ANALYSIS OF HIGHER ORDER MARKOV CHAINS

By

SENGODA GOUNDER GANESAN

Bachelor of Engineering
Madras University
Madras, India
1965

Master of Science
Madras University
Madras, India
1967

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1970

ANALYSIS OF HIGHER ORDER
MARKOV CHAINS

Thesis Approved:

James E Shamblin
Thesis Adviser

M. P. Terrell

E J Ferguson

R Leroy Sellers

D. Durham
Dean of the Graduate College

ACKNOWLEDGEMENT

I want to express my personal gratitude to Dr. James E. Shamblin for the suggestion of this thesis topic and his guidance and patience throughout the execution of this research while serving as committee chairman and thesis adviser by giving his many invaluable hours. I wish to express my sincere gratitude to Professor Wilson J. Bentley for his invaluable and much appreciated support and encouragement.

I am deeply indebted to Dr. Earl J. Ferguson, Dr. Marvin Palmer Terrell, and Dr. John Leroy Folks for their continual source of helpful suggestions.

The author wishes to thank Watumull Foundations for their Grant-in-aid which was used to defray the expenses of this dissertation and Miss Bell Edge for her efficient typing.

I want to express gratefulness and appreciation to my wife Devi, for her continued encouragement, understanding, and numerous sacrifices for the completion of this undertaking.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The stochastic process $(X_t; t=0, 1, 2, 3, \ldots)$ is a Markov chain if it has the following properties:

1. A finite set of possible outcomes.
2. The probability of the next outcome is dependent upon its immediately preceding outcome(s).
3. These probabilities are constant over time.

A stochastic process having the above Markov properties is said to be a first order Markov chain if its next outcome depends only upon the present outcome, a second order Markov chain if its next outcome depends upon its present outcome, and the outcome immediately preceding the present outcome, and so on.

First order Markov chains have been studies and reviewed in detail by Kemeny and Snell (8), Karlin (7), Breiman (2), Parzen (9), Hillier and Leiberman (5), Chung (3), and others. A short discussion of first order Markov chains, their solutions and properties will follow.

Markov chains can be analyzed using the fundamentals of classical probability theory. Markov chains can be analyzed by means of their transition probability matrices. Each element in the transition matrix represents the

probability of going from one state to another. For notation purposes, an element in a transition probability matrix will be called p_{ij}. This is the conditional probability that if the process is now in state i, it will be in state j on the next step. To be a transition matrix, the following conditions must be met:

1. Each element must be a probability, i.e., $0 \leq p_{ij} \leq 1$.
2. Each row must sum to exactly one, i.e., $\sum_{j=1}^{m} p_{ij} = 1$, where i, and j are for identifying the process's present and past outcomes respectively.

For the state space $S = (S_i;\ i=1, 2, .., m)$ the transition probability matrix P of the first order Markov chain $(X_t;\ t=0, 1, 2, ...)$ would appear as follows:

$$P = \begin{array}{c|cccccc} & S_1 & S_2 & S_3 & - & - & S_m \\ S_1 & p_{11} & p_{12} & p_{13} & - & - & p_{1m} \\ S_2 & p_{21} & p_{22} & p_{23} & - & - & p_{2m} \\ S_3 & p_{31} & p_{32} & p_{33} & - & - & p_{3m} \\ \vdots & \vdots & \vdots & \vdots & - & - & \vdots \\ \vdots & \vdots & \vdots & \vdots & - & - & \vdots \\ \vdots & \vdots & \vdots & \vdots & - & - & \vdots \\ S_m & p_{m1} & p_{m2} & p_{m3} & - & - & p_{mm} \end{array}$$

