \iff f(x) = \sum_{i=1}^{n} f_i \chi_{I_i}(x) ,$$

with the requirement that $\sum_{i=1}^{n} f_i = 1$. Such an f will be also represented by the row vector $f = (f_1, f_2, \ldots, f_n)$. The following theorem is one of the most important of this thesis.

<u>Theorem 2.1.2</u>: If τ satisfies the above assumptions, then there exists an $n \times n$ matrix M_τ such that $P_\tau f = f M_\tau$ for every $f \varepsilon F$.

<u>Proof</u>: Since P_τ is a linear operator, we have,

$$P_\tau f(x) = P_\tau \sum_{i=1}^{n} f_i \chi_{I_i}(x)$$

$$= \sum_{i=1}^{n} f_i P_\tau \chi_{I_i}(x) .$$

By Lemma 2.1.1 and Theorem 2.1.1 we obtain:

$$P_\tau f(x) = \sum_{i=1}^{n} f_i \sum_{j=1}^{n} \chi_{I_i}(\tau_j^{-1}(x)) \left|\frac{d\tau_j^{-1}(x)}{dx}\right| \chi_{\tau_j(I_j)}(x)$$

$$= \sum_{i=1}^{n} f_i \sum_{j=1}^{n} \chi_{I_i}(\tau_j^{-1}(x)) |\tau_j'|^{-1} \chi_{\tau_j(I_j)}(x) .$$

Since τ_j^{-1} has range I_j, $\chi_{I_i}(\tau_j^{-1}(x))$ will be zero for all $i \neq j$.

Thus

$$P_\tau f(x) = \sum_{i=1}^{n} f_i |\tau_i'|^{-1} \chi_{\tau_i(I_i)}(x) . \tag{2.3}$$

This proves that $P_\tau f$ is a piecewise constant function, i.e., $P_\tau f \in F$. Therefore, there exist constants $d_1, d_2, \ldots, d_n$ such that, for $x \in I_j$

$$P_\tau f(x) = d_j , \quad 1 \leq j \leq n . \tag{2.4}$$

Now we shall compute the constants d_j, for $j = 1,2,\ldots,n$. Let $x \in I_j$ then the right hand side of (2.3) is $f_i |\tau_i'|^{-1}$ if $I_j \subset \tau_i(I_i)$ and zero otherwise. If we define

$$m_{ij} = \begin{cases} |\tau_i'|^{-1} , & \text{if} \quad I_j \subset \tau_i(I_i) \\ 0 , & \text{otherwise} \end{cases} , \tag{2.5}$$

then for $x \in I_j$ (2.3) becomes

$$d_j = \sum_{i=1}^{n} f_i m_{ij} \quad , \quad 1 \le j \le n \ . \tag{2.6}$$

Using matrix representation, we can write (2.6) in the following form

$$(d_1, d_2, \ldots, d_n) = f M_\tau \ , \tag{2.7}$$

where f is a row vector with coefficients f_i, $1 \le i \le n$, and $M_\tau = (m_{ij})$ is a $n \times n$ matrix defined by (2.5).

Therefore using (2.4) and (2.7) we get

$$P_\tau f = f M_\tau \ .$$

Q.E.D.

As we mentioned in the introduction there exists a class of transformations such that each member of that class admits a unique absolutely continuous invariant measure. The existence of such a class is shown in [3] and is defined as follows:

Definition 2.1.1: We say that a point transformation $\tau: I \to I$, $I = [a,b]$, is in class $\mathcal{T}$ if it satisfies the following conditions for a fixed partition $\mathcal{J}$:

(1) τ is piecewise C^2 with respect to $\mathcal{J}$.

(2) $\inf_x |\tau'(x)| > 1$.

(3) τ takes partition points into partition points.

(4) The partition $\mathcal{J}$ has the communication property under τ.

In [14] it is proved that the invariant density f under τ is a fixed point of the Frobenius-Perron operator P_τ, i.e., is an invariant function under P_τ, $P_\tau f = f$. Moreover, in [10], it is proved that the invariant density f under P_τ is a piecewise constant function on the Markov partition $\mathcal{J} = \{I_i\}_{i=1}^{n}$ of $I = [0,1]$.

This permits us to consider the function f as an n-vector $f = (f_1, f_2, \ldots f_n)$ defined by $f_i = f(x)$ for each $x \varepsilon I_i$, $1 \le i \le n$. Now, using Theorem 2.1.2, the piecewise constant function f, which is invariant under P_τ, becomes an invariant vector under M_τ. Hence we can replace the problem of finding the invariant density f under τ using the operator P_τ, by the problem of finding the invariant density f under τ, using the matrix M_τ, which is easier because f is given as the solution of a system of linear equations. For that reason Theorem 2.1.2 is essential to the rest of this thesis.

We shall continue with some remarks to Theorem 2.1.2.

Remark 1: The matrix M_τ defined in the Theorem 2.1.2 is a non-negative $n \times n$ square matrix and, for $1 \le j \le n$ the non-zero entries in the jth row are contiguous and equal to $|\tau_j'|^{-1}$. We shall refer to M_τ as the matrix induced by τ.

Remark 2: Since the transformation τ is piecewise linear the derivative τ_i', for $1 \le i \le n$, is given by the following:

$$\left.\frac{d\,\tau_i(x)}{dx}\right|_{x \varepsilon I_i} = \pm \frac{m(\tau_i(I_i))}{m(I_i)} ,$$

where m is the Lebesgue measure on I. Therefore we have $|\tau_i'|^{-1} = |\pm m(I_i)/m(\tau_i(I_i))|$. Thus we have the following form for the entries of the matrix M_τ.

$$m_{ij} = \begin{cases} \dfrac{m(I_i)}{m(\tau_i(I_i))} , & \text{if } I_j \subset \tau_i(I_i) \\ 0 , & \text{otherwise} \end{cases}$$

Remark 3: Although τ determines M_τ uniquely, by Theorem 2.1.2, the converse is not true. Specifically, an M with equal, contiguous, non-negative entries in each row determines an equivalence class of 2^n transformations, since τ_j' can be positive or negative without altering M_τ. We illustrate that with the following example.

Example 2.1: Let M be a matrix as described above given by

$$M = \begin{bmatrix} .0 & 0 & .5 & .5 \\ .25 & .25 & .25 & .25 \\ 0 & .5 & .5 & 0 \\ .25 & .25 & .25 & .25 \end{bmatrix}$$

the transformations τ_1 and τ_2 show in Figure (2.1)

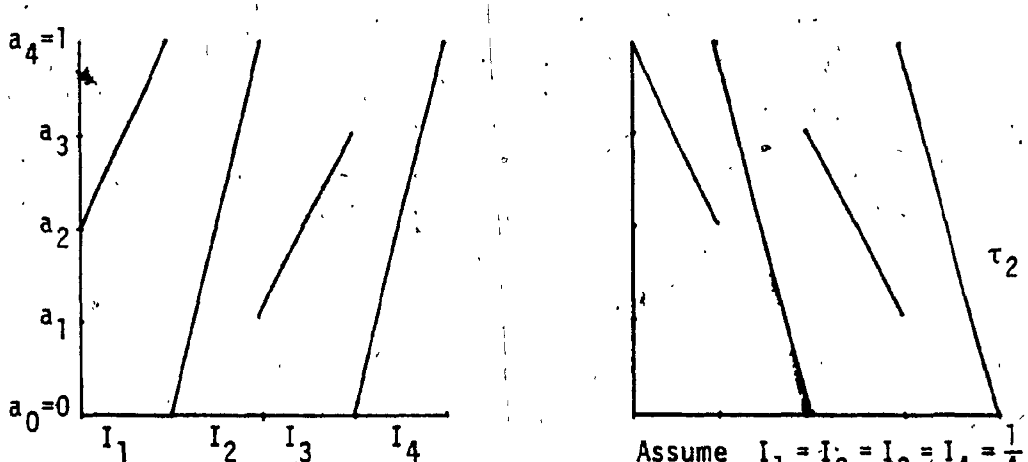


Fig. 2.1

are 2 of the 16 transformations in the equivalence class. Obviously, τ_i, $i = 1,2$, belong to $\mathcal{T}$, since $\inf_x |\tau_i'(x)| > 1$.

2.2. PROPERTIES OF M_τ

In this section we shall study some properties of M_τ. First we shall give another definition for irreducibility of a matrix, and we shall prove that the new definition and the definition 1.1.4 are equivalent. Using Markov chain terminology [20], we give the following definitions.

Definition 2.2.1: A sequence $(i_0, i_1, i_2, \ldots, i_{m-1}, j)$, for $m \geq 1$, $i_0 = i$, from the index set $\{1,2,\ldots,n\}$ of a non-negative matrix A, is said to form a chain of length m between the ordered pair (i,j) if

$$a_{i_0 i_1} a_{i_1 i_2} \cdots a_{i_{m-2} i_{m-1}} a_{i_{m-1} j} > 0 .$$

A chain for which $i = j$ is called a cycle of length m between i and itself.

We say that i leads to j, and write $i \to j$, if there exists an integer $m \geq 1$ such that $a_{ij}^{(m)} > 0$ or, equivalently, if there is a chain between i and j.

We say that i and j communicate if $i \to j$ and $j \to i$ and write $i \leftrightarrow j$.

Definition 2.2.2: If $i \to i$, d(i) is the period of the index i, if it is the greatest common divisor of those k for which $a_{ii}^{(k)} > 0$. Note that if $a_{ii} > 0$ then $d(i) = 1$.

Definition 2.2.3: An $n \times n$ non-negative square matrix A is said to be irreducible (in Markov chain terminology, all states (indices) communicate), if for every pair i,j of its index set, there exists a positive integer q_0 (which may depend on i and j

i.e., $q_0 \equiv q_0(i,j)$ such that $a_{ij}^{(q_0)} > 0$. Note that an irreducible matrix A cannot have a zero row or column.

An irreducible matrix is said to be periodic with period d, if the period of any one of its indices satisfies $d > 1$, and is aperiodic if $d = 1$.

In [12] the following result was proved.

Lemma 2.2.1: If A is a non-negative and irreducible $n \times n$ matrix, then:

$$(I + A)^{n-1} > 0 ,$$

where I is the unit $n \times n$ square matrix.

Proof: It is sufficient to show that for every column vector $y \geq 0$ $(y \neq 0)$ the inequality

$$(I + A)^{n-1} y > 0,$$

holds, (Setting $y = e_j$, $1 \leq j \leq n$, where e_j is the jth coordinate vector, we have $(I + A)^{n-1} > 0$.)

This inequality will be established if we can show that the vector $z = (I + A)y = y + Ay$ always has at least one more positive coordinate than y does. (For, by repeating this argument $n - 1$ times for an arbitrary non-zero $y \geq 0$, we get $(I + A)^{n-1} y > 0$.) Let us assume the contrary. Then, since $A \geq 0$, $Ay \geq 0$, z has at least as many positive coordinates as y. So we assume that y and z have the same positive coordinates. Without loss of generality, we may assume that the columns y and z have the form

$$y = \begin{bmatrix} u \\ 0 \end{bmatrix} , \quad z = \begin{bmatrix} v \\ 0 \end{bmatrix} , \quad (u > 0 \ , \ v > 0)$$

where the columns u and v are of the same dimension. (The columns y and z can be brought into this form by suitably renumbering the coordinates.) Setting:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

we have

$$\begin{bmatrix} u \\ 0 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \cdot \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix},$$

Thus

$$A_{21} u = 0 .$$

Since $u > 0$ and $A \geq 0$, it follows that $A_{21} = 0$, which contradicts the hypothesis that A is irreducible.

Q.E.D.

<u>Corollary</u>: If A is a non-negative and irreducible $n \times n$ square matrix, then for every index pair $(i,j), 1 \leq i,j \leq n$, there exists a positive integer q_o such that,

$$a_{ij}^{(q_o)} > 0 .$$

Moreover, q_o can always be chosen within the bounds:

$$\left.\begin{array}{ll} q_o \leq t-1 , & \text{if } i \neq j \\ q_o \leq t , & \text{if } i = j \end{array}\right\} , \qquad (2.8)$$

where t is the degree of the minimal polynomial $\psi(\lambda)$ of A.

Proof: Let $\rho(\lambda)$ denote the remainder on dividing $(\lambda+1)^{n-1}$ by $\psi(\lambda)$; i.e..

$$(\lambda+1)^{n-1} = \varphi(\lambda)\, \psi(\lambda) + \rho(\lambda) .$$

Then,

$$(A+I)^{n-1} = \varphi(A)\, \psi(A) + \rho(A)$$

or,

$$(A+I)^{n-1} = \rho(A) > 0 ,$$

Since $\psi(A) = 0$ and $(A+I)^{n-1} > 0$, by virtue of Lemma 2.2.1. Since the degree of $\rho(\lambda)$ is less than t, it follows from the inequality $\rho(A) > 0$ that for arbitrary i,j $(1 \le i,j \le n)$, at least one of the non-negative numbers,

$$\delta_{ij}, a_{ij}, a_{ij}^{(2)}, \ldots, a_{ij}^{(t-1)}$$

is not zero. Since $\delta_{ij} = 0$ for $i \neq j$, the first of the relations (2.8) follows. The other relation (for $i = j$) is obtained in a similar manner if the inequality $\rho(A)$ is replaced by $A\,\rho(A) > 0$.

Q.E.D.

Note that the product of an irreducible non-negative matrix and a positive matrix is itself positive.

Therefore, Definition 1.1.4 implies Definition 2.2.3. To show that Definition 2.2.3 implies Definition 1.1.4, we first note that Definition 1.1.4 can be stated as follows.

A $n \times n$ square matrix $A = (a_{ij})$ is called <u>reducible</u> if the index set $\{1,2,\ldots,n\}$ can be split into two complementary sets (without common indices) $\{i_1, i_2, \ldots, i_\mu\}$, $\{j_1, j_2, \ldots, j_\nu\}$, $\mu + \nu = n$, such that:

$$a_{i_\alpha j_\beta} = 0 \quad (\alpha = 1,2,\ldots,\mu;\ \beta = 1,2,\ldots,\nu) .$$

Lemma 2.2.2: Let A be a non-negative $n \times n$ square matrix such that, for each pair (i,j), $1 \le i,j \le n$, there exists a positive integer m, such that $a_{ij}^{(m)} > 0$ i.e., that the state i leads to j $(i \to j)$ or, in other words, there exists a chain of length m between the ordered pair (i,j):

$$a_{ii_1} a_{i_1 i_2} \cdots a_{i_{m-1} j} > 0.$$

Then the matrix A is irreducible.

Proof: Assume that A is reducible, then we can split the index set $\{1,2,\ldots,n\}$ into two complementary sets, without common indices, $I = \left\{i_1,i_2,\ldots,i_\mu\right\}$, $J = \left\{j_1,j_2,\ldots,j_\nu\right\}$, $\mu + \nu = n$, such that:

$$a_{i_\alpha j_\beta} = 0 \quad (\alpha = 1,2,\ldots,\mu,\ \beta = 1,2,\ldots,\nu).$$

Let us consider the ordered pair (i,j) where $i \in I$ and $j \in J$. For that pair, there exist a chain of length q such that:

$$a_{ii_1} a_{i_1 i_2} \cdots a_{i_{q-1} j} > 0 .$$

This is true only if $a_{i_k i_{k+1}} \neq 0$, $k = 0,1,\ldots,q$, $i_0 = i$, $i_{k+1} = j$. Since $i \in I$ and $a_{ii_1} \neq 0$, $i_1 \notin J$, i.e., $i_1 \in I$. Similarly, since $a_{i_1 i_2} \neq 0$, $i_2 \notin J$, i.e., $i_2 \in I$. Continuing in this way, we finally obtain, $i_{q-1} \in I$ and $a_{i_{q-1} j} \neq 0$, which implies that $j \notin J$. But this contradicts the hypothesis that $j \in J$. Therefore, we cannot split the index set $\{1,2,\ldots,n\}$ into two complementary sets, i.e. A is irreducible.

Q.E.D.

Therefore, Definitions 1.1.4 and 2.2.3 are equivalent.

Let us now study some properties of the matrix M_τ.

<u>Lemma 2.2.3</u>: The matrix M_τ, for $\tau \in \mathcal{T}$, τ piecewise linear, is irreducible.

<u>Proof</u>[1]: From the definition of M_τ, we have:

$$m_{ij} = \begin{cases} |\tau_i'|^{-1} & , \text{ if } I_j \subset \tau_i(I_i) \\ 0 & , \text{ otherwise} \end{cases} .$$

We claim that, if $I_j \subset \tau_i^k(I_i)$, then

$$m_{ij}^{(k)} = |(\tau_i^k)'|^{-1} .$$

We shall prove this by induction. For $k=1$, the result is true. Assume that it is true for $k=r$ and consider the case $k=r+1$. Then

$$|(\tau_i^{r+1})'|^{-1} = |[\tau_i(\tau_i^r)]'|^{-1}$$

$$= \left| \frac{d\tau_i(\tau_i^r(x))}{d\tau_i^r(x)} \cdot \frac{d\tau_i^r(x)}{dx} \right|^{-1} .$$

Consider now the term $\frac{d\tau_i(\tau_i^r(x))}{d\tau_i^r(x)}$. Putting $\tau_i^r(x) = y$, we get,

$$\frac{d\tau_i(\tau_i^r(x))}{d\tau_i^r(x)} = \frac{d\tau_i(y)}{dy} .$$

It is obvious that $\frac{d\tau_i(y)}{dy} \neq 0$ if and only if $y \in I_i$ and in that case,

---

[1]This result is due to the author.

since τ is piecewise linear, we have $\frac{d\tau_i(y)}{dy} = \tau_i'$. Therefore, we have

$$|(\tau_i^{r+1})'|^{-1} = \left|\tau_i' \cdot \frac{d\tau_i^r(x)}{dx}\right|^{-1}$$

$$= |\tau_i'|^{-1} \, |(\tau_i^r)'|^{-1}$$

$$= m_{ij} \cdot m_{ij}^{(r)} = m_{ij}^{(r+1)} .$$

Now, since $\tau \in \mathcal{T}$, τ has the communication property, (irreducibility in Markov chain terminology) i.e., for any $I_i, I_j \in \mathcal{J}$, there exist integers p and q such that:

$$I_i \subset \tau^p(I_j) \text{ and } I_j \subset \tau^q(I_i) .$$

Thus for any pair (i,j) of the index set there exists an integer q such that $m_{ij}^{(q)} = |(\tau_i^q)'|^{-1} > 0$, since $I_j \subset \tau_i^q(I_i)$. Hence $m_{ij}^{(q)} > 0$ and so, M_τ is irreducible.

Q.E.D.

<u>Lemma 2.2.4</u>: Let $\tau \in \mathcal{T}$, τ piecewise linear. Assume that the subintervals of the partition $\mathcal{J} = \{I_i\}_{i=1}^n$ are all of equal length. Then the matrix M_τ is row stochastic.

<u>Proof</u>: Let the length of each subinterval be denoted by Δ, i.e.,

$$\ell(I_i) = m(I_i) = a_i - a_{i-1} = \Delta , \quad i = 1,\ldots,n , \; a_0 = 0, \; a_n = 1.$$

$\ell(I_i) = m(I_i)$, since m is the Lebesgue measure. Suppose the Markov map τ maps I_i onto $I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}$. Then,

$\tau_i(I_i) = I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}$, and $m(\tau_i(I_i)) = a_{t+k} - a_{t-1} = (k+1)\Delta$.

Now since

$$m_{ij} = \begin{cases} |\tau_i'|^{-1} = \dfrac{m(I_i)}{m(\tau_i(I_i))} , & \text{if } I_j \subset \tau_i(I_i) \\ 0 , & \text{otherwise} \end{cases} ,$$

it follows that the ith row of M_τ has entries

$$m_{ij} = \frac{m(I_i)}{m(\tau_i(I_i))} = \frac{\Delta}{(k+1)\Delta} = \frac{1}{k+1} , \quad j = t,t+1,\dots,t+k$$

in columns $t,t+1,\dots,t+k$ and zero in all the remaining columns. Therefore,

$$\sum_{j=1}^{n} m_{ij} = \sum_{j=t}^{t+k} \frac{1}{k+1} = \sum_{i=1}^{k+1} \frac{1}{k+1} = 1 .$$

Since i was arbitrary, the matrix M_τ is row stochastic.

Q.E.D.

We shall close this chapter with a Lemma which says under what conditions M_τ is primitive. Before that we recall from matrix theory [20] the following.

<u>Theorem 2.2.1</u>: A non-negative square matrix A is primitive if and only if it is irreducible and has at least one aperiodic index.

<u>Proof</u>: Let A be primitive. Then is obviously irreducible and has $d(i) = 1$ for every i, since for any fixed i, and n large enough, we have $a_{ii}^{(n)} > 0$, $a_{ii}^{(n+1)} > 0$ and the greatest common divisor of n and $n+1$ is 1.

Conversely, suppose that a_{ii} is aperiodic for some i. Since we have:

$$a_{ii}^{(n+m)} = \sum_k a_{ik}^{(n)} a_{ki}^{(m)} \geq a_{ii}^{(n)} a_{ii}^{(m)} ,$$

the set $D = \left\{n \geq 1 : a_{ii}^{(n)} > 0\right\}$ is closed under addition. This, with the fact that its greatest common divisor is 1, implies that D contains all large positive integers say all $n \geq M$.

Now, consider any other index j and choose k and ℓ so that $a_{ij}^{(k)} > 0$ and $a_{ji}^{(\ell)} > 0$. Then for all $n \geq M$ we have:

$$a_{jj}^{(\ell+n+k)} \geq a_{ji}^{(\ell)} a_{ii}^{(n)} a_{ij}^{(k)} > 0 ,$$

where the first inequality follows from the rule of matrix multiplication and the non-negativity of the elements of A. This means that $a_{jj}^{(m)} > 0$ for all large m, implying a_{jj} is also aperiodic. Thus any element on the diagonal is aperiodic.