In the matrix P for a given state at the k^{th} period, a row exhaustively enumerates all possible states that the process can take. Thus a row is a probability vector. This is to be expected since a vector is simply a 1 x m matrix. Thus a transition matrix P is a matrix composed of rows of probability vectors. For purposes of notation, a row vector (probability vector) will be represented as V_i to represent the i^{th} row. The distribution of the process for the k^{th} stage can be analyzed by obtaining P^k. P^k is obtained by multiplying P successively. If V_i^k is the probability vector describing the probabilities of possible outcomes in k steps for the present state S_i, then this could be obtained if V_i^2 can be obtained. V_i^2 is obtained either by the classical probability method or by the matrix method. If the present, the next, and the time after next steps are represented by 0, 1 and 2, then the possible states in every step are as follows:

0	1	2
S_k	$(S_1, S_2, .., S_m)$	S_k $k=1, 2, .., m.$

For the above possible states in the respective stages

$$p_{ik}^{\,2} = \sum_{j=1}^{m} p_{ij}\, p_{jk}$$

where $p_{ik}^{\,2}$ = the conditional probability of the process being in k, time after next given that it is in i now.

Therefore,

$$V_i^2 = (p_{i1}^2, p_{i2}^2, \ldots, p_{im}^2).$$

In the matrix method V_i^2 is obtained from the product of V_i^1 and P, i.e.,

$$V_i^2 = V_i^1 P$$

$$= (p_{i1}, p_{i2}, \ldots, p_{im}) \times \begin{bmatrix} p_{11} & p_{12} & - & - & - & p_{1m} \\ | & | & - & - & - & | \\ | & | & - & - & - & | \\ | & | & - & - & - & | \\ p_{m1} & p_{m2} & - & - & - & p_{mm} \end{bmatrix}$$

$$= \left(\sum_{j=1}^{m} p_{ij}p_{j1}, \sum_{j=1}^{m} p_{ij}p_{j2}, \ldots, \sum_{j=1}^{m} p_{ij}p_{jm}\right)$$

$$= (p_{i1}^2, p_{i2}^2, \ldots, p_{im}^2)$$

therefore,

$$V_i^k = V_i^{k-1} P$$

$$= V_i^1 P^{k-1}$$

$$= (p_{i1}^k, \ldots, p_{im}^k)$$

where p_{ij}^{k} is the probability of the process being in the state j at t=k given that it is in the state i at t=0 (now). Thus the probability vector V_i^{k} describes the probability distribution of the process for the k^{th} step from now. Actually, if the results after k steps are desired, P^k gives even more complete information since it is composed of all the individual vectors V_i^{k}. Thus P^k gives the probabilities of being in any given state for all possible starting conditions or states.

If the system or process being modeled as a Markov chain has certain properties, it is possible to determine the probabilities of outcomes after steady state conditions have been reached. After the process has been in operation for a long period of time, a given outcome will result x percent of the time. At times, it is desirable to be able to determine these percentages. Perhaps the most detrimental assumed condition in this case is the requirement that the transition matrix contain probabilities which are constant over time. This requirement should always be kept in mind when this analysis is being made to insure that the results obtained are properly interpreted.

To insure that steady state conditions may be reached, the chain must be <u>ergodic.</u> An ergodic chain mathematically describes a process in which it is possible to eventually go from any one state to any other state. It is not necessary that this be accomplished in just one step but it is required that it is possible for any outcome to be

possible regardless of the present state.

A more restricted case of an ergodic chain is a regular chain. A regular chain may be defined as a Markov chain having a transition matrix P, which for some power of P, has only positive probability elements. Note that all regular chains will be ergodic but the reverse is not necessarily true.

The existence of steady state conditions in an ergodic Markov chain can be most easily demonstrated by computing P^k for various values of k. As k becomes larger, the values p_{ij}^{k} tend to a fixed limit and each probability vector V_i^k tends to become equal for all values of i. This suggests the following statements:

1. For a sufficiently large value of k, the probability vector V_i^k becomes equal for all i and does not change significantly for larger values of k.
2. Since $V_i^{k+1} = V_i^k P$, and $V_i^{k+1} = V_i^k$, then there exists a vector V* such that

$$V^* = V^*P.$$

The vector V* contains the probabilities which exist at steady state conditions. Let v_j be the j^{th} element in the probability vector V*. Since V* is a probability vector the following condition must still exist:

$$\sum_{j=1}^{m} v_j = 1 .$$

And from statement 2,

$(v_1, v_2, .., v_m)$ $P=(v_1, v_2, .., v_m)$. If this matrix product is expanded, there will be m equations. When added to the requirement that the sum of the probabilities equal 1, there are (m+1) equations and m unknowns. These may be solved for the m unknowns by discarding any 1 of the last m equations.

A special case of Markov chains that is used to describe those processes which cease upon reaching certain given conditions is called absorbing Markov chains. Several kinds of pertinent information may be obtained from the analysis of this type of chains. It is possible to determine the following data:

1. The expected number of steps before the process is absorbed.
2. The expected number of times the process is in any given non-absorbing state.
3. The probability of absorption by any given absorbing state.

The first step in the analysis is to rearrange the transition matrix so that four sub-matrices exist as indicated below:

$$P = \begin{array}{c|c} I & O \\ \hline A & N \end{array} .$$

These smaller matrices contain probability elements but taken individually do not constitute a transition matrix. If taken individually, they contain the following information concerning probabilities. Assume there are a absorbing states, n non-absorbing states and a + n = m total states.

I - an a x a identity matrix, representing the probabilities of staying within any absorbing state.

0 - an a x n zero matrix, representing the probabilities of going from any absorbing state to any other non-absorbing state.

A - an n x a matrix, containing the probabilities of going from any non-absorbing state to any other absorbing state.

N - an n x n matrix, containing the probabilities of going from any non-absorbing state to any other non-absorbing state.

One way of finding the expected number of steps before the process is absorbed would be to find the expected number of times the process would be in each non-absorbing state and sum these.

The expected number of times the process is in a non-absorbing state j is the sum of the following terms.

Expected number of times in j = (1) (probability of being in j at start) + (1) (probability of being in j after 1 step) + (1) (probability of being in j after 2 steps)+...

$= I + N + N^2 + \ldots$

For the larger power of N, the above geometric series converges to $(I - N)^{-1}$ (Kemeny and Snell, 8). Thus, for a given starting state, the matrix $(I - N)^{-1}$ gives the expected number of times a process is in each non-absorbing state before absorption.

To find the probability of absorption by any given absorbing state, a similar logic is used in the analysis. Let j signify some given absorbing state; let i signify some specified non-absorbing state.

Probability of ending in j = (probability of going from i to j in 1 step) + (probability of going from i to j in 2 steps) + (probability of going from i to j in 3 steps) +...

$= A + NA + N^2A + \ldots$

$= (I + N + N^2 + \ldots) A$.

For higher powers of N the quantity $(I + N + N^2 + \ldots)$ converges to $(I - N)^{-1}$. Therefore,

$(I + N + N^2 + \ldots) A = (I - N)^{-1} A$.

Higher order Markov chains are those Markov chains whose future outcomes depend upon one or more immediately preceding states. For example, in the case of the first order Markov chain the next outcome depends only upon the present state, in the second order Markov chain the future outcome depends upon the present state and the state

immediately preceding the present state, in the n order chains the future outcome will depend upon the present state and the (n-1) states immediately preceding the present state.

The objective of this research is to develop appropriate methods of analysis for higher order Markov chains, thus allowing them to be applied to O.R. problems. Second order chains are studied in detail. In the second chapter, a model is developed after discussing the notations used. n-dimensional matrix algebra is considered as a substitute to the classical probability theory. In the third chapter the n-dimensional matrix algebra developed in the second chapter is used to analyze the distribution of second order chains at various stages (steps or periods). Steady state probabilities, expected number of times the process is in a non-absorbing state and the probability of the process being absorbed are discussed in detail with examples.