Similarly, we have:

$$a_{ji}^{(n+\ell)} \geq a_{ji}^{(\ell)} a_{ii}^{(n)} > 0 \qquad \text{for } n \geq M .$$

Thus, the off-diagonal entries of A^n are likewise positive when n is large enough. Since A is a finite matrix. it is clear that for large values of n, A^n is strictly positive, and so, primitive.

Q.E.D.

<u>Lemma 2.2.5</u>: Let $\tau \in \mathcal{T}$ be piecewise linear and let it have a fixed point. Then the matrix M_τ is primitive.

<u>Proof</u>: By Lemma 2.2.3 M_τ is irreducible, and if we show that M_τ has an aperiodic index then Theorem 2.2.1 is applicable and so M_τ is primitive. Since τ has a fixed point, there exists $x_0 \in I_i$ such that $\tau_i(x_0) = x_0$ for some i. This implies that $I_i \subset \tau_i(I_i)$,

since τ is Markov and linear on I_i. It follows that $\tilde{m}_{ii} = |\tau_i'|^{-1} > 0$, i.e., the ith index is aperiodic.

Q.E.D.

CHAPTER III

INVARIANT DENSITIES FOR A CLASS OF PIECEWISE LINEAR TRANSFORMATIONS

3.1. THE PERRON-FROBENIUS THEOREM FOR IRREDUCIBLE MATRICES

In this section we shall prove a theorem known as Perron-Frobenius dealing with non-negative irreducible matrices. Since, as we know from Lemma 2.2.3, M_τ is a non-negative irreducible matrix, the Perron-Frobenius Theorem is essential to the rest of this thesis, as we shall see in section 3.2. We shall start with some preliminaries.

Let R be the field of real or complex numbers. Let $A \in R_{n\times n}$, the space of all $n\times n$ matrices, and $x \in R_n$, the space of all n column vectors, then the vector Ax is in R_n and is a member of the range of A. Let us consider the vectors x, which on multiplication by A are transformed into multiples of themselves: that is, consider those vectors $x \neq 0$ for which there exists a member μ_j of R, the scalar field, such that:

$$Ax = \mu_j x .$$

Such a non-zero vector x is called a (right) eigenvector of A and μ_j is the corresponding eigenvalue. The left eigenvector y of A corresponding to the eigenvalue μ_j is defined by the relation:

$$y^T(\mu_j I-A) = 0^T .$$

It is well known [12 page 54] that if $A \in R_{n\times n}$, the $\mu_j \in R$ is an eigenvalue of A if and only if μ_j is a zero of the characteristic polynomial of A:

$$C(\mu) = \det(\mu I-A) .$$

Let us now state and prove the Perron-Frobenius Theorem [5, 9, 12, 20].

<u>Theorem 3.1.1</u>: (Perron-Frobenius) A non-negative irreducible matrix $A = (a_{ij})$, $1 \le i,j \le n$, always has a positive eigenvalue r, which is a simple root of the characteristic equation. The absolute value of all the other eigenvalues do not exceed r. To the "maximal" eigenvalue r, there corresponds an eigenvector with positive coordinates. Moreover, if $0 \le B \le A$ and β is an eigenvalue of B, then $|\beta| \le r$. If $|\beta| = r$, then $B = A$.

<u>Proof</u>: Consider the real-valued function r defined on the non-zero vectors $x \ge 0$ by

$$r(x) = \min_{1 \le i \le n} \frac{(Ax)_i}{x_i},$$

where $(Ax)_i$ denotes the ith component of the vector Ax, i.e., $(Ax)_i = \sum_{j=1}^{n} a_{ij}x_j$, $i = 1,2,\ldots,n$. We interpret as ∞ the ratio if $x_i = 0$ for some i. Clearly,

$$0 \le r(x) < \infty .$$

Now, since:

$$r(x)x_i \le (Ax)_i = \sum_j a_{ij}x_j \quad \text{for each } i$$

$$r(x)x \le Ax ,$$

and so,

$$r(x)Ix \le IAx .$$

Since $IA \le IK$, where $K = \max_i \sum_j a_{ij}$ it follows that,

$$r(x) \le \frac{IKx}{Ix} = K = \max_i \sum_j a_{ij} ,$$

in other words, $r(x)$ is uniformly bounded above for all such x. Thus $r(x)$ is the largest real number q for which,

$$qx \le Ax .$$

We shall show that the function $r(x)$ assumes a maximum real value r for some real vector $z \ge 0$.

Let L denote the domain of the function r, that is the set of all non-zero, non-negative vectors of order n, and define

$$r = \sup_{x \varepsilon L} r(x) . \qquad (3.1)$$

From the definition of $r(x)$ we observe that the value of $r(x)$ does not change when we multiply the vector $x \ge 0$ $(x \neq 0)$ by a number $\lambda > 0$. Therefore, in the computation of this supremum, we can restrict ourselves to the closed and bounded set M of vectors x for which $x \ge 0$ and $\|x\| = \sum_i x_i^2 = 1$. Thus $M \subset L$ and,

$$r = \sup_{r \varepsilon M} r(x) .$$

It is known [19] that if the function $r(x)$ was continuous on M we could equate the sup $r(x)$ to max $r(x)$, $x \varepsilon M$, and since a continuous function takes on its maximum value on its domain, then the existence of a maximum (and so a supremum) would be guaranteed. However, $r(x)$ may have discontinuities at points where elements of x vanish. Therefore, we replace M by N where N consists of all vectors y of the form,

$$y = (I + A)^{n-1} x , \quad x \varepsilon M .$$

By Lemma 2.2.1 N consists of positive vectors only and so $N \subset L$. Moreover, the set N, like M, is bounded and closed[1], and $r(y)$ is continuous on N.

If we multiply both sides of the inequality,

$$r(x)x \le Ax \quad , \quad x \in M$$

by $(I+A)^{n-1} > 0$, we obtain

$$r(x)y \le Ay \; ; \; y = (I+A)^{n-1}x \in N .$$

Now $r(y)$ is the greatest number q such that $qy \le Ay$ and hence $r(x) \le r(y)$. Thus,

$$r = \sup_{x \in M} r(x) \le \max_{y \in N} r(y) .$$

But since $N \subset L$,

$$\max_{y \in N} r(y) \le \sup_{x \in L} r(x) = \sup_{x \in M} r(x) .$$

Hence

$$r = \max_{y \in N} r(y) \tag{3.2}$$

and there is a $z > 0$ such that $r = r(z)$.

It is possible to have other vectors in L for which $r(x)$ attains the value r. Any such vector is called an extremal vector of A.

It remains to be shown that:

(i) The number r defined by (3.1) is positive and is an eigenvalue of A.

(ii) Every extremal vector z is positive and is an eigenvector of A for the eigenvalue r, i.e., $Az = rz$.

---

[1] The set N is bounded and closed because it is a continuous map, $y=(1+A)^{n-1}x$, of the bounded, closed set of vectors x for which $x \ge 0$ and $\Sigma_i x_i^2 = 1$.

Consider the n-vector $u = (1,1,\ldots,1)^T$, then $r(u) = \min_i \sum_k a_{ik} > 0$[1]. Since $r \geq r(u)$, we deduce that $r > 0$.

Let z be an extremal vector and let $x = (I+A)^{n-1}z$. Without loss of generality, we may suppose that $z \in M$. By Lemma 2.2.1, $x > 0$, and clearly $x \in N$. We also assume that $Az - rz \neq 0$, then,

$$(I+A)^{n-1}(Az - rz) > 0 .$$

Hence $Ax - rx > 0$, $rx < Ax$, which implies that $r < r(x)$. But this contradicts the definition (3.2) of r. Therefore we must have $Az = rz$. Thus any extremal vector z is a right eigenvector of A for the eigenvalue r. Now, since $Az = rz$ we have

$$0 < x = (I+A)^{n-1}z = (1+r)^{n-1}z.$$

and since $r > 0$ we have $z > 0$.

We shall show now that the absolute value of all the other real or complex eigenvalues do not exceed r. Let

$$Ay = \alpha y \quad , \quad y \neq 0 \tag{3.3}$$

since $A \geq 0$, we have:

$$|\alpha||y| = |Ay| \leq A|y|, \qquad (|y| = (|y_1|,|y_2|,\ldots\ldots,|y_2|)) \tag{3.4}$$

Hence $|\alpha| \leq r(|y|) \leq r$, which we wanted to prove.

Now, let y be some eigenvector corresponding to r, $Ay = ry$ $(y \neq 0)$. Then, setting $\alpha = r$ in (3.3) and (3.4), we conclude that

---

[1] Since A is irreducible, it can have no columns consisting entirely of zeros.

$|y|$ is an external vector, so that $|y| > 0$, $y_i \neq 0$, $1 \le i \le n$. This implies that the dimension of the right eigenspace of r is 1, since, otherwise, we could find two linearly independent right eigenvectors z_1, z_2 and then determine numbers α, β such that $\alpha z_1 + \beta z_2$ has a zero coordinate. From what we have shown, this is impossible. This proves that the geometric[1] multiplicity of r is 1.

Let us consider now the adjoint matrix of the characteristic matrix $\lambda I - A$:

$$B(\lambda) = [B_{ik}(\lambda)]_1^n = \Delta(\lambda)(\lambda I - A)^{-1} ,$$

where $\Delta(\lambda)$ is the characteristic polynomial of A and $B_{ik}(\lambda)$, the algebraic complement of the element $\lambda\delta_{ki} - a_{ki}$ in the determinant $\Delta(\lambda)$. Now, since only one[2] eigenvector $z = (z_1, z_2, \ldots, z_n)^T$, $z_i > 0$ for $1 \le i \le n$, corresponds to the eigenvalue r, it follows that $B(r) \neq 0$, and that in every non-zero column of $B(r)$ all the elements are different from zero and are of the same sign. The same is true for the rows of $B(r)$, since A can be replaced by the transposed matrix A^T. From these properties of the rows and columns of A, it follows that all the $B_{ik}(r)$, $1 \le i, k \le n$, are different from zero and of the same sign σ. Therefore:

$$\sigma\Delta'(r) = \sigma \sum_{i=1}^{n} B_{ii}(r) > 0 ,$$

i.e., $\Delta'(r) \neq 0$ ($\Delta'(r)$ means the derivative of $\Delta(\lambda)$ at r) and r is a simple root of the characteristic equation $\Delta(\lambda) = 0$. In other words, the algebraic multiplicity of the eigenvalue r is 1.

---

[1]By geometric multiplicity of an eigenvalue r, we mean the maximal number of linearly independent right eigenvectors associated with r.

[2]Up to constant multiples.

Finally, let $y \neq 0$ be a right eigenvector of B corresponding to β. Then taking absolute value, we have:

$$|\beta||y| \leq B|y| \leq A|y| \; . \tag{3.5}$$

Now, taking the vector z, which we used above we get:

$$|\beta| z|y| \leq zA|y| = rz|y| \; ,$$

and since $z|y| > 0$, we get:

$$|\beta| \leq r \; .$$

Suppose now $|\beta| = r$. Then, from (3.5) we have, $r|y| \leq A|y|$. It follows that $A|y| = r|y| > 0$, while from (3.5) we get $r|y| = B|y| = A|y|$. So, it must follow from $B \leq A$ that $B = A$.

Q.E.D.

<u>Corollary 3.1.1</u>:

$$\min_i \sum_{j=1}^{n} a_{ij} \leq r \leq \max_i \sum_{j=1}^{n} a_{ij} \tag{3.6}$$

with equality on either side implying equality throughout (i.e., r can only be equal to the maximal or minimal row sum, if all row sums are equal).

<u>Proof</u>: From (3.2) we have that:

$$r = \max_{x>0} \min_i \frac{(Ax)_i}{x_i} \; . \tag{3.7}$$

Similarly we can show [9, vol. II, p. 64] that:

$$r = \min_{x>0} \max_i \frac{(Ax)_i}{x_i} \; , \tag{3.8}$$

and a combination of (3.7) and (3.8), together with $x = (1,1,\ldots,1)^T$ gives (3.6).

Now assume that one of the equalities in (3.6) holds, but not all row sums are equal. Then by increasing (or decreasing) the positive elements of A (but keeping them positive), we produce a new irreducible matrix, with all rows sums equal and the same r, in view of (3.6) which is impossible by hypothesis of the last part of Theorem 3.1.1.

<u>Remark</u>: We can state an analogous theorem with the corresponding corollary for primitive matrices. The proof remains the same.

<u>Lemma 3.1.1</u>: Let M be a non-negative, row stochastic and irreducible $n \times n$ matrix. Them M has 1 as the eigenvalue of maximum absolute value, and furthermore the algebraic and geometric multiplicities of this eigenvalue are also 1.

<u>Proof</u>: By Corollary 3.1.1 we have:

$$\min_i \sum_{j=1}^{n} m_{ij} \le r \le \max_i \sum_{j=1}^{n} m_{ij} ,$$

since M is stochastic $\sum_{j=1}^{n} m_{ij} = 1$ for each i. Therefore,

$$\min_i \sum_{j=1}^{n} m_{ij} = 1 = \max_i \sum_{j=1}^{n} m_{ij} .$$

This implies that $r = 1$. Moreover, Theorem 3.1.1 shows that the geometric and algebraic multiplicities are also 1.

Q.E.D.

We shall close this section with one more result from matrix theory. Before that, we recall [12] that two $n \times n$ matrices A and B are <u>similar</u> if there exists a non-singular $n \times n$ matrix T (non-singular means that $\det T \neq 0$), such that $A = T^{-1}BT$.

<u>Lemma 3.1.2</u>: Similar matrices have the same eigenvalues.

Proof: We shall prove that similar matrices have the same characteristic polynomial. Hence, it follows that they have the same eigenvalues.

Suppose that $A = T^{-1}BT$, i.e., A and B are similar, and let $C(\mu) = \det(\mu I - A)$ be the characteristic polynomial of A. Then:

$$\begin{aligned} C(\mu) &= \det(\mu I - T^{-1}BT) \\ &= \det(\mu T^{-1}T - T^{-1}BT) \\ &= \det(T^{-1}\mu T - T^{-1}BT) \\ &= \det T^{-1}(\mu I - B)T \\ &= (\det T^{-1}) \det(\mu I - B) \det(T) \\ &= \det(\mu I - B) \quad , \end{aligned}$$

since $\det(T^{-1}) \det(T) = 1$. Thus, the characteristic polynomials of A and B are the same, i.e., the eigenvalues of a matrix are invariant under similar transformation.

Q.E.D.

3.2. THE MATRIX M_τ AND ITS EIGENFUNCTIONS

As we show in Chapter II the piecewise linear Markov map τ, $\tau \in \mathcal{T}$, induces an $n \times n$ matrix M_τ defined as follows:

$$m_{ij} = \begin{cases} |\tau_i'|^{-1} & , \text{ if } I_j \subset \tau_i(I_i) \\ 0 & , \text{ otherwise} \end{cases}$$

Lemma 2.2.3 shows that M_τ is irreducible and Lemma 2.2.4 shows that, when the subintervals of the partition $\mathcal{J}$ are all of equal length, the matrix M_τ is stochastic. Therefore, by Lemma 3.1.1, the stochastic matrix M_τ has 1 as the eigenvalue of maximum absolute value, with geometric and algebraic multiplicities also 1.

In [6] it is proved that the same is true for any matrix M_τ, $\tau \in \mathcal{T}$, τ piecewise linear, and not necessarily stochastic. This implies that there exists a large class of non-negative, non-stochastic, irreducible matrices whose eigenvalues of maximum absolute value is 1.

Before we present this fact, we remark that pre- or post-multiplication by a non-singular diagonal matrix cannot change any non-zero elements of a matrix A to zero, and therefore such multiplication has not effect on irreducibility.

<u>Theorem 3.2.1</u>: Let $\tau \in \mathcal{T}$, be piecewise linear. Then the matrix M_τ induced by τ has 1 as the eigenvalue of maximum absolute value, the algebraic and geometric multiplicities of this eigenvalue are also 1.

<u>Proof</u>: Let $0 = a_0 < a_1 < \ldots < a_n = 1$ be a partition of $[0,1] = I$, and let $I_j = (a_{j-1}, a_j)$, $1 \le j \le n$. Define

$$\delta = \prod_{j=1}^{n} (a_j - a_{j-1}),$$

and

$$\delta_i = \frac{\delta}{a_i - a_{i-1}} = \prod_{\substack{j=1 \\ j \ne i}}^{n} (a_j - a_{j-1}) \qquad 1 \le i \le n .$$

Now, let us define the diagonal matrix D to have entries $d_{ii} = \delta_i$, $1 \le i \le n$. Then $E = D^{-1}$ is a diagonal matrix with entries

$e_{ii} = \delta_i^{-1}$, $1 \le i \le n$.

Suppose now that

$$\tau(I_i) = I_j \cup I_{j+1} \cup \ldots \cup I_{j+k} .$$

Then, by definition, the ith row of M_τ has entries equal to $\frac{m(I_i)}{m(\tau_i(I_i))} = \frac{a_i - a_{i-1}}{a_{j+k} - a_{j-1}}$ in columns $j, j+1, \ldots, j+k$, and zero in every remaining column. Now, let $B = DM_\tau D^{-1}$. Then,

$$\begin{aligned} b_{rs} &= d_{rr} m_{rs} d_{ss}^{-1} \\ &= \delta_r m_{rs} \delta_s^{-1} . \end{aligned}$$

We claim that B is row stochastic. To show that, let us consider the row sum of the ith row of B.

$$\begin{aligned} \sum_{s=1}^{n} b_{is} &= \sum_{s=1}^{n} \delta_i m_{is} \delta_s^{-1} \\ &= \sum_{s=j}^{j+k} \delta_i \frac{a_i - a_{i-1}}{a_{j+k} - a_{j-1}} \delta_s^{-1} \\ &= \frac{\delta}{a_i - a_{i-1}} \frac{a_i - a_{i-1}}{a_{j+k} - a_{j-1}} \left[\delta_j^{-1} + \delta_{j+1}^{-1} + \ldots + \delta_{j+k}^{-1}\right] \\ &= \frac{\delta}{a_{j+k} - a_{j-1}} \left[\frac{a_j - a_{j-1}}{\delta} + \frac{a_{j+1} - a_{j+1-1}}{\delta} + \ldots + \frac{a_{j+k} - a_{j+k-1}}{\delta}\right] \\ &= 1 . \end{aligned}$$

Therefore B is row stochastic, and since M_τ is irreducible, then so is B. Thus the result follows from Lemma 3.1.1 and Lemma 3.1.2.

Q.E.D.

As we mentioned in the introduction, (page 3) one of the main goals of this thesis is to find the unique invariant densities for Markov maps τ, belonging in $\mathcal{T}$. In the following we shall find the unique invariant density of a piecewise linear transformation $\tau \in \mathcal{T}$ when it satisfies two simple conditions.

Let $I = [0,1]$, and $0 = a_0 < a_1 < \ldots < a_n = 1$ be any partition of I. Let $I_i = (a_{i-1}, a_i)$, $1 \le i \le n$. Let us consider the piecewise linear transformation $\tau : I \to I$, $\tau \in \mathcal{T}$, defined by the conditions:

$$(\varepsilon_1) \qquad \tau(I_i) = I_{i+1} \quad , \quad 1 \le i < n$$

and

$$(\varepsilon_2) \qquad \tau(I_n) = \bigcup_{i=1}^{n} I_i = I \ .$$

The transformation τ induces a matrix M_τ which, according to Theorem 3.2.1, has 1 as the eigenvalue of maximum absolute value, with algebraic and geometric multiplicity both equal to 1. To that eigenvalue there corresponds a unique (left) eigenvector (since the geometric multiplicity is 1) π, i.e.,

$$\pi M_\tau = \pi \ .$$

In Chapter II, page 13, we show that the vector π, when considered as a piecewise constant function is the unique density invariant under τ.

The next theorem shows how we can construct invariant vectors π.

<u>Theorem 3.2.2</u>: Let $\tau \in \mathcal{T}$, piecewise linear, be a transformation on [0,1], satisfying conditions (ε_1) and (ε_2). Then it admits a unique (up to constant multiples) invariant density (eigenfunction) $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, where:

$$\pi_i = \pi\Big|_{I_i} = \frac{a_i - a_0}{a_i - a_{i-1}} \ , \ a_0 = 0 \ , \ a_n = 1 \ .$$

<u>Proof:</u> By condition (ε_1), the $n \times n$ matrix M_τ is given by:

$$m_{ij} = \begin{cases} \dfrac{a_i - a_{i-1}}{a_{i+1} - a_i} , & \text{if } j = i+1 \text{ (or } i = j-1) \\ 0 , & \text{otherwise} . \end{cases}$$

Also,

$$m_{nj} = \frac{a_n - a_{n-1}}{a_n - a_o} , \quad 1 \le j \le n \ ;$$

i.e., M_τ has nonzero entries on the superdiagonal and on the last row.