CHAPTER II

CONCEPT OF n-ORDER MARKOV CHAINS

The stochastic process $(X_t; t=0, 1, 2...)$ defined in the state space $S = (S_1, S_2, ..., S_m)$ is an n-order Markov chain if it has the transition probability matrix P having the following properties:

1. The element of P which is the probability of the next outcome is dependent upon the present state and (n-1) states immediately preceding the present state.
2. The elements of P are constant over time.

As in the case of first order Markov chains, the elements of the transition probability matrix of higher order Markov chains are also called one step transition probabilities since they describe the conditional probability of being in a particular state in the n^{th} step, given the states at t=0, 1, 2, .., n-1.

Notations

It is worthwhile to describe the notations before the detailed study of the Markov chains is undertaken.

P - transition probability matrix, or the matrix of one-step probabilities.

$p_{a,b,..,i,j}$ - element of the matrix of one-step transition probabilities. It is the conditional probability of the process being in the state $\underline{j}$ at the n^{th} step given that it was in the states $\underline{a},\underline{b},...,\underline{i}$ at t=0, 1, 2, .., n-1 respectively.

$p^{k}_{a,b,...,i,j}$ - probability of the process being in the state j at the $(n-1+k)^{th}$ step given that it was in the states $\underline{a}$, $\underline{b}$,...,$\underline{i}$ at t=0, 1, 2,..., n-1 respectively.

I - Identity matrix.

Model

It may be convenient to think of a Markov chain as a modeling and analysis technique suitable for a special case of probability problems. These probability problems may be analyzed theoretically using the fundamentals of classical probability theory. If a process is to be analyzed by an n-order Markov chain model, the process must have one step probabilities. These one-step probabilities constituting the transition probability matrix characterize the process during its transition from a given condition to any other state in one step. The elements of a transition matrix must satisfy the Markovian properties mentioned in the beginning of this chapter. Since the transition matrices of third or higher order Markov chains have more than three dimensions, it is advisable to consider the

transition matrix of a second order Markov chain for discussion.

Let $(X_t; t=0, 1, 2, \ldots)$ be a second order Markov chain with the state space S having only two discrete points (a, b), then its three dimensional transition probability matrix P in two dimensions will be as given below:

$$P = \begin{array}{cccc} & t=0 & t=1 & t=2 \\ & & & a \quad b \\ & & a & p_{aaa} \quad p_{aab} \\ & a & & \\ & & b & p_{aba} \quad p_{abb} \\ & & a & p_{baa} \quad p_{bab} \\ & b & & \\ & & b & p_{bba} \quad p_{bbb} \end{array} .$$

Each row in P is a probability vector describing the process exhaustively for the given present state and the state immediately preceding the present state. Since each row is a probability vector the sum of the probabilities in each row must be equal to one. The dependency of the process upon the state immediately preceding the present state (state at t=0) is shown in Figure 1, in three dimensions. In Figure 1, there are two 2 x 2 matrices, one for every possible state at t=0. If the process does not depend upon the outcome at time t=0, then these two 2 x 2