We know that M_τ has 1 as the eigenvalue of maximum absolute value. It only remains to show that $\pi M_\tau = \pi$. we have,

$$\sum_{r=1}^{n} \pi_r m_{rs} = \begin{cases} \pi_{s-1} m_{s-1,s} + \pi_n m_{ns} , & s \neq 1 \\ \pi_n m_{ns} , & s = 1 . \end{cases}$$

For $s = 1$, we have,

$$\begin{aligned} \pi_n m_{ns} &= \frac{a_n - a_o}{a_n - a_{n-1}} \cdot \frac{a_n - a_{n-1}}{a_n - a_o} \\ &= 1 \\ &= \frac{a_1 - a_o}{a_1 - a_{1-1}} = \pi_1 . \end{aligned}$$

For $s \neq 1$, we have,

$$\begin{aligned} \pi_{s-1} m_{s-1,s} + \pi_n m_{ns} &= \frac{a_{s-1}-a_o}{a_{s-1}-a_{s-2}} \cdot \frac{a_{s-1}-a_{s-2}}{a_s - a_{s-1}} + 1 \\ &= \frac{a_{s-1}-a_o}{a_s-a_{s-1}} + \frac{a_s-a_{s-1}}{a_s-a_{s-1}} \\ &= \frac{a_s-a_o}{a_s-a_{s-1}} = \pi_s . \end{aligned}$$

Hence, $\pi M_\tau = \pi$. Since the geometric multiplicity of the eigenvalue 1 is 1 , π is unique up to constant multiples.

Q.E.D.

Now we consider the inverse problem. Given any vector $g = (g_1, g_2, \ldots, g_n)$ of length n, is it possible to construct a Markov map τ whose induced matrix M_τ satisfies $gM_\tau = g$? The answer is affirmative if g satisfies the following two conditions:

(i) $g_1 = c$, where $c > 0$, any constant,

and

(ii) $g_i > c$, $2 \le i \le n$.

If (i) and (ii) are satisfied, we can show that for every such g, there exists a Markov map τ satisfying conditions (ε_1) and (ε_2) above and such that g is the unique eigenfunction associated with the eigenvalue 1 of the matrix M_τ . Without loss of generality, we can assume $c = 1$.

<u>Theorem 3.2.3</u>: Let $g = (g_1, g_2, \ldots, g_n)$ be any vector of length n, such that $g_1 = 1$ and $g_i > 1$, $2 \le i \le n$. Then there exists a Markov map whose induced matrix M_τ satisfies $gM_\tau = g$.

<u>Proof</u>: Let $a_0 = 0$. We choose $a_1, 0<a_1<1$ such that:

$$a_i = a_{i-1} + \frac{a_{i-1} - a_0}{g_i - 1} \le 1 ,$$

for all $2 \le i \le n$. It follows that:

$$a_i - a_0 = a_{i-1} - a_0 + \frac{a_{i-1} - a_0}{g_i - 1} ,$$

and

$$a_i - a_{i-1} = \frac{a_{i-1} - a_0}{g_i - 1} \; .$$

consider:

$$\pi_i = \frac{a_i - a_0}{a_i - a_{i-1}}$$

$$= \left[a_{i-1} - a_0 + \frac{a_{i-1} - a_0}{g_i - 1}\right] \Big/ \frac{a_{i-1} - a_0}{g_i - 1}$$

$$= \left[a_{i-1} - a_0 + \frac{a_{i-1} - a_0}{g_i - 1}\right] \cdot \left[\frac{g_i - 1}{a_{i-1} - a_0}\right]$$

$$= g_i - 1 + 1$$

$$g_i \; .$$

Clearly, $\pi_1 = 1 = g_1$. Therefore, from Theorem 3.2.2, we have:

$$gM_\tau = g.$$

Q.E.D.

We shall now show by means of an example how to construct the unique invariant density for a given Markov map.

Example 3.1: Let $\tau: [0,1]\to[0,1]$ be the piecewise linear transformation in $\mathcal{T}$, defined by $\tau(0) = 0$, $\tau(\frac{1}{2}) = 1$, $\tau(1) = \frac{1}{2^n}$, $n \geq 2$, (Fig. 3.1).

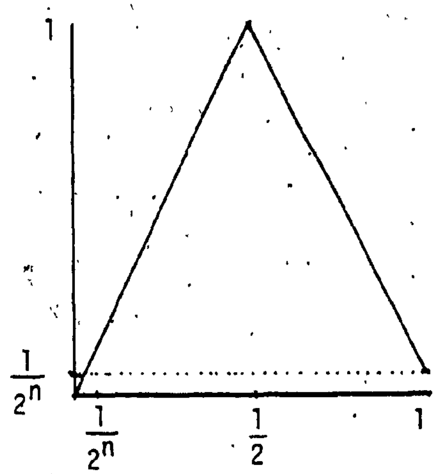


Fig. 3.1. $\tau:[0,1]\to[0,1]$

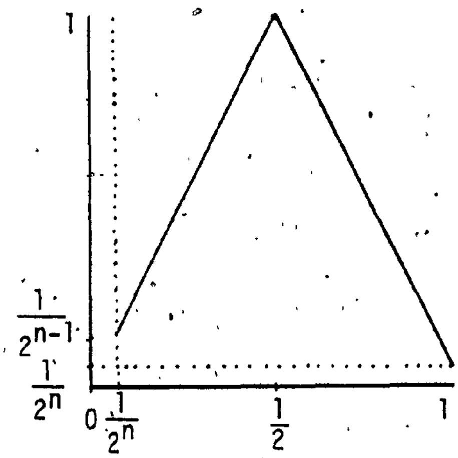


Fig. 3.2. $\tau:[\frac{1}{2^n},1]\to[\frac{1}{2^n},1]$.

Let $n \geq 2$ be fixed, and let us consider the restriction of τ on $[\frac{1}{2^n},1]$ (Fig. 3.2). Let $\mathcal{J}$ be a partition of $[\frac{1}{2^n},1]$ given by $a_o = \frac{1}{2^n} < \frac{1}{2^{n-1}} <\ldots< \frac{1}{2} < 1 = a_n$. Obviously $\mathcal{J}$ is a Markov partition, with respect to τ, of $[\frac{1}{2^n},1]$. Moreover $\tau(I_i) = I_{i+1}$, $1 \leq i \leq n$, and $\tau(I_n) = \bigcup_{i=1}^{n} I_n$, where $I_i = (a_{i-1},a_i)$ and $a_i = \frac{1}{2^{n-i}}$, $0 \leq i \leq n$, i.e., τ satisfies conditions (ε_1) and (ε_2).

The transformation τ induces an $n \times n$ matrix M_τ given by:

$$m_{ij} = \begin{cases} |\tau_i'|^{-1} = \left|\dfrac{m(I_i)}{m(\tau_i(I_i))}\right| = \left|\dfrac{\ell(I_i)}{\ell(\tau_i(I_i))}\right| , & \text{if } I_j \subset \tau_i(I_i) \\ 0 , & \text{otherwise} \end{cases} .$$

From the definition of τ, the ith row of M_τ, for $i < n$, has zero everywhere except for the $(i+1)$th entry, which is given by,

$$|\tau_i'|^{-1} = \left|\frac{\ell(I_i)}{\ell(\tau_i(I_i))}\right| = \left|\frac{\ell(I_i)}{\ell(I_{i+1})}\right|$$

$$= \left|\frac{\frac{1}{2^{n-i}} - \frac{1}{2^{n-i+1}}}{\frac{1}{2^{n-i-1}} - \frac{1}{2^{n-i}}}\right| = \left|\frac{\frac{2-1}{2^{n-i+1}}}{\frac{2-1}{2^{n-i}}}\right|$$

$$= \left|\frac{2^{n-i}}{2^{n-i+1}}\right|$$

$$= \frac{1}{2} .$$

The nth row of M_τ has all entries equal given by,

$$|\tau_i'|^{-1} = \left|\frac{\ell(I_n)}{\ell(\tau(I_n))}\right|$$

$$= \left|\frac{a_n - a_{n-1}}{a_n - a_o}\right|$$

$$= \frac{1 - \frac{1}{2}}{1 - \frac{1}{2^n}}$$

$$= \frac{2^{n-1}}{2^n - 1}$$

Therefore, the $n \times n$ matrix induced by τ is :

$$M_\tau = \begin{bmatrix} \frac{1}{2} & & & \\ & \frac{1}{2} & & 0 \\ & 0 & \ddots & \\ & & & \frac{1}{2} \\ \frac{2^{n-1}}{2^n-1} & \cdots\cdots & & \frac{2^{n-1}}{2^n-1} \end{bmatrix}$$

Since τ is Markov, M_τ is irreducible. Hence 1 is its eigenvalue of maximum absolute value. Indeed, the unique (up to constant multiples) eigenvector is given by $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, where

$$\pi_i = \frac{a_i - a_0}{a_i - a_{i-1}} = \frac{2^{n+1} - 2^{n+1-i}}{2^{n+1} - 2^n} = 2 - 2^{1-i} .$$

When viewed as a density, π can be expressed as:

$$f(x) = \sum_{i=1}^{n} \left[\frac{2^{n+1} - 2^{n+1-i}}{2^{n+1} - 2^n}\right] \chi_{I_i}(x)$$

$$= \sum_{i=1}^{n} (2 - 2^{1-i}) \chi_{I_i}(x) ,$$

where χ_{I_i} is the characteristic function on the set I_i, $1 \leq i \leq n$.

The area under f, for $n \to \infty$, is 1, since,

$$\lim_{n\to\infty} \int_{[\frac{1}{2^n},1]} f(x)dx = \lim_{n\to\infty} \int_{[\frac{1}{2^n},0]} \sum_{i=1}^{n} \pi_i \chi_{I_i}(x)dx$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \pi_i \int_{a_{i-1}}^{a_i} dx$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \frac{a_i - a_o}{a_i - a_{i-1}} \cdot (a_i - a_{i-1})$$

$$= \lim_{n\to\infty} (a_n - a_o)$$

$$= 1.$$

CHAPTER IV

MATRICES INDUCED BY PIECEWISE LINEAR TRANSFORMATION AND THEIR INVARIANT VECTORS

4.1. PRELIMINARIES

In Chapter II, we showed that a piecewise linear transformation $\tau \in \mathcal{T}$ induces a uniquely determined square matrix M_τ, which is non-negative, irreducible and row contiguous. We also showed that starting from M_τ we can determine an equivalent class of 2^n transformations, and also that the problem of finding the unique invariant densities under the piecewise linear transformations $\tau \in \mathcal{T}$ is equivalent to the problem of finding the invariant vectors under M_τ.

In Chapter III, we showed that each member from the class of non-negative square, irreducible (or primitive) stochastic matrices possesses a unique[1] invariant vector. That means that every square matrix M_τ which is non-negative, primitive, stochastic and contiguous possesses a unique invariant vector, which can serve as the unique invariant density for each one of the transformations in the class of 2^n transformations which are determined by M_τ.

In this Chapter, we shall characterize those classes of n-vectors which can serve as a unique invariant vector for a matrix A in a class A_n, defined below.

---

[1]Up to constant multiples.

Definition 4.1.1: We say that the non-negative, $n \times n$ square matrix $A \in A_n$, if A satisfies the following conditions:

(i) A is primitive.

(ii) A is row stochastic.

(iii) A is row contiguous.

Since we shall use some results from graph theory, to check the primitivity of a matrix A, and to construct A_n-invariant vectors, we introduce in the following section, some basic definitions from graph theory [2, 4, 11, 18].

4.2. ELEMENTS FROM GRAPH THEORY

Directed Graphs

A directed graph G (or digraph) consists of two things:

(i) A set V whose elements are called vertices, or points.

(ii) A set E of ordered pairs of vertices called arcs (edges).

We shall denote a directed graph by $G = (V,E)$. We can picture directed graphs by diagrams in the plane where each vertex v in V is represented by a dot and each arc $a = \langle u,v \rangle$ is represented by an arrow from the initial (beginning) point u of a to its terminal (ending) point v. If the initial point for an arc is the same vertex as its terminal point, then the arc is called a loop.

Since we are concerned only with directed graphs we shall generally omit the word "directed".

We say that $G = (V,E)$ is finite if its set V of vertices is finite and its set E of arcs is finite.

Let V' be a subset of V and let E' be a subset of E whose endpoints belong to V'. Then G' = (V',E') is a (directed) graph and is called a subgraph of G = (V,E).

A complete graph on n-vertices is a graph with the property that every pair of vertices is an arc. A k-clique of a graph G is a complete subgraph of G on k vertices.

The outdegree and indegree of a vertex v are equal, respectively, to the number of arcs beginning and ending at v. A vertex with zero outdegree is called a sink and a vertex with zero indegree is called a source.

A walk from v to w in a graph G is a sequence of vertices $v_1, v_2, \ldots, v_k$ such that $v = v_1$ and $w = v_k$ and $\langle v_{i-1}, v_i \rangle \in E$ for all $1 < i \leq k$. A semiwalk from v to w is a sequence of vertices $v_1, v_2, \ldots, v_k$ such that $v_1 = v$, $v_k = w$ and either $\langle v_{i-1}, v_i \rangle \in E$ or $\langle v_i, v_{i-1} \rangle \in E$ for all $1 < i \leq k$. A (semi) path is a (semi) walk with all vertices distinct.

A graph is strongly connected if for all $v, w \in V$, $v \neq w$ there is a path from v to w. It is weakly connected if every pair of distinct points is joined by a semipath. A weak component of a graph is a maximal weakly connected subgraph.

Rooted Trees

A cycle is a closed walk with distinct vertices except for the first and the last vertices. A semicycle of a graph G is a semipath from a vertex v to itself. A graph G is said to be acyclic or cycle-free if it contains no cycles.

A rooted tree is a weakly connected graph which does not contain semicycles with a designated vertex r, called the root.

A forest is an acyclic graph in which every weak component is a tree.

An in-tree is a rooted tree in which there is a path from every vertex to the root. It is known [11 pg. 201] that a weakly connected graph is an in-tree if and only if exactly one vertex has outdegree 0 and all others have outdegree 1.

A vertex w is said to be a child of the vertex v in a rooted tree if the (unique) semipath from the root r to w has v as its penultimate vertex (i.e., a semipath $r = v_1,\ldots,v_k,v,w$).

A vertex is a leaf of a rooted tree if it has no children.

The length of a path from the root r to v is called the level or depth of v.

Matrices and Graphs

Consider a finite graph G and let us label the vertices of the graph G by assigning subscripts to them so that $V = \{v_1,v_2,\ldots,v_m\}$. It is known [4] that we can represent that graph by a matrix in many different ways (adjacency matrix, reachability matrix, distance matrix). For the purpose of this thesis, we shall employ the adjacency matrix, which is defined as follows: $A = A(G) = (a_{ij})$, where,

$$a_{ij} = \begin{cases} 1 , & \text{if } \langle v_i,v_j\rangle \in E \\ 0 , & \text{otherwise} \end{cases}$$

$1 \leq i,j \leq m$. Thus a graph can be represented by a 0-1 square matrix and, conversely, a 0-1 square matrix can be considered as an adjacency matrix for some graph.

We elucidate that with the following example.

<u>Example 4.1</u>: A digraph and its adjacency matrix.

G :

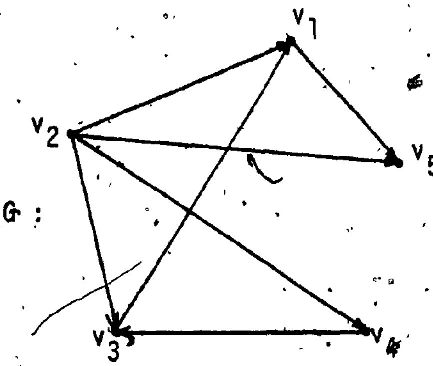


Fig. 4.1

A(G) :

	v_1	v_2	v_3	v_4	v_5	Row sum
v_1	0	0	0	0	1	1
v_2	1	0	1	1	1	4
v_3	1	0	0	0	0	1
v_4	0	0	1	0	0	1
v_5	0	0	0	0	0	0
Column sum	2	0	2	1	2	

From the above example we note that the row sums of $A(G)$ give the outdegrees of the vertices of G and the column sums are the indegrees.

We shall continue with some interesting observations for the adjacency matrix of a digraph G.

<u>Theorem 4.2.1</u>: Let $A = A(G)$ be an adjacency matrix of G. Then the (i,j)th entry of the matrix A^n gives the number of walks of length n from the vertex v_i to the vertex v_j.

<u>Proof</u>: We prove this by induction on n. Note that a walk of length 1 from v_i to v_j is precisely an arc $\langle v_i,v_j\rangle$. Thus the theorem holds for $n = 1$, since the (i,j)th entry of A gives the number of arcs $\langle v_i,v_j\rangle$ (note this number is zero or 1), which is the number of walks of length 1 from v_i to v_j.

Suppose $n > 1$, $A^{n-1} = (a_{ik})$ and $A = (b_{kj})$. Then $A^n = (c_{ij})$ is given by $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. By induction, a_{ik} equals the number of walks of length $n-1$ from v_i to v_j. Also, b_{kj} equals the number of walks (arcs) from v_k to v_j. Thus, $a_{ik} b_{kj}$ gives the number of walks of length n from v_i to v_j, where v_k is the next-to-last vertex in the walk. Thus, all walks of length n from v_i to v_j can be obtained by summing up the $a_{ik} b_{kj}$ for all k, that is, c_{ij} is the number of walks of length n from v_i to v_j. Thus, the theorem is proved.

Q.E.D.

Now, suppose that G has m vertices. Then any path of G cannot contain more than m vertices and hence must have length $m-1$ or less. Thus, G is strongly connected if for any vertices u and v, there exists a path from u to v and one from v to u, each of length $m-1$ or less.

<u>Corollary 4.2.1</u>: (A criterion for connectedness). Let $A = A(G)$ be an adjacency matrix of G. Let:

$$C = A + A^2 + \ldots + A^{m-1} .$$

Then G is strongly connected if and only if C has no zero entries off the main diagonal.

<u>Proof</u>: The (i,j)th entry of C is zero if and only if the (i,j)th entries of $A, A^2, \ldots, A^{m-1}$ are all zero. This means there is no walk from v_i to v_j of length $m-1$ or less, and so there is no path from v_i to v_j. Thus G is strongly connected if and

only if the (i,j)th entries of C, $i \neq j$, are non-zero.

Q.E.D.

It is well known [4 page 171] that a matrix $A = A(G)$ is irreducible if and only if G is strongly connected. It is also known (Theorem 2.2.1) that a non-negative square matrix A is primitive if and only if A is irreducible and has at least one aperiodic index. Thus, using Corollary 4.2.1 we get the following theorem.

<u>Theorem</u>[1] <u>4.2.2</u>: (A test of primitivity) A non-negative $(m \times m)$ matrix A with 0-1 entries is primitive if and only if the matrix

$$C = A + A^2 + \ldots + A^{m-1} .$$

has no zero entries off the main diagonal and $a_{ii} > 0$ for some i, $1 \le i \le m$.

<u>Proof</u>: By Corollary 4.2.1, G is strongly connected and so A is irreducible, and since A has an aperiodic index (the index i), Theorem 2.2.1 confirms that A is primitive.

Q.E.D.

4.3. PERMUTATION INVARIANT MATRICES

Induced 0-1 Matrices

Let A be a square matrix with non-negative entries. We can associate with A a 0-1 square matrix by setting all non-zero entries to 1. We denote this matrix by g(A). We remark that g(A) is an adjacency matrix for some graph. Let us recall that A_n is the class of square $n \times n$ matrices which are stochastic, primitive and each

[1] This result is due to the author.

row of which has its non-zero entries contiguous with equal value.

In the following we shall prove three lemmas which characterize the map $g: A_n \to g(A_n)$, where the map g is defined [8] as follows, for $A = (a_{ij}) \in A_n$:

$$g(a_{ij}) = \begin{cases} 0 \ , & \text{if } a_{ij} = 0 \\ 1 \ , & \text{if } a_{ij} \neq 0 \ , \end{cases}$$

$1 \leq i,j \leq n$. Obviously $g(A_n)$ is a subspace of the space of square 0-1 matrices.

Lemma 4.3.1: The mapping $g: A_n \to g(A_n)$ is 1-1.

Proof: Let $A, A' \in A_n$, where $A \neq A'$. Let the ith row of A contain k_i non-negative entries and let the ith row of A' contain k_i' non-negative entries. Without loss of generality, we assume in view of the fact that $A \neq A'$, that $k_i > k_i'$.

Now assume that $g(A) = g(A')$. The ith row of $g(A)$ must contain a contiguous block of k_i 1's and so must $g(A')$, since $g(A) = g(A')$. Now, since A' is stochastic and contiguous, it's ith row must contain a block of $\frac{1}{k_i}$'s in the same columns as $g(A')$, which means that A' must contain k_i non-negative entries in the ith row. This contradicts the assumption that A' contains $k_i' < k_i$ non-negative entries in the ith row.

Therefore, if $A \neq A'$, we have $g(A) \neq g(A')$, i.e., g is 1-1.

Q.E.D.

Lemma 4.3.2: Let A be an $n \times n$, square matrix with non-negative entries and let $g(A)$ be the corresponding 0-1 matrix.

Then $g(g(A)^k) = g(A^k)$ for any $k > 0$.

<u>Proof</u>: Let A and B be two, non-negative $n \times n$ matrices. We know that $g(A)$ and A both have non-zero entries in the same positions. The same applies for $g(B)$ and B. Let us consider the two products $g(A)\,g(B)$ and $A\,B$, and let us denote by α_{ij} and β_{ij} $i,j = 1,\ldots,n$ the entries of $A.B$ and $g(A)\,g(B)$, respectively. Then,

$$\beta_{ij} = \sum_{k=1}^{n} g(a)_{ik}\, g(b)_{kj} ,$$

is not zero if and only if for at least one k, $1 \le k \le n$ both $g(a)_{ik}$ and $g(b)_{kj}$ are non-zero. But then for the same k's, we have both a_{ik} and b_{kj} are non-zero, and so $a_{ik}\, b_{kj}$ is non-zero. Thus $\sum_{k=1}^{n} a_{ik} \cdot b_{kj} = \alpha_{ij}$ is non-zero and therefore β_{ij} is non-zero if and only if α_{ij} is, as well.

Now we shall show, by induction on k, that $g(g(A)^k) = g(A^k)$. For $k = 1$, it is obvious that $g(g(A)) = g(A)$. Assume that the statement is true for $k = n-1$, i.e., $g(g(A)^{n-1}) = g(A^{n-1})$. We shall show that it is true for $k = n$. From the equality $g(g(A)^{n-1}) = g(A^{n-1})$, we see that $g(g(A)^{n-1})$ and $g(A^{n-1})$ have non-zero entries in the same positions. But the same is true for the matrices $g(A)^{n-1}$ and A^{n-1}. From the definition of g, it is clear that g does not change any non-zero entries to zero. Then, by the above argument, $g(A)^{n-1}\, g(A)$ and $A^{n-1} A$ have non-zero entries in the same positions, i.e., $g(A)^n$ and A^n have non-zero entries in the same positions and so, by the definition of g, $g(g(A)^n) = g(A^n)$. Thus the equality is true and for $k = n$, and the induction is complete.

Q.E.D.

Lemma 4.3.3: Let $A \in A_n$. Then $g(A)$ is primitive.

Proof: Since A is primitive, there exists an integer $k > 0$ such that $A^k > 0$, i.e.,

A^k has non-zero entries $\Longleftrightarrow$

$g(A^k)$ has non-zero entries $\Longleftrightarrow$

$g(g(A^k))$ has non-zero entries $\Longleftrightarrow$

$g(A)^k$ has non-zero entries $\Longleftrightarrow$

$g(A)$ is primitive.

Q.E.D.

Lemma 4.3.3 tells us that instead of checking the primitivity of a non-negative matrix A it is equivalent to check the primitivity of $g(A)$ which is much easier from a computational point of view.

As we mentioned at the beginning of this section, the 0-1 matrix $g(A)$ can serve as the adjacency matrix for some graph G. Since the rows and columns of $g(A)$ correspond to an arbitrary labelling of the vertices of G, it follows that if $g(A)$ is an adjacency matrix of G, then so is $Pg(A)P^T$ where P is a $n \times n$ permutation matrix. Moreover, if $g(A)$ and $g(B)$ are any two adjacency matrices for G, there exists a permutation matrix P such that $g(B) = Pg(A)P^T$ [4]. Therefore it is clear that we shall be interested primarily in those properties of $g(A)$ which are invariant under permutations of the rows and columns. Since $Pg(A)P^T = g(PAP^T)$ for $A \in A_n$, we shall define a class of matrices $P_n \subset A_n$, which is closed under pre- and post-multiplication by permutation matrices. P_n well be the main object of our investigation in the sequel.

Definition 4.3.1 : Let $P_n \subset A_n$ be the class of matrices $A \varepsilon A_n$ with the property that $PAP^T \varepsilon A_n$ for any permutation matrix P.

The following results characterize those vectors π for which there exists a matrix $A \varepsilon P_n$ such that $\pi A = \pi$. We refer to such vectors as P_n-invariant vectors [8].

Lemma 4.3.4 [8]: Let $A \varepsilon P_n$ and let $G = (V,E)$ be the digraph associated with g(A). Then,

i) G contains a k-clique, K, for some $k \geq 1$, such that there exists an arc from each vertex of the k-clique to each vertex of G;

ii) if v is a vertex of G not in K, v has outdegree 1;

iii) the subgraph of G, $\overline{G}$, formed by deleting all arcs <v,w>, where v is a vertex of K, is a forest of k in-trees with roots in K.

Note: part (ii) simply says that the ith row of g(A), $i > k$ (for $k \geq 1$) contains only one non-zero entry.

Proof: We claim that each row of g(A) consists entirely of 1's or of a single 1. If there was a row, say the ith row, which contains a contiguous block of k_i 1's ($k_i \neq 1, n$), we could find a permutation matrix P such that one of the rows of $Pg(A)P^T$, say the jth, has non-contiguous entries of 1's. This implies that the jth row of $g(PAP^T)$ is non-contiguous and similarly the jth row of PAP^T is non-contiguous. This implies that $A \notin P_n$, which contradicts the hypothesis.

We now show that $k \neq 0$. If $k = 0$, then every row in $g(A)$ has only one non-zero entry. Since $g(A)$ is strongly connected, it must be a permutation matrix. Then $g(A)^m$ cannot be primitive for any m. Hence $A \notin A_n$, contradicting the hypothesis that $A \in P_n \subset A_n$.

We can, therefore, choose a permutation matrix P such that $Pg(A)$ has all its rows of 1's in the first k positions. Then, since $A \in P_n$, it follows that $PAP^T \in A_n$, and so $g(PAP^T)$ has all its rows of 1's in the first k positions (note that $g(PAP^T) = Pg(A)P^T$) .

Let K be the k-clique defined by the $k \times k$ matrix of 1's in the upper left corner of $g(PAP^T)$, so (i) and (ii) of the lemma are established.

Now, for part (iii) we claim that each weak component (maximal weakly connected subgraph) of the graph $\bar{G}$ corresponding to $g(PAP^T)$ with the first k rows replaced by zeros is an in-tree rooted at a vertex of K. It suffices to show that each weakly connected component has exactly one vertex of outdegree 0, since each vertex of $\bar{G}$ has outdegree of at most 1 and the only vertices of outdegree 0 in $\bar{G}$ are those of K.

If each vertex has outdegree 1, the component would clearly contain a cycle. Moreover, there would be no arc from the cycle to any vertex of K. This contradicts the fact that G is strongly connected. Therefore, each weakly connected component has at least one vertex of outdegree 0.

Suppose now that v and w are two vertices of outdegree 0 in the same weak component. Let $v = v_1, v_2, \ldots, v_n = w$ be a semiwalk in $\bar{G}$.

Since v has outdegree 0, $<v_2,v_1> \varepsilon E$. On the other hand, since w has outdegree 0, $<v_{n-1},v_n> \varepsilon E$. Let i be the smallest index such that $<v_i,v_{i+1}> \varepsilon E$. Clearly $1 < i < n$. Since $<v_{i-1},v_i> \notin E$, $<v_i,v_{i-1}> \varepsilon E$ and $<v_i,v_{i+1}> \varepsilon E$, and hence v_i has outdegree 2. Therefore $r = w$.

Now, since there are k vertices of outdegree 0 in G, there are k such in-trees.

Q.E.D.

The converse of Lemma 4.3.4 is also true. Before proving that, we note that the inverse mapping $g^{-1}: g(A_n) \to A_n$ is well-defined. This is true since, by Lemma 4.3.1, the map g is 1-1. Moreover, if $A \varepsilon g(P_n)$, then $g^{-1}(A)$ has rows consisting entirely of $\frac{1}{n}$'s or containing exactly one 1, and, to those rows from $g^{-1}(A)$ which consist entirely of $\frac{1}{n}$'s there correspond rows of A which consist entirely of 1's. This is obvious because $g^{-1}(A)$ is stochastic and contiguous.

Thus, if each row of A contains exactly one 1, then it follows that $g^{-1}(A) = A$.

<u>Lemma 4.3.5</u>: Let G be a graph with properties (i), (ii), (iii) of Lemma 4.3.4, and let A be an adjacency matrix for G. Then $g^{-1}(A) \varepsilon P_n$.

<u>Proof</u>: From (i) and (ii), it is clear that A has rows consisting entirely of 1's or containing exactly one 1; i.e., $g^{-1}(A)$ is stochastic and contiguous. Therefore, it suffices to show that $g^{-1}(A) \varepsilon A_n$. This will be true if $g^{-1}(A)$ is primitive, i.e., if G is strongly connected and there exists an aperiodic vertex (i.e., a loop).

Let v and w be any two vertices of G. If v is a vertex of K, we know from (i) that there is a path (namely an arc) from v to w. Otherwise, from (iii) we see that there exists a path to some vertex u of K, and this path can be extended to w since there is an arc $\langle u,w \rangle$ by (i). Hence, G is strongly connected.

Since $k \geq 1$, A has at least one row with no zero entries. Thus there exists a vertex i such that $a_{ii} > 0$, and so, A is primitive.

Q.E.D.

A Special Labelling Procedure

There are many ways of labelling a graph. We show the following labelling procedure [8].

Consider a graph G satisfying Lemma 4.3.4, and let $\overline{G}$ be the forest composed of k in-trees, numbered $T_1, T_2, \ldots, T_k$ where T_i has n_i vertices.

We label the root of T_1 as 1, and we let $2,3,\ldots,m_1$ label the vertices of the first level of T_1, starting from the left. Then $m_1+1, m_1+2, \ldots, m_2$ label the vertices of level 2 and so on, until all the n_1 vertices of T_1 are labelled. We label T_2 as n_1+1 and use the same procedure to label all the vertices of T_2. In general, the root of T_i is labelled as b_i+1 where $b_i = \sum_{j=1}^{i-1} n_j$, and its vertices are labelled, as above, consecutively, adding 1 to the label, as we proceed from left to right, along each level of T_i and then down to the next level, as shown in Figure 4.2.

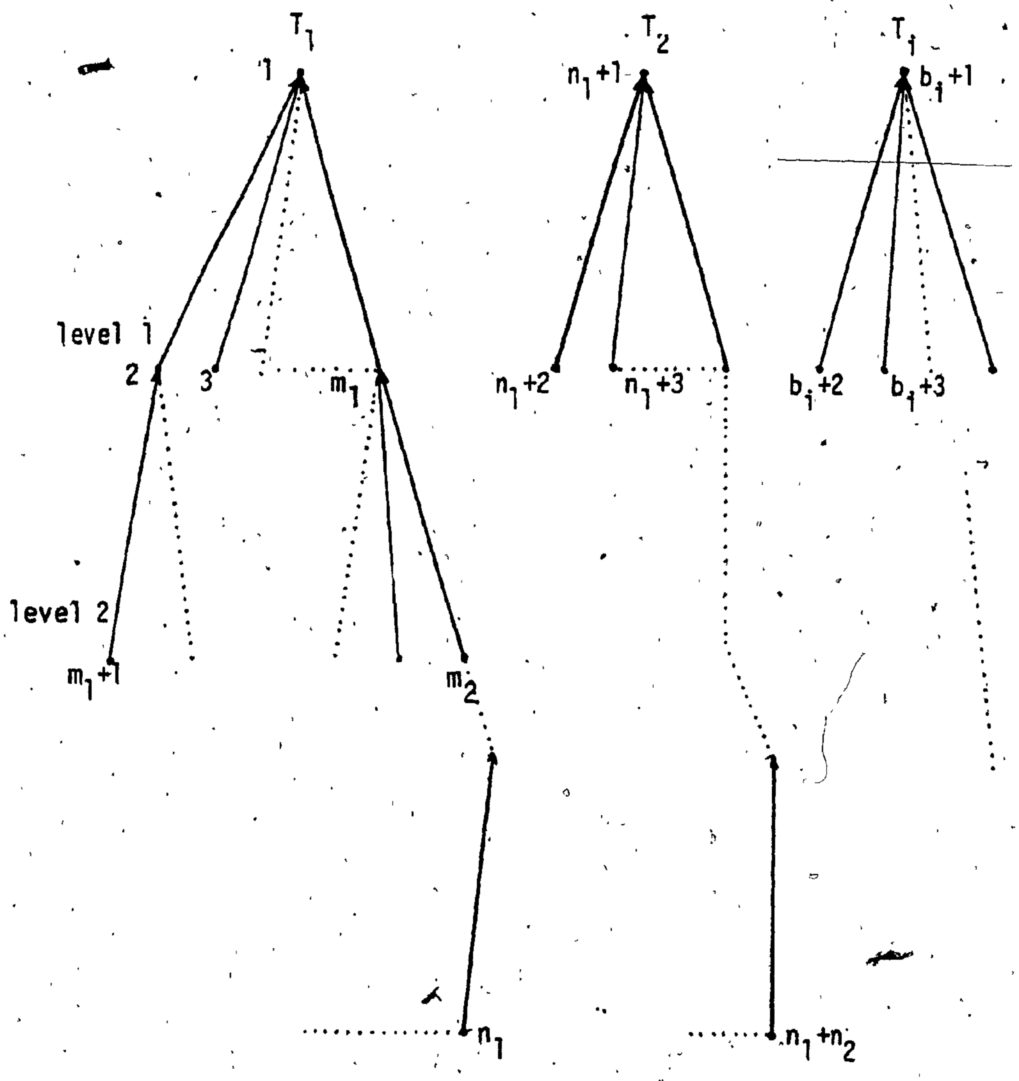


Fig. 4.2

The graph of the forest $\overline{G}$.

With this labelling scheme, the adjacency matrix for $\overline{G}$ is a block diagonal matrix of the form:

$$\overline{A} = \begin{bmatrix} A_1 & & & \\ & A_2 & 0 & \\ 0 & & \ddots & \\ & & & A_k \end{bmatrix},$$

where $A_i = (a^i_{rs})$ is the adjacency matrix for T_i, and is the form

$$A_i = \begin{bmatrix} 0 & & & & \\ 1 & & & & \\ \vdots & & & & \\ 1 & & & & \\ & 1 & & 0 & \\ & \vdots & & & \\ & 1 & & & \\ & & 1 & & \\ & & \vdots & & \\ & & 1 & \cdots\cdots & \\ & & 0 & & 1 \end{bmatrix}$$

The first row of A_i has zero entries only, since the first row of A_i corresponds to the root of the in-tree T_i, which has zero outdegree. Each other row of A_i has exactly one 1, since each other row of A_i corresponds to a vertex of the in-tree T_i, different from the root, and we know that each such vertex has outdegree 1. We can also see that if a vertex labelled β in T_i has indegree equal to t, and $\beta + p+1,\ldots,\beta + p+t$, $p < n_i$ are the labelled vertices in T_i, from which there starts an arc ending at the vertex β, then $a^i_{rs} = 1$ for $r = \beta+p+1,\ldots,\beta+p+t$.

Now let $B = \left\{1, 1+n_1, 1+n_1+n_2, \ldots, 1+n_1+n_2+\cdots+n_{k-1}\right\} = \left\{1 + \sum_{j=1}^{i-1} n_j\right\}_{i=1}^{k-1}$,

i.e., B contains the integers which are the numbers corresponding to the roots of the above forest $\overline{G}$ according to the above labelling procedure.

Let U be the $n \times n$ matrix, $n = n_1+n_2+\ldots+n_k$, in which the jth row, $j \in B$, consists entirely of 1's and all other rows are zero, i.e.:

$$U = \begin{bmatrix} 1 & 1 & 1 & \cdots\cdots\cdots & 1 \\ & & 0 & & \\ 1 & 1 & 1 & \cdots\cdots\cdots & 1 \\ & & 0 & & \\ & & & & \\ 1 & 1 & 1 & \cdots\cdots\cdots & 1 \\ & & 0 & & \end{bmatrix} \begin{matrix} \text{1th row} \\ \\ (1+n_1)\text{th row} \\ \vdots \\ \\ (1+n_1+\ldots+n_{k-1})\text{th row} \\ \\ \end{matrix}$$

Now, let $\overline{P} = \overline{A} + U$, i.e.:

$$
\overline{P} = \left[\begin{array}{cccccccc}
1 & 1 & 1 & \cdots & \cdots & \cdots & \cdots & 1 \\
1 & & & & & & & \\
\vdots & & & & & & & \\
1 & 1 & & & & & & \\
 & \vdots & & & 0 & & & \\
 & 1 & & & & & & \\
 & & 1 & & & & & \\
 & 0 & \vdots & & & & & \\
 & & 1 & & & & & \\
1 & 1 \cdots 1 & 1 & \cdots & \cdots & \cdots & \cdots & 1 \\
 & & 1 & & & & & \\
 & & \vdots & & 0 & & & \\
 & & 1 & & & & & \\
 & 0 & & 1 & & & & \\
 & & & \vdots & & & & \\
 & & & 1 & & & & \\
1 & 1 \cdots & \cdots & 1 & 1 & \cdots & \cdots & 1 \\
 & & & & 1 & & & \\
 & & & & \vdots & & & \\
 & 0 & & & & & 0 & \\
1 & 1 \cdots & \cdots & \cdots & \cdots & \cdots & 1 \cdots & 1 \\
 & & & & & & 1 & \\
 & & & & & & \vdots & 0 \\
 & & & & & & 1 & \\
 & & 0 & & & & & \\
 & & & & & & & 1
\end{array}\right]
\begin{array}{l}
\text{1th row} \\ \\ \\ \\ \\ \\ \\ \\ \\
(1+n_1)\text{th row} \\ \\ \\ \\ \\ \\ \\
(1+n_1+n_2)\text{th row} \\ \\ \\ \\
(1+n_1+\ldots+n_{k-1})\text{th row} \\ \\ \\ \\ \\
\text{nth row}
\end{array}
$$

From the labelling procedure of G and the construction of $\overline{P}$ it follows that $\overline{P}$ is an adjacency matrix of G.

Consider now the matrix $g^{-1}(\overline{P})$, where,

$$g^{-1}(\overline{P}) = \left[\begin{array}{cccccccc} \frac{1}{n} & \frac{1}{n} & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{1}{n} \\ 1 & & & & & & & \\ \vdots & & \ddots & & 0 & & & \\ 1 & & & & & & & \\ & 1 & & & & & & \\ \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} & \cdots & \cdots & \cdots & \frac{1}{n} \\ & & & 1 & & & & \\ & 0 & & & & 0 & & \\ & & & & & & & \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ & & & & & 1 & & \\ & & 0 & & & & & 1 \end{array}\right] \begin{array}{l} \text{1th row} \\ \\ \\ \\ \\ (1+n_1)\text{th row} \\ \\ \\ \\ (1+n_1+\ldots+n_{k-1})\text{th row} \\ \\ \\ \end{array}$$

By Lemma 4.3.5, $g^{-1}(\overline{P}) \in P_n$.

Characterization of P_n-invariant Vectors

For matrices of the form of $g^{-1}(\overline{P})$, it is not difficult to characterize the vectors π such that $\pi\, g^{-1}(\overline{P}) = \pi$. This leads to the following characterization of P_n-invariant vectors:

<u>Theorem 4.3.1</u>: [8] Let $\overline{G} = (V,E)$ be a forest of k rooted in-trees. Let c be a constant and let $w: V \to \mathbb{R}$ be a real valued function which assigns weights to the vertices of $\vec{G}$ as follows:

(i) if $v \in V$ has indegree 0, then $w(v) = c$.

(ii) if $v \in V$ has indegree $m \neq 0$ and $v_1, v_2, \ldots, v_m$ are the vertices from which there starts an arc ending at v, then

$$w(v) = c + \sum_{i=1}^{m} w(v_i) .$$

Suppose $v_1, v_2, \ldots, v_n$ is a labelling of the elements of V, then the vector $(w(v_1), w(v_2), \ldots, w(v_n))$ is P_n-invariant. Moreover, for any P_n-invariant vector π, there is a forest $\overline{G}$ and a labelling of the vertices such that $\pi = (w(v_1), w(v_2), \ldots, w(v_n))$.

<u>Proof</u>: Let us label the vertices of $\overline{G}$ following the above procedure, and let this be $v_1, v_2, \ldots, v_n$. Let $w(v)$ be the weight assigned to the vertex v, as defined in (ii) of the statement of the theorem, let $\overline{A}$ and $\overline{P}$ be the adjacency matrix for $\overline{G}$ and G, respectively, as constructed above. We claim that $\pi = (w(v_1), w(v_2), \ldots, w(v_n))$ satisfies $\pi \cdot g^{-1}(\overline{P}) = \pi$. To see this we assume the following:

a) Let v be a root of the T_ith in-tree of $\overline{G}$ possessing n_i vertices $1 \le i \le k$. Assume that with the above labelling procedure this in-tree has q levels, and let the jth level has h_j vertices, $j = 1,2,\ldots,q$. Then it is clear that $\sum_{j=1}^{q} h_j = n_i - 1$ and, by (ii), for the root v, we have

$$w(v) = c + \sum_{i=1}^{h_1} w(v_i) \quad ,$$

and

$$w(v_i) = c + \sum_{r=1}^{t_i} w(v_r) \quad ,$$

where t_i represents the number of vertices in level 2, from which there originates an arc ending at v_i. It is clear that $\sum_{i=1}^{h_1} t_i = h_2$.

Thus we have the following:

$$w(v) = c + \sum_{i=1}^{h_1} \left(c + \sum_{r=1}^{t_i} w(v_r)\right)$$

$$= c + h_1 c + \sum_{i=1}^{h_1} \sum_{r=1}^{t_i} w(v_r)$$

$$= c + h_1 c + \sum_{r=1}^{h_2} w(v_r') \quad ,$$

where v_r' is a new rearrangement of the elements of level 2.