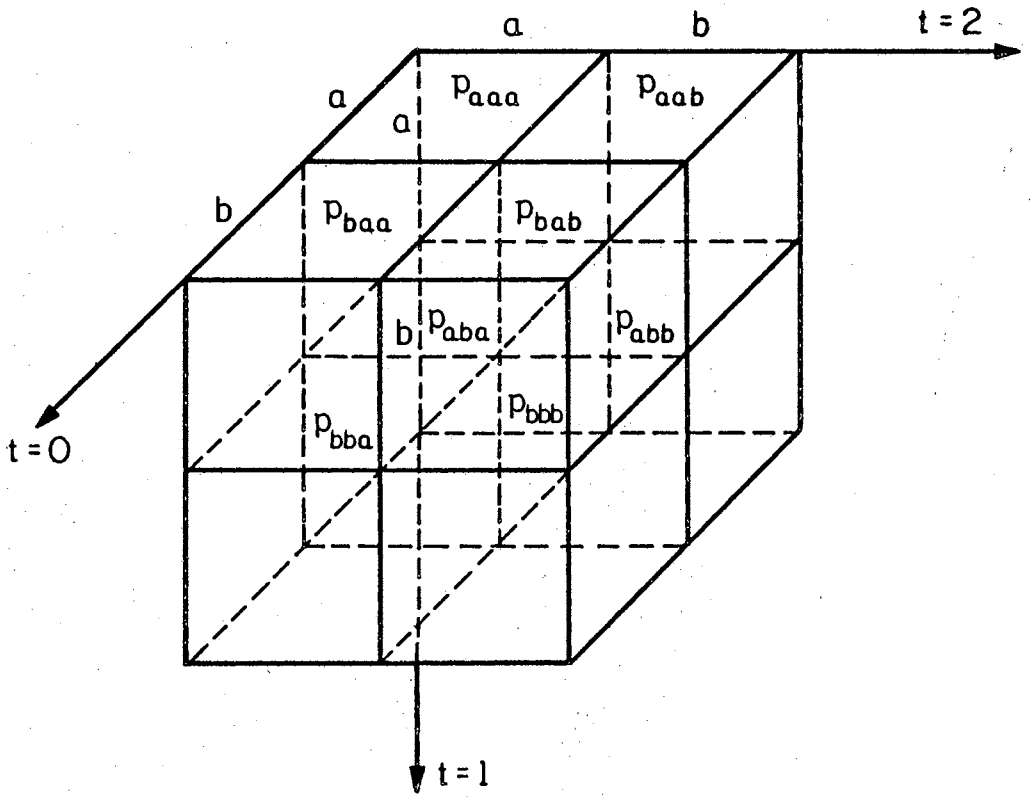

Figure 1. The Transition Probability Matrix P of A Second Order Markov Chain with the State Space S=(a,b)

matrices reduce to a single 2 x 2 matrix independent of the outcome at t=0 and it is a first order Markov chain.

The one-step probabilities can be used to analyze the distribution of the process at different steps using the classical probability theory. An n-dimensional matrix method is developed as a substitute to the classical probability theory. Since the discussion about the steady state and the absorption characteristics of the process requires the knowledge of n-dimensional matrix algebra, it is deferred to the next chapter.

Solution of Higher Order Markov Chains and n-Dimensional Matrix Algebra

Let $(X_t ; t=0, 1, 2, \ldots)$ be a second order Markov chain with the state space S= (a, b) and the transition probability matrix P having the following numerical values:

$$P= \begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{abb}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} = \begin{bmatrix} \underline{0.8} & \underline{0.2} \\ \underline{0.6} & \underline{0.4} \\ \underline{0.5} & \underline{0.5} \\ \underline{0.3} & \underline{0.7} \end{bmatrix}$$

P consists of four probability vectors, V_1 =(0.8, 0.2), V_2 =(0.6, 0.4), V_3 =(0.5, 0.5), and V_4 =(0.3, 0.7). The second stage distribution of the process can be analyzed either by the classical method or by the matrix method. The classical and matrix methods are both discussed to

show that they give the same result. However, the matrix method is computationally superior to the classical method.

Classical Method

Suppose p_{aba}^{2} is the probability of the process being in the state a time after next (i.e., at time t=3) given that it was in the state a in the step immediately preceding the present stage (i.e., at t=0) and is in b now (i.e., at t=1). This probability can be calculated by analyzing the possible states that the process may take during the intermediate stages between the present and the final stages by letting the first (the stage immediately preceding the present stage), the second (the present), and the third (the stage immediately following the present stage) stages to be 0, 1, and 2

stage t=0 1 2 3

state m=a b a or b a .