Continuing in this way, we get:

$$w(v) = c + h_1 c + h_2 c + \ldots + h_{q-1} c + \sum_{e=1}^{h_q} w(v_e) \quad .$$

But since each vertex on the q level has indegree 0, by (i) we have $w(v_e) = c$. Therefore:

$$w(v) = c + h_1 c + h_2 c + \ldots + h_q c$$

$$= c + \left(\sum_{j=1}^{q} h_j\right) c$$

$$= c + (n_i - 1)c$$

$$= n_i c \quad .$$

Now, let $B = g^{-1}(\overline{P})$ and $D = g^{-1}(U)$, if A_i is an $n_i \times n_i$ matrix as defined above, since each row of A_i contains exactly one 1, then $g^{-1}(A_i) = A_i$. Therefore we have

$$B = g^{-1}(\overline{P})$$
$$= g^{-1}(\overline{A} + U)$$
$$= g^{-1}(\overline{A}) + g^{-1}(U)$$
$$= \overline{A} + D .$$

Now, let $\pi = (\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)})$, where $\pi^{(i)} = (w(v_{b_i+1}), w(v_{b_i+2}), \ldots, w(v_{b_i+n_i}))$, $i = 1,2,\ldots,k$. We note that the vector π has k components, i.e., a number equal to the number of roots (or in-trees) in the forest $\overline{G}$, and that each component $\pi^{(i)}$ has length n_i, i.e., equal to the number of vertices in the in-tree T_i. Therefore, it is clear that π has length $n = n_1 + n_2 + \ldots + n_k$. Then,

$$\pi B = \pi(\overline{A} + D)$$
$$= \pi \overline{A} + \pi D$$
$$= (\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)}) \overline{A} + \pi D$$
$$= (\pi^{(1)}A_1, \pi^{(2)}A_2, \ldots, \pi^{(k)}A_k) + \pi D$$
$$= (\pi^{(1)}A_1 + \pi D, \pi^{(2)}A_2 + \pi D, \ldots, \pi^{(k)}A_k + \pi D),$$

but,

$$\pi D = (\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)})\ g^{-1}(U)$$

$$= (w(v_{b_1+1}), w(v_{b_1+2}), \ldots, w(v_{b_1+n_1}), w(v_{b_2+1}), \ldots, w(v_{b_k+n_k}))g^{-1}(U)$$

$$= (w(v_1), w(v_2), \ldots, w(v_{n_1}), w(v_{n_1+1}), \ldots, w(v_n))g^{-1}(U)$$

$$= (w(v_1), \ldots, w(v_{1+n_1}), \ldots, w(v_{1+n_1+n_2}), \ldots, w(v_n)) \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots\cdots\cdots & \frac{1}{n} \\ & & 0 & \\ \frac{1}{n} & \frac{1}{n} & \cdots\cdots\cdots & \frac{1}{n} \\ & & 0 & \\ \frac{1}{n} & \frac{1}{n} & \cdots\cdots\cdots & \frac{1}{n} \\ & & \vdots & \\ & & \vdots & \\ & & 0 & \end{bmatrix}$$

$$= w(v_1)\frac{1}{n} + 0\ldots0 + w(v_{1+n_1})\frac{1}{n} + 0\ldots0 + \ldots + w(v_{1+n_1+\ldots+n_{k-1}})\frac{1}{n} + 0\ldots+0$$

$$= \frac{1}{n}w(v_{b_1+1}) + \frac{1}{n}w(v_{b_2+1}) + \ldots + \frac{1}{n}w(v_{b_k+1})$$

$$= \frac{1}{n}\sum_{i=0}^{k} w(v_{b_i+1})$$

$$= \frac{1}{n}\sum_{i=0}^{k} n_i\ c$$

$$= c$$

Thus, we obtain,

$$\pi B = (\pi^{(1)}A_1+c,\ \pi^{(2)}A_2+c, \ldots,\ \pi^{(k)}A_k+c) . \tag{4.1}$$

b) Let v_r be a vertex of T_i with indegree 0. Then

$$\pi_r = w(v_r) = c ,$$

and

$$(\pi B)_r = c + (\pi^{(i)}A_i)_{r-b_i} , \quad \text{from (4.1)} ,$$

$$= c + \sum_{j=1}^{n_i} \pi_j^{(i)} a^i_{j,r-b_i}$$

$$= c .$$

Since v_r has indegree zero and so $a_{j,r-b_i} = 0$, j, $j = 1,2,\ldots,n_i$.

Hence,

$$(\pi B)_r = \pi_r .$$

c) Let v_r be a vertex of T_i, with indegree m, and let $v_{i+1},\ldots,v_{i+m}$ be the vertices from which there starts an arc ending at v_r. Then,

$$\pi_r = w(v_r)$$

$$= c + \sum_{\ell=1}^{m} w(v_{i+\ell}) \quad \text{by (ii)} .$$

Also, in view of (4.1),

$$(\pi B)_r = c + (\pi^{(i)}A_i)_{r-b_i}$$

$$= c + \sum_{j=1}^{n_i} \pi_j^{(i)} a^i_{j,r-b_i}$$

$$= c + \sum_{\ell=1}^{m} \pi^{(i)}_{i+\ell-b_i} \cdot 1$$

$$= c + \sum_{\ell=1}^{m} w(v_{i+\ell})$$

$$= \pi_r \ .$$

Hence,

$$(\pi B)_r = \pi_r \ .$$

Therefore, the claim is established.

Now let π' be a vector of weights corresponding to a different labelling of the vertices. Then it is obvious that the elements of π' are a permutation of those of π, so there exists a permutation matrix P, such that,

$$\pi' = \pi P$$

Thus,

$$\begin{aligned} \pi' &= \pi P \\ &= \pi B P \\ &= \pi P P^T B P \\ &= \pi' P^T B P \ , \end{aligned}$$

and, since $B \in P_n$ (note $B = g^{-1}(\overline{P})$), $P^T BP \in P_n$. Therefore from $\pi' P^T BP = \pi'$, we deduce that π' is P_n-invariant.

Conversely, let π' be a P_n-invariant vector and let $B' \in P_n$ be a matrix such that $\pi'B' = \pi'$. We know that there exists a permutation matrix P such that $PB'P^T = B$ is of the form $g^{-1}(\overline{P})$. Let $\pi = \pi'P^T$ and $c = \frac{1}{n} \sum_{i=1}^{k} \pi_{b_i+1}$. Accordingly if v_r is a vertex of $\overline{G}$ with indegree m and $v_{i+1},\ldots,v_{i+m}$ vertices of $\overline{G}$ from which there starts an arc ending at v_r, then

$$\pi_r = c + \sum_{\ell=1}^{m} w(v_{i+\ell})$$

$$= c + \sum_{\ell=1}^{m} \pi_{i+\ell} .$$

Therefore, π consists of weights as described in (ii) and $\pi = \pi P^T$ is a permutation of those weights.

Q.E.D.

Examples

Example 4.2:

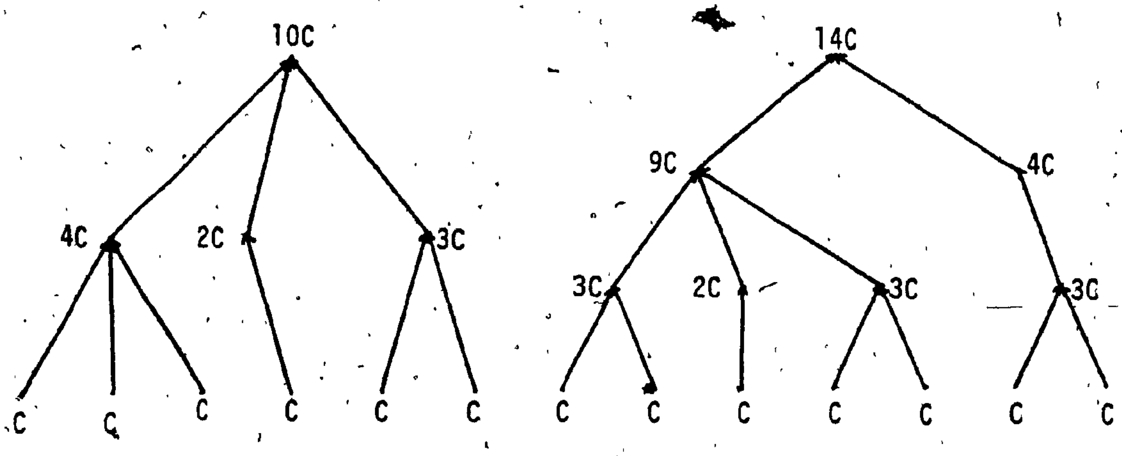


Fig. 4.3.

Fig. 4.3 shown the weight assignment for a forest consisting of two in-trees.

Example 4.3:

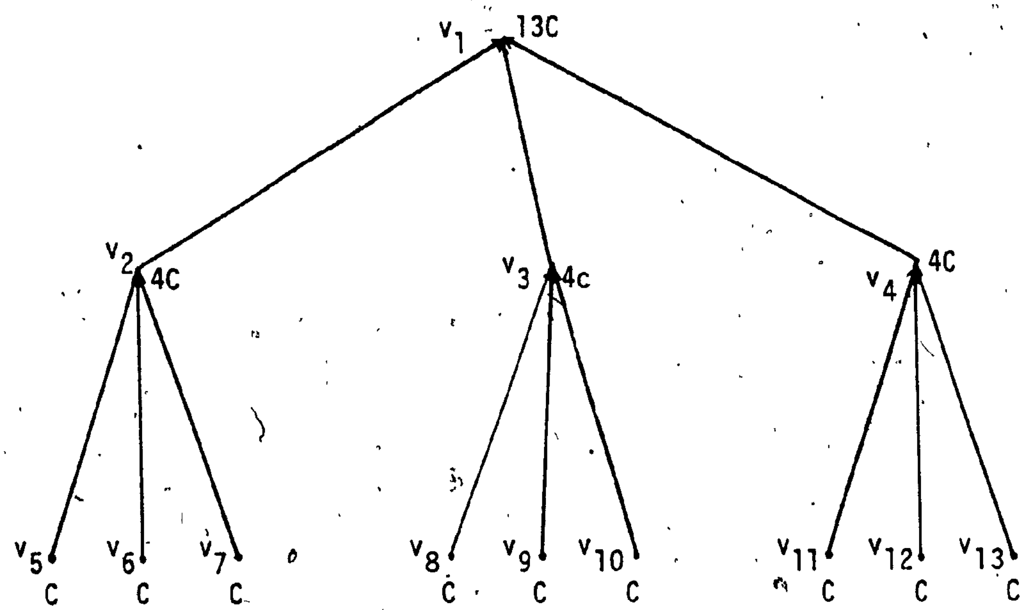


Fig. 4.4

Fig. 4.4 shows the labelling of the vertices of an in-tree and the weight assignment of each vertex.

$$B = \begin{bmatrix} \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\pi = (13,\ 4,\ 4,\ 4,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1).$$

It can be easily shown that $\pi B = \pi$, i.e., that the vector π is P_n-invariant.

<u>Example 4.4.</u>:

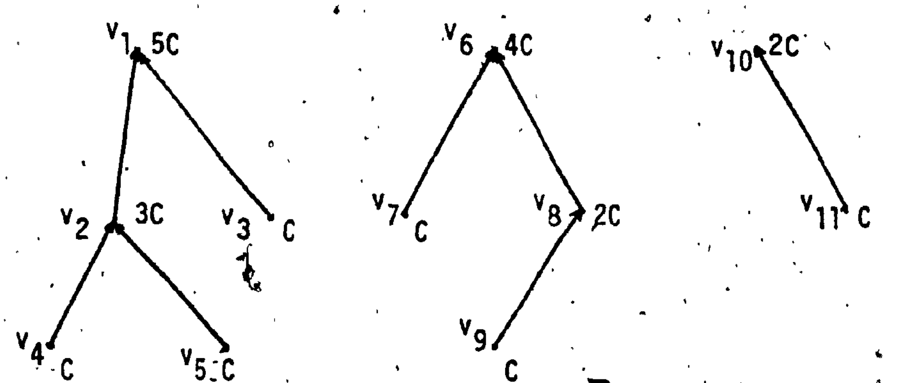


Fig. 4.5

$$B = \begin{bmatrix} \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$\pi = (5, 3, 1, 1, 1, 4, 1, 2, 1, 2, 1)$, and $\pi B = \pi$

<u>Example 4.5:</u>

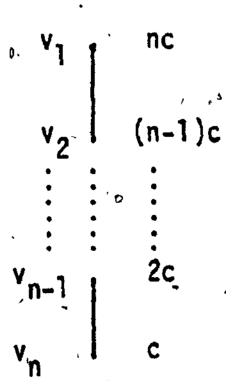
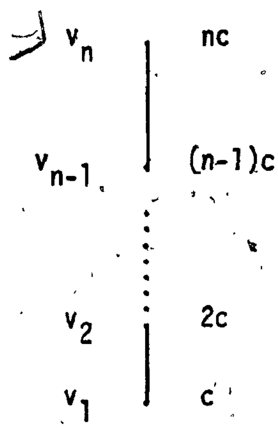


Fig. 4.6

$$B = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \cdots & \frac{1}{n} \\ 1 & & & & \\ & 1 & & 0 & \\ & & \ddots & & \\ & 0 & & 1 & \\ & & & 1 & 0 \end{bmatrix}$$

$\pi = (n, n-1, \ldots, 2, 1)$ and $\pi B = \pi$

Example 4.6: Since the vector $\pi = (n, n-1, \ldots 2, 1)$ is P_n-invariant, $\pi' = P\pi$ is also P_n-invariant, where P is a permutation matrix. Therefore, the vector $\pi = (1, 2, \ldots, n-1, n)$ is P_n-invariant, with the corresponding forest $\overline{G}$ and the matrix B' given by,



$$B' = \begin{bmatrix} 0 & 1 & & & \\ & & 1 & & 0 \\ & & & \ddots & \\ & & 0 & & \\ & & & & 1 \\ \frac{1}{n} & \frac{1}{n} & \cdots & \cdots & \frac{1}{n} \end{bmatrix}$$

Fig. 4.7

4.4. EXTENSION OF P_n-INVARIANT VECTORS

It is known (example 4.6) that the vector $\pi = (1, 2, \ldots, n)$ is P_n-invariant, and therefore it is the unique invariant density of a transformation τ which can be constructed. Suppose now, that we need a vector π with the jth term equal to mj rather than j, where m is a positive integer. We shall see that it is possible to construct a new τ for which this is the case.

Theorem 4.4.1: Let n, m and j be fixed positive integers satisfying $1 < m < n$, $m \le j+1 < n$. Then the n-vector

$$\pi = (1, 2, \ldots, m-1, j-(m-1), 2j-(m-2), \ldots, (m-1)j-1, mj, j+1, \ldots, n)$$

is A_n-invariant, i.e., there exists a matrix $A \in A_n$ such that

$$\pi A = \pi \, .$$

<u>Proof</u>: Let us construct the matrix $A = (a_{ij})$ as follows:

$$a_{i,i+1} = 1 \quad , \quad i \neq j$$

$$a_{j,k} = \begin{cases} \frac{1}{m} & , \quad j-(m-2) \le k \le j+1 = j-(m-(m+1)) \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$a_{n,s} = \frac{1}{n} \quad , \quad s = 1,2,\ldots,n \, .$$

i.e.,

$$A = \begin{bmatrix} 0 & 1 & & & & & \\ & & \ddots & & 0 & & \\ & & & 1 & & & \\ & & \frac{1}{m} \cdots\cdots & \frac{1}{m} & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \\ \frac{1}{n} & \frac{1}{n} & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \frac{1}{n} \end{bmatrix}$$

First, we shall show that $A \in A_n$;

(i). By construction, A is contiguous.

(ii) A is stochastic, since, for $k = n$,

$$\sum_{i=1}^{n} a_{n,i} = \sum_{i=1}^{n} \frac{1}{n} = 1 ,$$

for $k \neq j$,

$$\sum_{i=1}^{n} a_{k,i} = a_{k,k+1} = 1,$$

and for $k = j$,

$$\sum_{i=1}^{n} a_{ji} = \sum_{i=j-(m-2)}^{j+1} \frac{1}{m} = \sum_{t=1}^{m} \frac{1}{m} = 1.$$

(iii) To show that A is primitive, we consider the matrix B',

where B' is the matrix associated with π' as constructed in example 4.6. It is known that B' is primitive.

Clearly, A has a non-zero entry corresponding to each non-zero entry of B' (plus others in the jth row). Hence, g(A) is strongly connected and possesses an aperiodic vertex. Thus, A is primitive.

We claim now that $\pi A = \pi$. Let $m \leq j$. Then, for $i = 1$,

$$\sum_{r=1}^{n} \pi_r a_{ri} = \pi_n \frac{1}{n} = 1 = \pi_1 .$$

For $1 < i \leq j-(m-1)$,

$$\begin{aligned}\sum_{r=1}^{n} \pi_r a_{r,i} &= \pi_{i-1} a_{i-1,i} + \pi_n \frac{1}{n} \\ &= \pi_{i-1} + 1 \\ &= i-1 + 1 = i = \pi_i .\end{aligned}$$

For $i = j-(m-k)$, $2 \le k \le m$, i.e., for $j-(m-2) \le i \le j-(m-m) = j$,

$$\sum_{r=1}^{n} \pi_r a_{r,j-(m-k)} = \pi_{j-(m-(k-1))} a_{j-(m-(k-1)),j-(m-k)} + \pi_j a_{j,j-(m-k)} + \pi_n \frac{1}{n}$$

$$= [(k-1)j-(m-(k-1))] + mj \frac{1}{m} + 1$$

$$= kj-(m-k)$$

$$= \pi_{j-(m-k)}$$

For $j+1 \le i \le n$,

$$\sum_{r=1}^{n} \pi_r a_{r,i} = \pi_{i-1} a_{i-1,i} + \pi_n \frac{1}{n}$$

$$= \pi_{i-1} + 1$$

$$= i-1 + 1 = i = \pi_i .$$

Thus, we have shown that $\pi A = \pi$.

Q.E.D.

Corollary 4.4.1: Let n,m be fixed positive integers and let $j = m-1$. Then,

$$\pi = (m, 2m, 3m, \ldots, (m-1)m, m, m+1, \ldots, n)$$

is A_n-invariant.

Proof: Let $m = j+1$ in Theorem 4.4.1. Then the vector $\pi = (1,2,\ldots,j-(m-1),2j-(m-2),\ldots,mj,j+1,\ldots,n)$ becomes $\pi = (m,2m,\ldots,(m-1)m,m,m+1,\ldots,n)$.

Q.E.D.

Example 4.7: Let n = 6, m = 3, j = 4. Then the vector $\pi = (1,2,7,12,5,6)$ is A_n-invariant, where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

Example 4.8: Let n = 6, m = 4, j = m-1 = 3. Then the vector $\pi = (4,8,12,4,5,6)$ is A_n-invariant, where

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

Example 4.9: Letting m = 2, Theorem 4.4.1 proves that the vector $\pi = (1,2,\ldots,j-1,2j,\ j+1,\ldots,n)$ is A_n-invariant where,

$$A = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & 0 & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & \frac{1}{2} & \frac{1}{2} & & & \\ & & & & & 1 & & \\ & 0 & & & & & \ddots & \\ & & & & & & & 1 \\ \frac{1}{n} & \frac{1}{n} & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{1}{n} \end{bmatrix} \quad j\text{th row}$$

<u>Corollary 4.4.2</u>: Let $\pi = (\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)})$, where $\pi^{(i)}$ is a vector of length n_i and is obtained from the vector $(k_i, k_i+1, \ldots, k_i+n_i-1)$, where $k_i = \sum_{j=1}^{i-1} n_j$, by multiplying k_i+j_i by m_i, where $1 < m_i < n_i$, $m_i \le j_i < n_i$. Then the vector π is A_n-invariant.

<u>Proof</u>: The construction of Theorem 4.4.1 can be applied to each of the rows k_i+j_i of A. Since none of the modified entries overlap, the proof goes through exactly as in Theorem 4.4.1.

Q.E.D.

4.5. A SPECIAL CONSTRUCTION

In this section, we discuss an interesting and useful subclass of A_n-invariant vectors.

Theorem 4.5.1: Let $m,n,k \in N$ such that $m = \frac{n}{k}$. Then the vector $\pi = (m,m+1,\ldots,n)$ is A_n-invariant.

Proof: Let us construct a square matrix A with $(k-1)m+1 = n-(m-1)$ rows, as follows:

for $0 \le p \le k-2$,

$$a_{pm+1,j} = \frac{1}{(p+1)n} \quad , \quad k-p \le j \le k + (p+1)(m-1)$$

$$a_{pm+r,j} = \begin{cases} 1 & , \quad 2 \le r \le m,\ j = k + p(m-1) + r-1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

for $p = k-1$

$$a_{pm+1,j} = a_{(k-1)m+1,j} = \frac{1}{k} \quad , \quad 1 \le j \le k .$$

i.e.,

$$
\begin{array}{c}
1 \ldots k \ldots m \ldots k{+}m{-}1 \ldots k{+}2m{-}2 \ldots\ldots (k{-}1)m{+}1 \leftarrow \text{columns}
\end{array}
$$

$$
A = \left[\begin{array}{ccccccc}
\frac{1}{m} & \cdots\cdots & \frac{1}{m} & & & & \\
& 1 & & & & & \\
0 & & & & 0 & & \\
& & & 1 & & & \\
\frac{1}{2m} & \cdots\cdots & \frac{1}{2m} & \cdots\cdots & \frac{1}{2m} & & \\
& & & & 1 & & \\
& 0 & & & & & \\
& & & & & 1 & \\
\frac{1}{3m} & \cdots\cdots\cdots & & & \frac{1}{3m} & \cdots\cdots & \frac{1}{3m} \\
& & & & & & 1 \\
& & 0 & & & & \\
\frac{1}{k} & \cdots & \frac{1}{k} & & & &
\end{array}\right]
\begin{array}{l}
\text{1th row} \\ \\ \\ \\ \text{mth row} \\ \\ \\ \\ \text{2mth row} \\ \\ \\ \text{[(k-1)m+1]th row}
\end{array}
$$

A is clearly contiguous. It is also stochastic, since for $i = pm+1$, $0 \le p \le k-2$,

$$\sum_{j=1}^{(k-1)m+1} a_{ij} = \sum_{j=1}^{(k-1)m+1} \frac{1}{(p+1)m} = \sum_{j=k-p}^{k+(p+1)(m-1)} \frac{1}{(p+1)m} = \sum_{j=1}^{(p+1)m} \frac{1}{(p+1)m} = 1,$$

for $i = (k-1)m+1$,

$$\sum_{j=1}^{(k-1)m+1} a_{ij} = \sum_{j=1}^{k} \frac{1}{k} = 1,$$

for $i = pm+r$, where $0 \le p \le k-2$ and $2 \le r \le m$,

$$\sum_{j=1}^{(k-1)m+1} a_{ij} = a_{pm+r,k+p(m-1)+r-1} = 1.$$

Thus A is stochastic. We shall now show that A is primitive. To do this we shall prove that the graph G, which corresponds to $g(A)$ is strongly connected and has an aperiodic index.

a) We claim that for all $i \neq \ell$, there exists a path f[illegible] vertex v_i to the vertex v_ℓ, where $\ell = (k-1)m+1$. It is clear that for each $i \neq \ell$, there exists $j > i$ such that $g(a)_{ij} = \overline{P}_{ij} = 1$, since for each i, the ith row of A contains at least one non-zero entry.

Let $f(i)$ be the smallest such j. Similarly for that $f(i)$ (assuming that $f(i) \neq \ell$, since where $f(i) = \ell$ the same result is true), there exists $j > f(i)$ such that $\overline{P}_{f(i),j} = 1$. Let $f^2(i)$ be the smallest such j. Continuing the process, we have $i, f(i), f^2(i), \ldots, f^k(i)$, which defines a path, starting at i and visits vertices of strictly increasing number. Since there are a finite number of vertices, the sequence must terminate. Moreover, it is clear that the sequence terminates when $f^k(i) = \ell$.

b) Since there are arcs from vertex number ℓ to vertices $1,2,\ldots,k$, the set of vertices $v^{(-1)} = \{1,\ldots,k,\ell\}$ is strongly connected. From (a), there exists a path from all other vertices into this set (specifically ℓ). Therefore, it suffices to show that there are paths from this set to the remaining vertices.

Let $v^{(s)} = \{k+s(m-1)+1, k+s(m-1)+2, \ldots, k+s(m-1)+m-1\}$, where $0 \le s \le k-2$. Now, there exists an arc from vertex 1 to each vertex in

$v^{(o)} = \{k+1, k+2, \ldots, k+m-1\}$ (which is obvious from the matrix A). Therefore, $v^{(-1)} \cup v^{(o)}$ is strongly connected. We proceed by induction. Suppose $v^{(-1)} \cup v^{(o)} \cup \ldots \cup v^{(i-1)}$, $i < k-3$, is strongly connected. We claim that $\bigcup_{j=-1}^{i} v^{(j)}$ is strongly connected. To prove this, we note that there are arcs from the vertex $im+1$ to all vertices in $v^{(i)}$. Therefore, it suffices to show that $im+1 \in \bigcup_{j=-1}^{i-1} v^{(j)}$. By the construction, this is true if $im+1 \leq k+i(m-1)$ (i.e., if $i \leq k-1$), which is certainly true.

Now, we must show that $\overline{P} = g(A)$ has an aperiodic vertex. For $k > 2$, P can take on the value $k-2$. Hence, from

$$\overline{P}_{pm+1,j} = \frac{1}{(p+1)m}, \quad k-p \leq j \leq k+(p+1)(m-1),$$

we obtain

$$\overline{P}_{(k-2)m+1,j} = \frac{1}{(k-1)m}, \quad 2 \leq j \leq k+(k-1)(m-1) .$$

Therefore, the $[(k-2)m+1]$th row of $\overline{P}$ intersects the diagonal since $k+(k-1)(m-1) \geq (k-2)m+1$ for $k > 2$. Hence, there exists an i such that $\overline{P}_{ii} > 0$. Consequently $\overline{P}$ is primitive, and so is A.

To complete the proof, we have to show that $\pi A = \pi$, for $j = 1$

$$\sum_{i=1}^{(k-1)m+1} \pi_i a_{ij} = \pi_{(k-1)m+1}\, a_{(k-1)m+1,j}$$

$$= n \cdot \frac{1}{k} = m = \pi_1 ,$$

for $2 \le j \le k$,

$$\sum_{i=1}^{(k-1)m+1} \pi_i a_{i,j} = \sum_{p=o}^{k-2} \pi_{pm+1} a_{pm+1,j} + \frac{n}{k}$$

$$= \sum_{p=k-j}^{k-2} \pi_{pm+1} \frac{1}{(p+1)m} + m \quad \text{(since } k-p \ge j \text{ or } p \ge k-j\text{)}$$

$$= \sum_{p=k-j}^{k-2} \frac{(p+1)m}{(p+1)m} + m$$

$$= j-1 + m$$

$$= \pi_j ,$$

for $(k+1) + s(m-1) \le j \le k + (s+1)(m-1)$, where $0 \le s \le k-2$,

$$\sum_{i=1}^{(k-1)m+1} \pi_i a_{i,j} = \sum_{p=s}^{k-2} \pi_{pm+1} a_{pm+1,j} + \pi_{j-k+s+1} a_{j-k+s+1,j}$$

$$= \sum_{p=1}^{k-s-1} 1 + \pi_{j-k+s+1} a_{j-k+s+1,j}$$

$$= (k-s-1) + (j-k+s+m)$$

$$= m+j-1$$

$$= \pi_j .$$

Hence, $\pi A = \pi$. This completes the proof that π is A_n-invariant.

Q.E.D.

<u>Remark 1</u>: For $k=2$, the matrix A becomes an $(m+1) \times (m+1)$ matrix of the form.

$$A = \begin{bmatrix} \frac{1}{m} & \frac{1}{m} & \cdots\cdots & \frac{1}{m} \\ & 1 & & \\ & & 1 \quad 0 & \\ & 0 & & \\ & & & 1 \\ \frac{1}{2} & \frac{1}{2} & & \end{bmatrix}$$

where the 1's are on the superdiagonal. It is clear that the (m+1)th entries on the diagonals of A^2 and A^3 are greater that 0. Thus, the set $\{\eta \geq 1: a^{(n)}_{m+1,m+1} > 0\} \supset \{2,3\}$, and therefore, g.c.d $\{n \geq 1: a^{(n)}_{m+1,m+1} > 0\}$ is equal to 1. Hence, the (m+1)th vertex is aperiodic. The graph of $g(\bar{A})$ is strongly connected. Therefore $\overline{P}$ is primitive and so is A.

Remark 2: For $k = n$, the matrix A becomes a skew triangular matrix,

$$A = \begin{bmatrix} & & & 1 \\ & & \frac{1}{2} & \frac{1}{2} \\ & 0 \quad \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ & & & \\ \frac{1}{n} & \cdots\cdots & & \frac{1}{n} \end{bmatrix}$$

Clearly, $A \neq P_n$. On the other hand, the matrix obtained for the same vector using the construction of section 4.3 is

$$A' = \begin{bmatrix} & 1 & & & \\ & & \ddots & & 0 \\ & 0 & & \ddots & \\ & & & & 1 \\ \frac{1}{n} & \cdots & \cdots & \cdots & \frac{1}{n} \end{bmatrix}$$

where the 1's are on the superdiagonal, i.e., $A' \in P_n$.

Example 4.10: Let $m = 4$, $n = 12$. Then we have $k = 3$ and $\dot{\pi} = (4,5,6,7,8,9,10,11,12)$. The vector π is A_n-invariant, where the associated matrix A is given by.

$$\begin{bmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We can generalize Theorem 4.5.1. This generalization is given as a corollary.

<u>Corollary 4.5.1</u>: Let $m,n,j \in N$ be fixed, such that $n = m+\ell_j$, j divides m (j/m), and $m = \frac{n}{k}$ (i.e., $(k-1)m = \ell_j$). ~~Then the vector~~

$$\pi = (m, m+j, m+2j, \ldots, n)$$

is A_n-invariant.

<u>Proof</u>: We construct the $[(k-1)\frac{m}{j} + 1] \times [(k-1)\frac{m}{j} + 1]$ matrix A in a directly analogous manner to that in Theorem 4.5.1. Namely, as follows.

For $0 \le p \le k-2$:

$$a_{p\frac{m}{j}+1,s} = \begin{cases} \frac{1}{(p+1)m} & , \ k-p \le s \le (p+1)(\frac{m}{j}-1) + k \\ 0 & , \ \text{otherwise} \end{cases}$$

$$a_{p\frac{m}{j}+r,s} = \begin{cases} 1 & , \ \text{if } j = k+p(\frac{m}{j}-1) + r-1 \\ 0 & , \ \text{otherwise} \end{cases} ,$$

where $2 \le r \le \frac{m}{j}$.

For $p = k+1$

$$a_{(k-1)\frac{m}{j}+1,s} = \begin{cases} \frac{1}{k} & , \ 1 \le s \le k \\ 0 & , \ \text{otherwise} \end{cases}$$

The proof that $A \in A_n$ is analogous to the proof of this fact in Theorem 4.5.1.

Q.E.D.

Theorem 4.6.2: Let π and σ be A_n-invariant vectors of lengths n_1 and n_2 respectively, such that $\pi A^{(1)} = \pi$ and $\sigma A^{(2)} = \sigma$. Suppose that the following two conditions hold, for some i and j:

(i) $a^{(1)}_{i,n_1} = \frac{1}{x}$, $a^{(2)}_{j,1} = \frac{1}{y}$, $x,y \in N$, $x \le n_1$, $y \le n_2$.

(ii) $\pi_i = kx$, $\sigma_j = ky$, where $k \in N$.

Then (π, σ) is A_n-invariant.

Proof: First we note that since $A^{(1)}$ and $A^{(2)}$ belong to A_n, $A^{(1)}$ and $A^{(2)}$ are both stochastic and contiguous. Then from (i) we have,

$$a^{(1)}_{i,n_1} = \frac{1}{x} \text{ implies that } a^{(1)}_{i,s} = \frac{1}{x} \text{ for } n_1-x+1 \le s \le n_1 ,$$

and

$$a^{(2)}_{j,1} = \frac{1}{y} \text{ implies that } a^{(2)}_{j,s} = \frac{1}{y} \text{ for } 1 \le s \le y .$$

Let us define the matrix A as follows:

$$A = \begin{bmatrix} \bar{A}^{(1)} & D^{(1)} \\ D^{(2)} & \bar{A}^{(2)} \end{bmatrix}$$

where,

$$\bar{a}^{(1)}_{r,s} = a^{(1)}_{r,s},\ \forall\ r \neq i,\ \bar{a}^{(1)}_{i,s} = \begin{cases} \frac{1}{x+y} & , \quad n_1-x+1 \le s \le n_1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$d^{(1)}_{r,s} = 0 \qquad ,\ \forall\ r \neq i,\ d^{(1)}_{i,s} = \begin{cases} \frac{1}{x+y} & , \quad 1 \le s \le y \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$\bar{a}^{(2)}_{r,s} = a^{(2)}_{r,s} \qquad ,\ \forall\ r \neq j\ ,\ \bar{a}^{(2)}_{j,s} = \begin{cases} \frac{1}{x+y} & , \quad 1 \le s \le y \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$d^{(2)}_{r,s} = 0 \qquad ,\ \forall\ r \neq j\ ,\ d^{(2)}_{j,s} = \begin{cases} \frac{1}{x+y} & , \quad n_1-x+1 \le s \le n_1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

Let $(\pi \cdot \sigma)A = (\rho,\mu)$, where ρ and μ are of lengths n_1 and n_2 respectively. Then,

$$\rho = \sum_{r=1}^{n_1} \pi_r\, \bar{a}^{(1)}_{rs} + \sum_{r=1}^{n_2} \sigma_r\, d^{(2)}_{rs}$$

$$= \begin{cases} \sum_{r=1}^{n} \pi_r\, a^{(1)}_{rs} = \pi_s & , \quad \text{if} \quad s \notin [n_1-x+1,\ n_1] \\ \sum_{\substack{r \neq i \\ r \neq j}} \pi_r\, a^{(1)}_{rs} + \pi_i \bar{a}^{(1)}_{is} + \pi_j \bar{a}^{(1)}_{js} + \sigma_j d^{(2)}_{js} & , \quad \text{if} \quad s \in [n_1-x+1,\ n_1] \end{cases}$$

$$= \begin{cases} \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \notin [n_1-x+1,\, n_1] \\ \sum_{r\neq i,j} \pi_r \, a_{rs}^{(1)} + \pi_j \, a_{js}^{(1)} + \frac{\pi_j}{x+y} + \frac{\sigma_j}{x+y} & , \quad \text{if } s \in [n_1-x+1,\, n_1] \end{cases}$$

$$= \begin{cases} \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \notin [n_1-x+1,\, n_1] \\ \sum_{r\neq i,j} \pi_r \, a_{rs}^{(1)} + \pi_j \, a_{js}^{(1)} + \frac{kx+ky}{x+y} & , \quad \text{if } s \in [n_1-x+1,\, n_1] \end{cases}$$

$$= \begin{cases} \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \notin [n_1-x+1,\, n_1] \\ \sum_{r\neq i} \pi_r \, a_{rs}^{(1)} + k & , \quad \text{if } s \in [n_1-x+1,\, n_1] \end{cases}$$

$$= \begin{cases} \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \notin [n_1-x+1,\, n_1] \\ \sum_{r\neq i} \pi_r \, a_{rs}^{(1)} + \frac{\pi_i}{x} & , \quad \text{if } s \in [n_1-x+1,\, n_1] \end{cases}$$

$$= \begin{cases} \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \notin [n_1-x+1,\, n_1] \\ \sum_{r\neq i} \pi_r \, a_{rs}^{(1)} + \pi_i \, a_{is}^{(1)} & , \quad \text{if } s \in [n_1-x+1,\, n_1] \end{cases}$$

$$= \begin{cases} \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \notin [n_1-x+1,\, n_1] \\ \sum_{r=1}^{n_1} \pi_r \, a_{rs}^{(1)} = \pi_s & , \quad \text{if } s \in [n_1-x+1,\, n_1] \end{cases}$$

$$= \pi_s \quad \text{for each} \quad s,\ 1 \le s \le n_1 .$$

Similarly, we can prove that $\mu_s = \sigma_s$ for each $s, 1 \le s \le n_2$. Therefore, we have $(\rho,\mu) = (\pi\ \sigma)$, i.e.,

$$(\pi,\sigma)A = (\pi,\sigma).$$

Now, from the definition of A, we can see that A is stochastic and contiguous. To see that A is primitive, we consider the matrix $\overline{A} = g(A)$ and the corresponding graph $G = g(A)$. We label the vertices of $g(A)$ by $a_1,\ldots,a_{n_1}$ and $b_1,\ldots,b_{n_2}$ as in the proof of Theorem 4.6.1. Since $g(\overline{A}^{(1)}) = g(A^{(1)})$ and $g(\overline{A}^{(2)}) = g(A^{(2)})$, in order to show that $g(A)$ is strongly connected, it suffices to show that for r, $1 \le r \le n_1$ and for s, $1 \le s \le n_2$, there are paths $a_r \to b_s$ and $b_s \to a_r$. We omit this, since it is analogous to the proof of Theorem 4.6.1.

Finally, let $Q^{(n)}$ represent the upper left $n_1 \times n_1$ block of A^n. Since $g(\overline{A}^{(1)}) = g(A^{(1)})$, it is clear that $Q^{(n)} \ge (A^{(1)})^n$ and hence $\overline{A}$ has one aperiodic state. Thus, A is primitive. Therefore, (π,σ) is A_n-invariant.

Q.E.D.

<u>Example 4.13</u>: Let $\pi = (1, 2, 3, 4, 5)$ and $\sigma = (1, 2, 3, 4)$ be A_n-invariant, with

$$A^{(1)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}, \quad \text{and} \quad A^{(2)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

since $a_{55}^{(1)} = \frac{1}{5}$, and $a_{41}^{(2)} = \frac{1}{4}$, we have $i = 5$, $j = 4$, $x = 5$, $y = 4$. The condition (ii) is satisfied for $k = 1$. Then the vector $(\pi,\sigma) = (1, 2, 3, 4, 5, 1, 2, 3, 4)$ is A_n-invariant with A given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix}$$

We can extend this example to any two vectors $\pi^{(1)}$ and $\pi^{(2)}$, where the n_ith entry of $\pi^{(i)}$ is n_i and the n_ith row of $A^{(i)}$ consists of equal entries $\frac{1}{n_i}$, $i = 1,2$. Therefore, this example applies to all vectors of section 4.4.

4.7. PERMUTATION RESULTS

Many of the matrices constructed in the preceding sections are not in P_n (for example, see section 4.5), i.e. they are not permutation invariant. It is of interest, however, to investigate the extent to which these vectors can be permuted.

Let $A \varepsilon A_n$ and let P be a permutation matrix. It is known that post-multiplication by P (i.e., PA) rearranges the rows of A and so post-multiplication does not have any effect on the contiguity. On the other hand, pre-multiplication by P^T, rearranges the columns of A and this may destroy the contiguity. We also know that, given any $A \varepsilon A_n$, PAP^T is the matrix which corresponds to a relabelling of the vertices of $g(A)$.

The following two results show that certain permutations are admissable.

<u>Theorem 4.7.1</u>: $\pi = (\pi_1, \ldots, \pi_n)$ is A_n-invariant if and only if $\pi' = (\pi_n, \pi_{n-1}, \ldots, \pi_1)$ is, as well.

<u>Proof</u>: Consider the skew-diagonal matrix:

$$P = \begin{bmatrix} & & & 1 \\ & 0 & 1 & \\ & \cdot & & \\ & & 0 & \\ 1 & 1 & & \end{bmatrix}$$

and assume that π is A_n-invariant with A the matrix induced by π, i.e., $\pi A = \pi$. Then $(PA)P^T$ just reverses the order of the columns of PA. Thus the contiguity of PA still stands and therefore, $PAP^T \in A_n$. Note that $\pi' = \pi P^T$, and so, we have,

$$\pi' PAP^T = \pi P^T PAP^T = \pi AP^T = \pi P^T = \pi' .$$

Hence, π' is A_n-invariant. Similarly, we can show that π is A_n-invariant if π' is.

Q.E.D.

<u>Theorem 4.7.2</u>: Suppose that $\pi A = \pi$, where $A \in A_n$ is an $n \times n$ matrix which can be written in block form as (A_1, A_2, A_3) and A_i is an $n \times p_i$ matrix with $o \le p_i \le n$,

$\sum_{i=1}^{3} p_i = n$. Let A_2 have the property that all its rows contain either 1 or p_2 non-zero elements. Then,

$$\pi' = (\pi_1, \ldots, \pi_{p_1}, \sigma_1, \ldots, \sigma_{p_2}, \pi_{p_1+p_2+1}, \ldots, \pi_n)$$

is A_n-invariant, where $(\sigma_1, \ldots, \sigma_{p_2})$ is a permutation of $(\pi_{p_1+1}, \ldots, \pi_{p_2})$.

<u>Proof</u>: Let P_2 be a $p_2 \times p_2$ permutation matrix such that,

$$(\pi_{p_1+1}, \ldots, \pi_{p_2}) P_2^T = (\sigma_1, \ldots, \sigma_{p_2}) .$$

Let us define

$$P = \begin{bmatrix} I_1 & & 0 \\ & P_2 & \\ 0 & & I_3 \end{bmatrix} ,$$

where I_i is the $p_i \times p_i$ identity matrix $i = 1,3$. Then, $PAP^T \varepsilon A_n$, $\pi P^T = \pi'$ and

$$\pi' PAP^T = \pi P^T PAP^T = \pi AP^T = \pi P^T = \pi' ,$$

i.e., π' is A_n-invariant.

Q.E.D.

CHAPTER V

MARKOV MAPS AND IRREDUCIBLE (OR PRIMITIVE) MATRICES

5.1. PRELIMINARIES

As we have seen, the irreducibility or the primitivity of a non-negative matrix has been essential to our work. Therefore, it is interesting to know if a given matrix is irreducible (primitive) or not. Direct verification for large matrices (i.e., using the definition of an irreducible (primitive) matrix) can be very time consuming, even on fast computers. For that reason, we would like to relate the irreducibility (primitivity) of large matrices to the irreducibility (primitivity) of smaller ones. In this chapter we shall analyze this idea using Markov maps.

Until now, we are familiar with Markov maps from [0,1] into itself, which are piecewise linear and belong to the class $\mathcal{T}$. We showed (Theorem 2.1.2) that each $\tau \in \mathcal{T}$, τ piecewise linear, determines uniquely a non-negative square matrix, which by Lemma 2.2.3 is irreducible. In this chapter we consider the following class of Markov maps.

<u>Definition 5.1.1</u>: Let $\mathcal{J} = \{I_i\}_{i=1}^n$ be a partition of I (I any interval), and let $\tau: I \to I$ be a Markov map with respect to $\mathcal{J}$. We say that τ is in the class W if:

(i) τ is piecewise C^2 with respect to $\mathcal{J}$

(ii) $\inf_x |\tau'(x)| > 0$.

It follows that τ is piecewise monotonic with respect to $\mathcal{J}$

(Lemma 2.1.2) and nonsingular. Obviously $\mathcal{T} \subset W$.

In section 5.2 we define the transition matrix, B_τ, for each Markov map and we study the relations between M_τ and B_τ for $\tau \in \mathcal{T}$, τ piecewise linear.

Finally in section 5.3, we consider the following problem: Let $\mathcal{J}_1 = \{I_i\}_{i=1}^{n_1}$ be a Markov partition with respect to τ of the interval $I = [a,b]$, let $\mathcal{J}_2 = \{K_i\}_{i=1}^{n_2}$ be any other partition of I (assume $n_1 < n_2$). Let $g(B_\tau^{(1)})$ and $g(B_\tau^{(2)})$ be the 0-1 matrices (where $B_\tau^{(1)}$ and $B_\tau^{(2)}$ are the transition matrices and g the function defined in 4.3) induced by $\mathcal{J}_1$ and $\mathcal{J}_2$, respectively. Suppose that $g(B_\tau^{(1)})$ is irreducible (primitive). Under what conditions on τ will $g(B_\tau^{(2)})$ be irreducible (primitive)? We shall show that if τ satisfies a certain easily verifiable condition, then the above question is answered in the affirmative [7], in other words, that the irreducibility (primitivity) of $g(B_\tau^{(1)})$ implies the irreducibility (primitivity) of $g(B_\tau^{(2)})$.