For the above possible states in the corresponding periods,

$$p_{aba}^{2} = p_{aba}\, p_{baa} + p_{abb}\, p_{bba} .$$

The rest of the second-stage probabilities can be calculated in the same way.

For p_{aba}^{3} the possible states in the corresponding periods will be:

stage t=0 1 2 3 4

stage m=a b a or b a or b a .

For the above possible states

$$p_{aba}^{3} = p_{aba}\, p_{baa}\, p_{aaa} + p_{aba}\, p_{bab}\, p_{aba}$$

$$+ p_{abb}\, p_{bba}\, p_{baa} + p_{abb}\, p_{bbb}\, p_{bba}. \qquad (1)$$

The rest of the third-stage probabilities can be calculated in the same way.

Matrix Method

The distribution of higher order Markov chains at various stages will be analyzed by the matrix method after the discussion of the process of higher order matrix multiplication.

The process of multiplication with matrices of 3 or more dimensions proceeds in a similar manner to the multiplication of two-dimensional matrices. An element in the product matrix is the result of the product of two uniquely defined vectors.

Even in a matrix of more than two dimensions, a vector is still identified by a fixed position in every dimension but one and it contains all elements in that dimension which are common to the fixed positions of the other dimension. For example, in a 2-dimensional matrix, a row vector is that vector in the i_1th row which contains all i_2= 1, 2, ... elements in the i_2th row. Note that an element is uniquely defined by an $i_1 i_2$ subscript.

In a three dimensional matrix, an element is uniquely identified by three subscripts, perhaps labeled $i_1 i_2 i_3$ with each letter denoting a position in that given dimension. A vector may be defined by fixing two of these dimensional variables and while containing all elements in the third dimension having the fixed subscripts in common. Referring to Figure 2, element (3,2,2) is the third element in the i_1 dimension, the second element in the i_2 dimension and the second element in the i_3 dimension. A vector might be labeled as (i_1=4, i_2=3, i_3= 1,2, ...) which says that it contains all elements having i_1=4, i_2=3 in common.

To obtain an element in the product matrix, the dot product of two vectors must be obtained. For example, in a three-dimensional matrix, element (i_1=2, i_2=3, i_3=5) is the dot product of the vectors (i_1=2, i_2=3, i_3=1, 2, 3 ...) and (i_1=3, i_2=1, 2, 3 ..., i_3=5). In terms of notation, the product element (2,3,5) = $\sum_{k=1}^{m} (a_{23k})(b_{3k5})$ for a cubic matrix of m elements on a side. In 5 dimensions, the product element (2,3,5,4,6) = $\sum_{k=1}^{m} (a_{2354k})(b_{354k6})$. In the case of n-dimensional matrices of m elements on a side, the product element (1,2,4,.., m-1,5) is the dot product of the row vector (i_1=1, i_2=2, i_3=4, ..., i_{n-1}=m-1, i_n= 1,2,..., m) and the column vector (i_1=2, i_2=4,..., i_{n-2}= m-1, i_{n-1}=1,2,..., m, i_n=5).

As an example of matrix multiplication, consider the product of two four-dimensional matrices A and B

AB=C

Let each matrix contain two elements along each dimension. Since it is not possible to display a matrix in the full 4 dimensions, a two-dimensional array identified by subscripts will be used.

$$A= \begin{bmatrix} a_{1111}=2 & a_{1112}=3 \\ a_{1121}=1 & a_{1122}=4 \\ a_{1211}=3 & a_{1212}=2 \\ a_{1221}=5 & a_{1222}=4 \\ a_{2111}=2 & a_{2112}=3 \\ a_{2121}=1 & a_{2122}=1 \\ a_{2211}=7 & a_{2212}=5 \\ a_{2221}=6 & a_{2222}=6 \end{bmatrix} .$$