5.2. MATRIX REPRESENTATION OF A MARKOV MAP

Let us consider an interval $I = [a,b]$ of the real line, and let us take a finite partition $\mathcal{J} = \{I_i\}_{i=1}^{n}$ of I. Let $\tau: I \to I$ be a Markov map with respect to the partition $\mathcal{J}$. We define the transition matrix $B = B_\tau$ for the Markov map τ as follows:

$$b_{ij} = \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_i)} \qquad i,j = 1,2,\ldots,n ,$$

where m is the Lebesgue measure on I. Clearly b_{ij} is the proportion of the interval I_i which is mapped onto the interval I_j.

It is known (Theorem 2.1.2) that when τ is piecewise linear and belongs to the class $\mathcal{T}$, then τ induces a matrix $M = M_\tau$, which is defined as follows:

$$m_{ij} = \begin{cases} |\tau_i'|^{-1} = \dfrac{m(I_i)}{m(\tau_i(I_i))} , & \text{if } I_j \subset \tau(I_i) \\ 0 , & \text{otherwise} . \end{cases}$$

Thus, for a piecewise linear transformation $\tau \in \mathcal{T}$ we can construct two matrices, namely $B = B_\tau$ and $M = M_\tau$. In the following we shall study the relations between B_τ and M_τ.

In general B_τ and M_τ are not equal. We can see that noticing that all the non-zero entries of M_τ on the ith row are contiguous and equal to $|\tau_i'|^{-1} = \dfrac{m(I_i)}{m(\tau_i(I_i))}$, but the non-zero entries of B_τ on the same ith row are not necessarily equal since $I_i \cap \tau^{-1}(I_j)$ varies, and so $m(I_i \cap \tau^{-1}(I_j))$, as I_j varies. We can see that easily from the following example:

<u>Example 5.1</u>: Let $\tau: [0,1] \rightarrow [0,1]$ be a piecewise linear transformation defined by $\tau(0) = 0$, $\tau(\frac{1}{4}^-) = \frac{1}{2}$, $\tau(\frac{1}{4}^+) = 0$, $\tau(\frac{1}{2}^-) = 1$, $\tau(\frac{1}{2}^+) = 0$, $\tau(1) = 1$, and shown in Fig. 5.1.

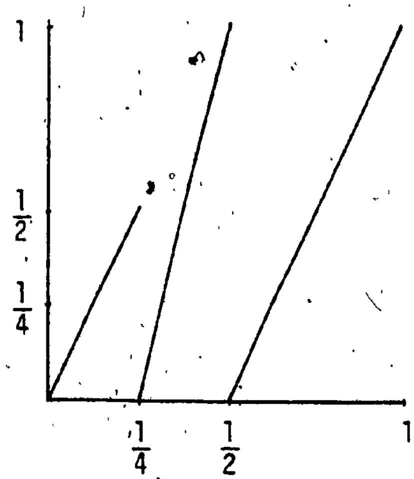


Fig. 5.1

It is easy to see that

$$M_\tau = \begin{bmatrix} .5 & .5 & 0 \\ .25 & .25 & .25 \\ .5 & .5 & .5 \end{bmatrix}$$

and

$$B_\tau = \begin{bmatrix} .5 & .5 & 0 \\ .25 & .25 & .5 \\ .25 & .25 & .5 \end{bmatrix}$$

Obviously $M_\tau \neq B_\tau$.

With the following Lemma we shall show that under certain conditions $M_\tau = B_\tau$

Lemma[1] 5.2.1: Let $\tau \in \mathcal{T}$, be a piecewise linear transformation of I onto I, $I = [a,b]$. The transition matrix $B = B_\tau$ and the induced matrix $M = M_\tau$ are precisely equal when the partition $\mathcal{J}$ of I has all subinterval of equal length.

Proof: We have from definitions of B_τ and M_τ that

$$b_{ij} = \frac{m(I_i \cap \tau_j^{-1}(I_j))}{m(I_i)}, \quad i,j = 1,2,\ldots,n$$

and

$$m_{ij} = \begin{cases} \dfrac{m(I_i)}{m(\tau_i(I_i))}, & \text{if } I_j \subset \tau_i(I_i) \\ 0, & \text{otherwise} \end{cases}.$$

We consider two cases.

Case 1: Let $m_{ij} = 0$. Then $I_j \not\subset \tau_i(I_i)$, which implies that $I_j \cap \tau_i(I_i) = \emptyset$ or $\tau_j^{-1}(I_j) \cap I_i = \emptyset$. Thus, $m(\tau_j^{-1}(I_j) \cap I_i) = m(\emptyset) = 0$. Therefore, $b_{ij} = 0$. Conversely, let $b_{ij} = 0$. Then $m(I_i \cap \tau_j^{-1}(I_j)) = 0$. This implies that $I_i \cap \tau_j^{-1}(I_j) = \emptyset$, since the intersection of two open intervals of the real line is $\emptyset$ or an open interval, and the measure of an open interval cannot be zero; equivalently, $\tau(I_i) \cap I_j = \emptyset$. Thus, $I_j \not\subset \tau(I_i)$, and so $m_{ij} = 0$. Therefore, the matrices B_τ and M_τ, both have the same zero entries.

---

[1] All the lemmas in that section are due to the author.

Case 2: Let i be arbitrary, and $m_{ij} \neq 0$ for some $j, j = 1,2,\ldots,n$. Let the length of each subinterval be denoted by Δ, i.e., $m(I_i) = \ell(I_i) = a_i - a_{i-1} = \Delta$. ($m(I_i) = \ell(I_i)$ since m is the Lebesque measure). Suppose the Markov map τ maps I_i onto $I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}$. Then

$$\tau_i(I_i) = I_t \cup I_{t+1} \cup \ldots \cup I_{t+k} ,$$

and $m(\tau_i(I_i)) = (k+1)\Delta$. Thus we get that

$$m_{ij} = \frac{m(I_i)}{m(\tau_i(I_i))} = \frac{\Delta}{(k+1)\Delta} = \frac{1}{K+1} , \; j = t, \; t+1, \ldots, t+k . \tag{5.1}$$

Now from $\tau_i(I_i) = I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}$, we get:

$$I_i = \tau^{-1}(I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}) .$$

Let $j = t + \ell$, $0 \le \ell \le k$. Then $\tau^{-1}(I_j) = \tau^{-1}(I_{t+\ell})$, and

$$\begin{aligned} I_i \cap \tau^{-1}(I_j) &= \tau_i^{-1}(I_t \cup I_{t+1} \cup \quad \cup I_{t+k}) \cap \tau^{-1}(I_{t+\ell}) \\ &= \tau^{-1}(I_{t+\ell}) . \end{aligned}$$

Thus,

$$\begin{aligned} m(I_i \cap \tau^{-1}(I_j)) &= m(\tau^{-1}(I_{t+\ell})) \\ &= \frac{m(I_{t+\ell})}{|\tau_i'|} \\ &= \frac{\Delta}{k+1} \end{aligned}$$

Therefore,

$$b_{ij} = \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_j)} = \frac{\frac{\Delta}{k+1}}{\Delta} = \frac{1}{k+1}\ , \quad j=t,t+1,\ldots,t+k, \quad (5.2)$$

and from (5.1) and (5.2) we get that $m_{ij} = b_{ij}$ for i arbitrary and $j = t, t+1,\ldots,t+k$. This completes the proof.

Q.E.D.

Example 5.2: Let us consider the Example 2.1. We have

$$M_\tau = \begin{bmatrix} 0 & 0 & .5 & .5 \\ .25 & .25 & .25 & .25 \\ 0 & .5 & .5 & 0 \\ .25 & .25 & .25 & .25 \end{bmatrix} .$$

Using the Definition of B_τ we find that

$$B_\tau = \begin{bmatrix} 0 & 0 & .5 & .5 \\ .25 & .25 & .25 & .25 \\ 0 & .5 & .5 & 0 \\ .25 & .25 & .25 & .25 \end{bmatrix} .$$

Therefore $M_\tau = B_\tau$.

Remark: From the proof of Lemma 5.2.1 it is clear that M_τ and B_τ both have the same zero entries not only in the case where we consider partition $\mathcal{J}$ with all subintervals of equal length, but also in the case where $\mathcal{J}$ is any arbitrary Markov partition. Thus M_τ and B_τ both have the same non-zero entries. Therefore, we can state the following lemma.

<u>Lemma 5.2.2</u>: Let $\tau \in \mathcal{T}$, τ piecewise linear, and let $g(B)$ and $g(M)$ be the 0-1 matrices induced (see section 4.3) by B_τ and M_τ, respectively. Then $g(B) = g(M)$.

<u>Proof</u>: We must show that $g(B)$ and $g(M)$ both have the same non-zero entries. Let us denote by C and T the matrices $g(B)$ and $g(M)$ respectively.

Let $c_{ij}=1$, for some i,j, $1 \le i,j \le n$, then $b_{ij} \neq 0$ or $m(I_i \cap \tau^{-1}(I_j)) \neq 0$, which implies that $I_i \cap \tau^{-1}(I_j) \neq \emptyset$ or $I_j \cap \tau(I_i) \neq \emptyset$. This implies that $I_j \subset \tau(I_i) = \tau_i(I_i)$, since any other relation between I_j and $\tau(I_i)$ contradicts the fact that τ takes partition points to partition points, i.e., $I_j \subset \tau_i(I_i)$. Thus $t_{ij} = 1$. Therefore, $C = T$ or $g(B) = g(M)$.

Q.E.D.

It is clear that the matrix B is irreducible (primitive) if and only if the matrix $g(B)$ is irreducible (primitive). Thus we can prove the following result.

<u>Lemma 5.2.3</u>: The transition matrix B_τ, of a piecewise linear transformation $\tau, \tau \in \mathcal{T}$, is irreducible matrix.

<u>Proof</u>: To show that B_τ is irreducible we shall show that $g(B_\tau)$ is. From Lemma 5.2.2 we have that $g(B_\tau) = g(M_\tau)$. Therefore, if we show that $g(M_\tau)$ is irreducible the result is proved. But $g(M_\tau)$ is irreducible since M_τ is in view the Lemma 2.2.3. Hence B_τ is irreducible.

Q.E.D.

5.3. IRREDUCIBILITY AND PRIMITIVITY OF TRANSITION MATRICES

Let us consider a Markov map $\tau \in W$, $\tau: I \to I$ with respect to the partition $\mathcal{J}_1 = \{I_i\}_{i=1}^n$, and let $\mathcal{J}_2 = \{K_j\}_{j=1}^m$ be any other partition of I (not necessarily Markov with respect to τ). We say that the condition (A) is satisfied if for each $K_j \in \mathcal{J}_2$ there exist m_j and some i, $1 \le i \le n$ such that:

$$\tau^{m_j}(K_j) \supset I_i \qquad (A)$$

Theorem 5.3.1: Let $\mathcal{J}_1$ be a Markov partition of I with respect to τ, $\tau \in W$, and let $\mathcal{J}_2$ be any other partition of I. Let $g(B_\tau^{(1)})$ and $g(B_\tau^{(2)})$ be the 0-1 matrices induced by $\mathcal{J}_1$ and $\mathcal{J}_2$, respectively. If condition (A) is satisfied, then,

(i) $g(B_\tau^{(1)})$ irreducible implies that $g(B_\tau^{(2)})$ is irreducible, and

(ii) $g(B_\tau^{(1)})$ primitive implies that $g(B_\tau^{(2)})$ is primitive.

Proof: Let $E = g(B_\tau^{(1)})$ and $D = g(B_\tau^{(2)})$.

(i) Let E be irreducible. Then, by definition (2.1.4) for each pair (i,j) there exists t such that $e_{ij}^t > 0$. We can see that $e_{ij} > 0$, if and only if $m(I_i \cap \tau^{-1}(I_j)) > 0$, or if and only if $I_i \cap \tau^{-1}(I_j)$ contains an open interval, or if and only if $\tau(I_i) \cap I_j$ contains an open interval. Similarly, $e_{ij}^t > 0$ if and only if $m(I_i \cap \tau^{-t}(I_j)) > 0$ and thus, if and only if $\tau^t(I_i) \cap I_j$ contains an open interval.

So it suffices to show that for all i,j there exists an m such that $\tau^m(K_i) \cap K_j$ contains an open interval. By condition (A), there exists an m_i and p, $1 \le p \le n$ such that $\tau^{m_i}(K_i) \supset I_p$. Now, $K_j \cap I_\ell \ne \emptyset$ for some ℓ. By the irreducibility of E, there is an m' such that $\tau^{m'}(I_p) \supset I_\ell$. Then

$$\tau^{m'+m_i}(K_i) \supset \tau^{m'}(I_p) \supset I_\ell \supset K_j \cap I_\ell \neq \emptyset .$$

Therefore, $\tau^{m'+m_i}(K_i) \cap K_j = \tau^m(K_i) \cap K_j \neq \emptyset$, i.e., contains an open interval.

(ii) Let E be primitive. Then exists a k such that $e_{ij}^k > 0$ for all i,j, which means that $m(I_i \cap \tau^{-k}(I_j)) > 0$, or that $I_i \cap \tau^{-k}(I_j)$ contains an open interval for each i and j, which means that $\tau^k(I_i) = I$ for each i, the same is true for $k+1$, i.e., $\tau^{k+1}(I_i) = I$ for each i. Let $m' = m+m_j$. Then by condition (A), we have,

$$\tau^{m_j}(K_j) \supset I_i ,$$

or

$$\tau^m(\tau^{m_j}(K_j)) \supset \tau^m(I_i)$$

or

$$\tau^{m'}(K_j) \supset I = \tau^m(I_i) , \tag{5.3}$$

and since,

$$\tau^{m'}(K_j) \subset I \tag{5.4}$$

for each K_j, $j = 1,2,\ldots,n$,

$$\tau^{m'}(K_j) = I .$$

Moreover,

$$\tau^{m'+1}(K_j) = I .$$

It follows that $D^{m'}$ and $D^{m'+1}$ both have non-zero entries in the jth position on the diagonal. Hence D is primitive.

Q.E.D.

Let $J_1 = \{I_i\}_{i=1}^n$ be a fixed partition of $I = [0,1]$ and let $\tau \in W$. We define

$$\mu = \min_{1 \le i \le n} m(I_i) \quad \text{and} \quad \sigma = \max_{1 \le i \le n} m(I_i)$$

Let $c = \mu/\sigma$. Clearly $0 < c \le 1$. When the partition consists of equal intervals, $c = 1$.

Let $\tau_i = \tau\big|_{I_i}$ and let ℓ_i be the straight line joining the endpoints of $\tau_i(x)$. Define,

$$\beta_i = \inf_{x \in I_i} |\tau_i'(x)|$$

and

$$\delta_i = \frac{\beta_i}{\ell_i'} .$$

Obviously, $0 < \delta_i \le 1$. If τ_i is linear, then $\delta_i = 1$.

<u>Lemma 5.3.1</u>: Let K be an interval contained in I_p, where $\tau(I_p) = \bigcup_{j=1}^{q} I_{p'+j}$, for some p' and q. Then

$$1 \ge \frac{m(\tau(K))}{\sum_{j=1}^{q} m(I_{p'+j})} \ge \delta_p \, \eta(K) ,$$

where,

$$\eta(K) = \frac{m(K)}{m(I_p)} .$$

<u>Proof</u>: The first inequality is obvious, since $\tau(K) \subset \tau(I_p)$ implies that $m(\tau(K)) \le m(\tau(I_p)) = m(\bigcup_{j=1}^{q} I_{p'+j}) = \sum_{j=1}^{q} m(I_{p'+j})$.

To show the second inequality, let us consider the straight line ℓ_p joining the endpoints of $\tau(I_p)$. Then,

$$\ell_p' = \frac{m(\tau(I_p))}{m(I_p)} = \frac{m(\tau(K))}{m(K)} .$$

Since τ is monotone, we have

$$\ell_p' \geq \beta_k = \inf_{x\in K}|\tau'(x)| \geq \beta_p = \inf_{x\in I_p}|\tau'(x)| .$$

Then

$$m(\tau(K)) = \ell_p' \, m(K) \geq \beta_p \, m(K) . \tag{5.5}$$

Now, since

$$\delta_p = \frac{\beta_p}{\ell_p'} = \beta_p \Big/ \frac{m(\tau(I_p))}{m(I_p)} = \frac{\beta_p \, m(I_p)}{m(\tau(I_p))}$$

or

$$\delta_p = \frac{\beta_p \, m(I_p)}{m(\bigcup_{j=1}^{q} I_{p'+j})} = \frac{\beta_p \, m(I_p)}{\sum_{j=1}^{q} m(I_{p'+j})} ;$$

we obtain

$$\beta_p = \frac{\delta_p}{m(I_p)} \sum_{j=1}^{q} m(I_{p'+j}) .$$

Thus, using (5.5), we get

$$m(\tau(K)) \geq \frac{\delta_p}{m(I_p)} \sum_{j=1}^{q} m(I_{p'+j}) \, m(K) ,$$

and

$$\frac{m(\tau(K))}{\sum_{j=1}^{q} m(I_{p'+j})} \geq \frac{\delta_p}{m(I_p)} m(K) = \delta_p \, \eta(K) .$$

Q.E.D.

Remark: Since $K \subset I_p$, $m(K) \leq m(I_p)$. This means that $\eta(K) \leq 1$.

Corollary 5.3.1: If $\tau(I_p) = I_{p'}$, then

$$\eta(\tau(K)) \geq \delta_p \, \eta(K) .$$

Proof: $\tau(I_p) = I_{p'}$ implies that $q = 0$. Then $m(\tau(I_p)) = m(I_{p'})$, and so

$$\begin{aligned} \eta(\tau(K)) &= \frac{m(\tau(K))}{m(\tau(I_{p'}))} \\ &= \frac{m(\tau(K))}{m(I_{p'})} \\ &\geq \delta_p \, \eta(K) \sum_{j=1}^{q} m(I_{p'+j})/m(I_{p'}) \\ &= \delta_p \, \eta(K). \end{aligned}$$

Q.E.D.

Definition 5.3.1: A Markov partition $\mathcal{J}$ is said to be a strong Markov partition if the transition matrix B, is irreducible and not a permutation matrix.

Theorem 5.3.2: Let $\mathcal{J}_1 = \{I_i\}_{i=1}^{n}$ be a strong Markov partition with respect to τ, $\tau \in W$, and let $\mathcal{J}_2 = \{K_j\}_{j=1}^{m}$ be any other partition of I. Then condition (A) is satisfied if

$$\delta_1 \delta_2 \ldots \delta_n (1+c) > 1 .$$

Proof: We first define a sequence $K^{(1)}, K^{(2)}, \ldots, K^{(j)}, \ldots$ of intervals contained in the elements of J_1 such that,

$$K^{(i+1)} \subset \tau(K^{(i)}),$$

i.e.,

$$K^{(j)} \subset \tau(K^{(j-1)}) \subset \tau^2(K^{(j-2)}) \subset \ldots \subset \tau^{j-1}(K^{(1)}).$$

We shall show that the corresponding sequence $\{n(K^{(i)})\}$ attains the value 1 in a finite number of steps, because if $n(K^{(i)}) = 1$ for some i, then $\frac{m(K^{(i)})}{m(I_i)} = 1$, or $K^{(i)} = I_i$, and since $K^{(i)} \subset \tau^{i-1}(K^{(1)})$, we have

$$I_i \subset \tau^{i-1}(K^{(1)}).$$

Therefore, condition (A) is satisfied.

Let $K \in J_2$ and let $I' \in J_1$ be such that $K \cap I' \neq \emptyset$. Let $K^{(i)} = K \cap I'$. Assuming that $K^{(i)} \subsetneq I_p$ (if $K^{(i)} = I_p$ we are through), we define $K^{(i+1)}$ as follows.

Case 1: $\tau: I_p \to I_{p'}$. Then $K^{(i+1)} = \tau(K^{(i)}) \subset I_{p'}$. Note that, $\tau(K^{(i)}) \subsetneq \tau(I_p)$, since $|\tau_p'| > 0$. So $K^{(i+1)} \neq I_{p'}$.

Case 2: $\tau: I_p \to \bigcup_{j=1}^{q} I_{p'+j}$. There are three possibilities.

(1) $\tau(K^{(i)}) \subset I_{p'+\ell}$ for some $1 \le \ell \le q$. Then $K^{(i+1)} = \tau(K^{(i)}) \subseteq I_{p'+\ell}$. In this case, either $K^{(i+1)} = I_{p'+\ell}$ and we terminate the sequence with $n(K^{(i+1)}) = 1$ or $K^{(i+1)} \subsetneq I_{p'+\ell}$.

(2) $\tau(K^{(i)}) \subseteq I_{p'+\ell} \cup I_{p'+\ell+1}$ and $\tau(K^{(i)})$ intersects both intervals. Let $\Delta^{(k)} = \tau(K^{(i)}) \cap I_{p'+\ell+k}$, $k = 0$ or 1. If for $k = 0$ or 1, $\Delta^{(k)} = I_{p'+\ell+k}$, let $K^{(i+1)} = \Delta^{(k)}$, and we terminate the sequence with $\eta(K^{(i+1)}) = 1$. (If this condition holds for both $\Delta^{(o)}$ and $\Delta^{(1)}$, we choose any one for $K^{(i+1)}$.) Otherwise, let $K^{(i+1)} = \Delta^{(k)}$ where $\eta(\Delta^{(k)}) \geq \eta(\Delta^{(k')})$, $k' \neq k$.

(3) $\tau(K^{(i)}) \subset \bigcup_{s=-1}^{t} I_{p'+\ell+s}$ for some $1 < \ell < q$, $1 \leq t \leq q-\ell$. Then $I_{p'+\ell} \subseteq \tau(K^{(i)})$, since τ is piecewise continuous. Let $K^{(i+1)} = I_{p'+\ell}$; then the sequence is terminated with $\eta(K^{(i+1)}) = 1$.

We now obtain, from Case 1 and 2 (1), (2), (3) a set of inequalities.

Case 1: By the Corollary of Lemma 5.3.1, we have

$$\eta(\tau(K^{(i)})) = \eta(K^{(i+1)}) \geq \delta_p \, \eta(K^{(i)}) . \tag{5.6}$$

Case 2: (1) By Lemma 5.3.1, we have,

$$\delta_p\ \eta(K^{(i)}) \le \frac{m(\tau(K^{(i)}))}{\sum_{j=1}^{q} m(I_{p'+j})}$$

$$= \frac{m(K^{(i+1)})}{m(I_{p'+\ell})+\mu}$$

$$\le \frac{\frac{m(K^{(i+1)})}{m(I_{p'+\ell})}}{1+\frac{\mu}{m(I_{p'+\ell})}}$$

$$\le \frac{\eta(K^{(i+1)})}{1+c}$$

Thus,

$$\delta_p(1+c)\ \eta(K^{(i)}) \le \eta(K^{(i+1)}) \le 1\ ,$$

where the last inequality is obtained from the remark of Lemma 5.3.1.

Case 2: (2) Since $\tau(K^{(i)}) \subset \Delta^{(o)} \cup \Delta^{(1)}$, we have from Lemma 5.3.1,

$$\delta_p\ \eta(K^{(i)}) \le \frac{m(\Delta^{(o)})}{\sum_{j=1}^{q} m(I_{p'+j})} + \frac{m(\Delta^{(1)})}{\sum_{j=1}^{q} m(I_{p'+j})}$$

$$\le \eta(\Delta^{(o)}) + \eta(\Delta^{(1)})$$

$$\le 2\eta(K^{(i+1)})\ .$$

Thus,

$$\frac{\delta_p}{2}\ \eta(K^{(i)}) \le \eta(K^{(i+1)}) \le 1\ .$$

It ~~is~~ clear from this construction that the sequence $\{K^{(i)}\}$ terminates only if there exists p such that $K^p \in J_1$. Then, condition (A) is satisfied.

Now we shall show that the sequence $\{K^{(i)}\}$ must terminate. First, we shall see that if the sequence is infinite, then Case 2 (2) cannot occur more than once. Suppose that it occurs twice, i.e., $K^{(i)}$ and $K^{(j)}$ are both derived by application of the Case 2 (2) construction. Then,

$$\tau(K^{(i-1)}) = (a, X_{p'+\ell}) \cup (X_{p'+\ell}, b) ,$$

where $X_{p'+\ell}$ is a partition point of J_1. Then, by the definition of $K^{(i)}$, it is equal to $\Delta^{(o)}$ or $\Delta^{(1)}$, so that $X_{p'+\ell}$ is an end point of $K^{(i)}$. Now, τ maps partition points to partion points, and since τ is monotonic on each subinterval it is easy to see that $K^{(j-1)}$ must be an interval with one of its endpoints a partition point of J_1, say X_p. Then, by our assumption,

$$\tau(K^{(j-1)}) = (a', x_r) \cup (x_r, b') ,$$

where x_p is mapped into either a' or b'. Since τ is Markov, a' or b' is a partition point. Thus $K^{(j)} \varepsilon J_1$ and the sequence terminates. Contradiction.

Now, let s be the index such that $K^{(s)}$ is obtained from $K^{(s-1)}$ by applying the construction of Case 2 (2), (if this case is impossible we put $s = 1$) and let us consider the sequence of numbers, $\eta(K^{(s)}), \eta(K^{(s+1)}),\ldots$. We note that if this sequence did not terminate, we would have $\eta(K^{(s+i)}) < 1$ for all $i \geq 0$.

Now it is clear that for all $p \geq s$, $K^{(p+1)}$ can only be derived from the construction of Case 1 or Case 2 (1). For any such p, let k be the smallest index such that $K^{(p+k)}$ is derived by applying

Case 2 (1). Let $K^{(p+j)} \subseteq I_{i_j} \in J_1$ for all $0 \le j \le k-1$. There can be no repetitions among the sequence $\{I_{i_j}\}_{j=0}^{k-1}$ since the matrix $g(B_\tau^{(1)})$ is irreducible and hence, the elements of J_1 are aperiodic. Therefore, we must have $k \le n$. If $k = n$, then $B_\tau^{(1)}$ would be a permutation matrix, contradicting the assumption. Thus $k \neq n$. Therefore, by inequality (5.6), we have,

$$\eta(K^{(p+k-1)}) \ge \delta_{i_0} \cdots \delta_{i_{k-2}} \eta(K^{(p+1)}) .$$

Since $\eta(K^{(p+1)}) \ge \delta_{i_{k-1}} (1+c)\, \eta(K^{(p)})$ (by 2(1)) we obtain,

$$\eta(K^{(p+k-1)}) \ge \delta_{i_0} \cdots \delta_{i_{k-2}} \delta_{i_{k-1}} (1+c)\, \eta(K^{(p)}),$$

and, since, $\eta(K^{(p+k)}) \ge \delta_{i_k} \eta(K^{(p+k-1)})$ (by (5.6), we finally get)

$$\eta(K^{(p+k)}) \ge \delta_{i_0} \cdots \delta_{i_{k-1}} \delta_{i_k} (1+c)\, \eta(K^{(p)}) ,$$

or

$$\eta(K^{(p+k)}) \ge \delta_1 \cdots \delta_n (1+c)\, \eta(K^{(p)}) ,$$

since the δ_{i_j} are all distinct and $\delta_{i_j} \le 1$.

Consider now the subsequence $\{\eta(K^{(s)}), \eta\left(K^{(s+k_1)}\right), \eta\left(K^{(s+k_1+k_2)}\right), \ldots, \eta\left(K^{(s+\sum_{i=1}^{g} k_i)}\right)$, where k_g is the index defined above, corresponding to $p = s + \sum_{i=1}^{g} k_i$. Then,

$$\eta\left(K^{(s+\Sigma_{i=1}^{g} k_i)}\right) \geq \delta_1 \delta_2 \ldots \delta_n (1+c)\, \eta(K^{(s)})$$

$$> 1 ,$$

for g sufficiently large, since $\delta_1 \delta_2 \ldots \delta_n(1+c) > 1$ by hypothesis. This contradicts the statement that $\eta(K^{(s+i)}) < 1$, for all $i \geq 0$.

Q.E.D.

Remark 1: We had assumed that the transition matrix $B_\tau^{(1)}$ is not a permutation matrix. This is a necessary condition as we shall see from the following example.

Example 5.3: Let $\tau : [0,3] \to [0,3]$ be the piecewise linear map defined by $\tau(0) = 1$, $\tau(2^-) = 3$, $\tau(2^+) = 0$, $\tau(3) = 1$, and shown in Fig. 5.2.

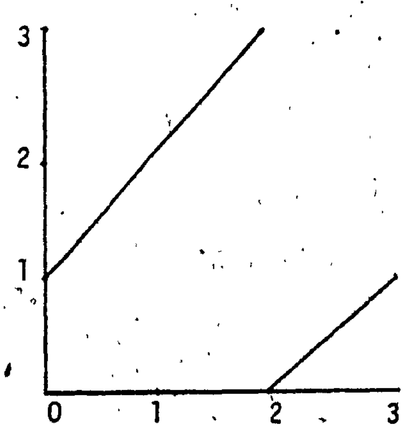


Fig. 5.2

Then $J_1 = \{(0,1), (1,2), (2,3)\}$, and

$$g(B_\tau^{(1)}) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} .$$

B_τ^1 is a permutation matrix and irreducible.

Consider now the partition

$$J_2 = \{(0.5), (.5,1), (1,1.5), (1.5,2), (2,2.5), (2.5,3)\} .$$

The 0-1 matrix induced by J_2 i.e., the $g(B_\tau^{(2)})$ is,

$$g(B_\tau^{(2)}) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

But $g(B_\tau^{(2)})$ is not an irreducible matrix as we can easily see from the associated directed graph. It consists of two disjoint cycles $1 \to 3 \to 5 \to 1$ and $2 \to 4 \to 6 \to 2$, which is not strongly connected.

Remark 2: If τ is piecewise linear, then $\delta_i = 1$, $i = 1,\ldots,n$, and in that case, the sufficiency condition in Theorem 5.3.2 reduces to $(1+c) > 1$, which is always true. Therefore, Theorem 5.3.2 is true when τ is piecewise linear.

If we combine Theorems 5.3.1 and 5.3.2, we obtain the following main result.

<u>Theorem 5.3.3</u>: Let $\mathcal{J}_1$ be a strong Markov partition of I with respect to τ, $\tau \in W$, satisfying the following condition,

$$\delta_1\delta_2\ldots\delta_n(1+c) > 1.$$

Furthermore, let $\mathcal{J}_2$ be any other partition of I. Let $g(B_\tau^{(1)})$ and $g(B_\tau^{(2)})$ be the 0-1 matrices induced by $\mathcal{J}_1$ and $\mathcal{J}_2$, respectively. Then $g(B_\tau^{(2)})$ is irreducible. If $g(B_\tau^{(1)})$ is also primitive, then so is $g(B_\tau^{(2)})$.

<u>Proof</u>: This follows from Theorems 5.3.1 and 5.3.2

Q.E.D.

Theorem 5.3.3 may not hold if $\mathcal{J}_1$ is not Markov. To prove this, consider the following example.

<u>Example 5.4</u>: Consider the piecewise linear map $\tau: [0,1]\to[0,1]$ defined by the conditions.

$$\tau(0) = 0,\ \tau(.5^-) = 1,\ \tau(.5^+) = .9 \text{ and } \tau(1) = .4$$

and, shown in Fig. 5.3,

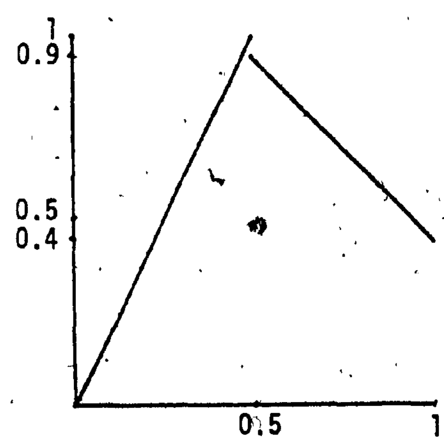


Fig. 5.3

Then $\mathcal{J}_1 = \{(0,.5), (.5,1)\}$, and,

$$g(B_\tau^{(1)}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} .$$

Clearly, $\mathcal{J}_1$ is not a Markov partition, but the induced matrix $g(B_\tau^{(1)})$ is, obtiously, primitive.

Now, consider the partition $\mathcal{J}_2 = \{(0,.4), (.4,.5), (.5,.9),(.9,1)\}$. The matrix $g(B_\tau^{(2)})$, induced by $\mathcal{J}_2$ is

$$g(B_\tau^{(2)}) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} .$$

Now, we consider the matrix $C = g(B_\tau^{(2)}) + g(B_\tau^{(2)})^2 + g(B_\tau^{(2)})^3$.

$$C = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 3 & 5 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 3 & 3 \\ 0 & 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 8 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 6 & 5 \\ 0 & 3 & 5 & 3 \end{bmatrix}$$

Since the matrix C contains zero entries off the main diagonal, according to Corollary 4.2.1 the corresponding graph to $g(B_\tau^{(2)})$ is not strongly connected, and so $B_\tau^{(2)}$ is not irreducible.

<u>Example 5.5</u>: Consider the continuous map $\tau: [0,1] \quad [0,1]$ defined by

$$\tau(x) = \begin{cases} 2x + \frac{1}{2} & , \quad 0 \le x \le \frac{1}{4} \\ -2x + \frac{3}{2} & , \quad \frac{1}{4} \le x \le \frac{3}{4} \\ 2x - \frac{3}{2} & , \quad \frac{3}{4} \le x \le 1 \end{cases}$$

and shown in Fig. 5.4.

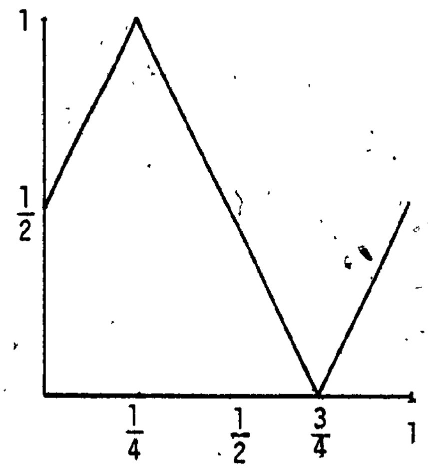


Fig. 5.4

Then $J_1 = \{(0, \frac{1}{4}), (\frac{1}{4}, \frac{3}{4}), (\frac{3}{4}, 1)\}$ and

$$g(B_\tau^{(1)}) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

It is clear that $g(B_\tau^{(1)})$ is primitive $(g(B_\tau^{(1)})^2 > 0)$. Now, let $J_2 = \{(0, \frac{1}{4}), (\frac{1}{4}, \frac{1}{2}), (\frac{1}{2}, \frac{3}{4}), (\frac{3}{4}, 1)\}$ be a second partition, which induces the matrix

$$g(B_\tau^{(2)}) = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

$g(B_\tau^{(2)})$ is not primitive (much less, irreducible). Clearly, J_1 is not a Markov partition.

Example 5.6: Let $\tau: [1\to5]\to[1,5]$ be the piecewise linear map defined by,

$$\tau(1) = 3,\ \tau(2) = 5,\ \tau(3) = 4,\ \tau(4) = 2 \text{ and } \tau(5) = 1,$$

and shown in Fig. 5.5.

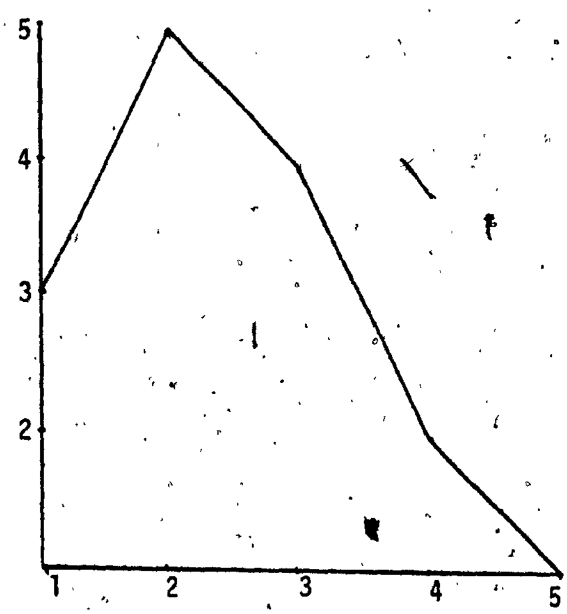


Fig. 5.5

Then, $J_1 = \{(1,2), (2,3), (3,4), (4,5)\}$. Let J_2 be another partition of $[1,5]$, given by

$$J_2 = \{(1,1.4),(1.4,2),(2,2.4),(2.4,3),(3,3.8),(3.8,4),(4,4.6),(4.6,5)\}.$$

J_1 is obviously a Markov partition, inducing the matrix

$$g(B_\tau^{(1)}) = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix $g(B_\tau^{(1)})$ is primitive, since by Theorem 4.2.2 we have

$$C = g(B_\tau^{(1)}) + g(B_\tau^{(1)})^2 + g(B_\tau^{(1)})^3 = \begin{bmatrix} 1 & 2 & 4 & 3 \\ 2 & 0 & 1 & 2 \\ 1 & 3 & 3 & 2 \\ 2 & 1 & 2 & 2 \end{bmatrix}$$

and $g(B_\tau^{(1)})_{33} = 1 > 0$. Hence the matrix

$$g(B_\tau^{(2)}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

induced by J_2, is primitive, by Remark 2 in section 5.3.

Example 5.7: Let τ and J_1 be defined as they were in Example 5.6. Let n be divisible by 8, and let J_2 be an equal length partition of [1,5] consisting of n subintervals. Then the matrix induced by J_2 is

$$g(B_\tau^{(2)}) = \begin{bmatrix} 0 & 0 & A_n & B_n \\ 0 & 0 & 0 & I'_n \\ 0 & C_n & D_n & 0 \\ I'_n & 0 & 0 & 0 \end{bmatrix}$$

where each block is an $\frac{n}{4} \times \frac{n}{4}$ matrix and the non-zero blocks are defined as follows.

Let H_n and G_n be $\frac{n}{8} \times \frac{n}{8}$ band matrices given by

$$H_n = \begin{bmatrix} 1 & 1 & & & \\ & 1 & 1 & 0 & \\ & & & \ddots & \\ & 0 & & & \\ & & & 1 & 1 \end{bmatrix}, \quad G_n = \begin{bmatrix} & & & 1 & 1 \\ & & 0 & & \\ & & & & \\ & 1 & 1 & & 0 \\ 1 & 1 & & & \end{bmatrix}$$

Define, $A_n = \left(\frac{H_n}{0}\right)$, $B_n = \left(\frac{0}{H_n}\right)$, $C_n = \left(\frac{G_n}{0}\right)$ and $D_n = \left(\frac{0}{G_n}\right)$. I'_n is the $\frac{n}{4} \times \frac{n}{4}$ skew diagonal matrix. Since $g(B_\tau^{(1)})$ is primitive, so is $g(B_\tau^{(2)})$.

Example 5.8: Consider the map $\tau: [0,1] \to [0,1]$, defined as follows:

$$\tau(x) = \begin{cases} 2x , & 0 \le x \le \frac{1}{2} \\ -2x + 2 , & \frac{1}{2} \le x \le 1 \end{cases} .$$

Let $\mathcal{J}_1 = \{(0, \frac{1}{2}), (\frac{1}{2}, 1)\}$. Obviously, $\mathcal{J}_1$ is a Markov partition which induces the primitive matrix

$$g(B_\tau^{(1)}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} .$$

Now, let us consider the partition

$$\mathcal{J}_2 = \left\{\left(\frac{1}{n+1}, \frac{1}{n}\right), \ldots, \left(\frac{1}{3}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{2}{3}\right), \left(\frac{2}{3}, \frac{3}{4}\right), \ldots, \left(1 - \frac{1}{n}, 1 - \frac{1}{n+1}\right)\right\} ,$$

where $n+1$ is even. $\mathcal{J}_2$ is not a Markov partition (because $\frac{2}{7}$ is not a partition point). The $2n \times 2n$ matrix induced by $\mathcal{J}_2$ is

$$g(B_\tau^{(2)}) = \begin{bmatrix} A_n & B_n \\ \overline{A}_n & \overline{B}_n \end{bmatrix} ,$$

where A_n, B_n, $\overline{A}_n$, $\overline{B}_n$ are $n \times n$ matrices, defined as follows:

$$A_n = \begin{bmatrix} 1 & 1 \ldots 1 & 0 & & \\ & & 1 & & \\ & & \vdots & 0 & \\ & P & 1 & & \\ & & & \ddots & 1 \\ & & 0 & & 1 \\ & & & & 0 \\ & & & & 0 \end{bmatrix}, \quad B_n = \begin{bmatrix} & & & & \\ & & & & \\ & & & 0 & \\ & & & & \\ 1 & & & & \\ 0 & 1 & 1 & & \end{bmatrix}$$

where, in A_n, the first row consists of $\frac{n+3}{2}$ 1's followed by 0's. Also, $\overline{A}_n$, $\overline{B}_n$ are obtained from A_n, B_n, respectively, by interchanging the rows, i.e., the first row becomes the last, the second row, the second of last; and so on. By Theorem 5.3.3 $g(B_\tau^{(2)})$ is primitive.

<u>Remark</u>: As simple as the transformation of this example is, it can generate by appropriate partition choices, large classes of matrices. All these matrices are primitive.

REFERENCES

[1] Bellman, R., Introduction to Matrix Analysis, McGraw-Hill (1960).

[2] Bondy, J.A. and Murty, V.S.R., Graph Theory with Application, the Macmillan Press Ltd., London, 1976.

[3] Boyarsky, A. and Scarowsky, M., On a class of transformations which have unique absolutely continuous invariant measure. Trans. Amer. Math. Soc. 255 (1979), 243-262.

[4] Dulmage, A.L. and Mendelsohn, N.S., Graphs and Matrices, in Graph Theory and Theoretical Physics (F. Harary, ed.), Academic Press, 167-229, 1967.

[5] Franklin, J.N., Matrix Theory, Prentice-Hall Inc., Englewood Cliffs, N.J., 1968.

[6] Friedman, N. and Boyarsky, A., Matrices and Eigenfunctions induced by Markov maps, Linear Algebra and its Applications, (to appear).

[7] Friedman, N., and Boyarsky, A., Irreducibility and Primitivity using Markov maps, Linear Algebra and its Applications, (to appear).

[8] Friedman, N., and Boyarsky, A., Construction of Ergodic Transformations, Advances in Mathematics, (to appear).

[9] Gantmacher, The Theory of Matrices (vol. I and II), Chelsea Publ. Comp., New York, 1960.

[10] Haddad, G., Upper Bounds for the Number of Absolutely Continuous Measures Invariant under Transformations, Master's Thesis, Concordia University, 1980.

[11] Harary, F., Graph Theory, Addison-Wesley, Massachusetts, 1969.