$$B= \begin{bmatrix} b_{1111}=6 & b_{1112}=1 \\ b_{1121}=6 & b_{1122}=2 \\ b_{1211}=2 & b_{1212}=4 \\ b_{1221}=4 & b_{1222}=1 \\ b_{2111}=6 & b_{2112}=6 \\ b_{2121}=6 & b_{2122}=3 \\ b_{2211}=1 & b_{2212}=4 \\ b_{2221}=3 & b_{2222}=6 \end{bmatrix} .$$

$$AB=C=\begin{bmatrix} c_{1111}=\sum_{k=1}^{2} a_{111k}\,b_{11k1} & c_{1112}=\sum_{k=1}^{2} a_{111k}\,b_{11k2} \\ =(2)(6)+(3)(6)=30 & =(2)(1)+(3)(2)=8 \\ c_{1121}=\sum_{k=1}^{2} a_{112k}\,b_{12k1} & c_{1122}=\sum_{k=1}^{2} a_{112k}\,b_{12k2} \\ =(1)(2)+(4)(4)=18 & =(1)(4)+(4)(1)=8 \\ c_{1211}=\sum_{k=1}^{2} a_{121k}\,b_{21k1} & c_{1212}=\sum_{k=1}^{2} a_{121k}\,b_{21k2} \\ =(3)(6)+(2)(6)=30 & =(3)(6)+(2)(3)=24 \\ c_{1221}=\sum_{k=1}^{2} a_{122k}\,b_{22k1} & c_{1222}=\sum_{k=1}^{2} a_{122k}\,b_{22k2} \\ =(5)(1)+(4)(3)=17 & =(5)(4)+(4)(6)=44 \\ c_{2111}=\sum_{k=1}^{2} a_{211k}\,b_{11k1} & c_{2112}=\sum_{k=1}^{2} a_{211k}\,b_{11k2} \\ =(2)(6)+(3)(6)=30 & =(2)(1)+(3)(2)=8 \\ c_{2121}=\sum_{k=1}^{2} a_{212k}\,b_{12k1} & c_{2122}=\sum_{k=1}^{2} a_{212k}\,b_{12k2} \\ =(1)(2)+(1)(4)=6 & =(1)(4)+(1)(1)=5 \\ c_{2211}=\sum_{k=1}^{2} a_{221k}\,b_{21k1} & c_{2212}=\sum_{k=1}^{2} a_{221k}\,b_{21k2} \\ =(7)(6)+(5)(6)=72 & =(7)(6)+(5)(3)=57 \\ c_{2221}=\sum_{k=1}^{2} a_{222k}\,b_{22k1} & c_{2222}=\sum_{k=1}^{2} a_{222k}\,b_{22k2} \\ =(6)(1)+(6)(3)=24 & =(6)(4)+(6)(6)=60 \end{bmatrix}.$$

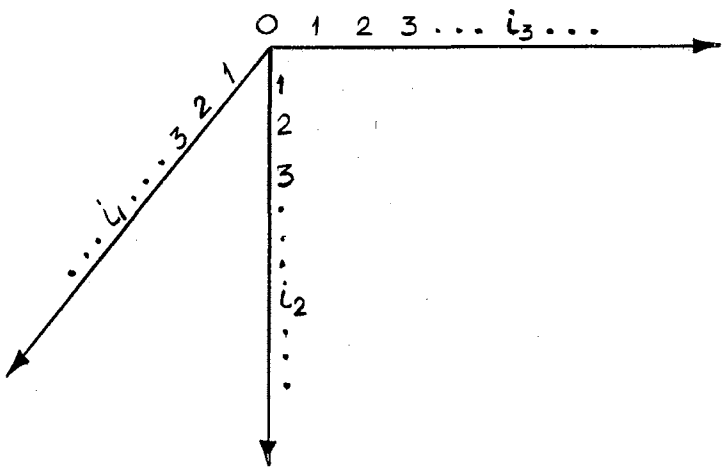

Figure 2. A Representation of a Three-Dimensional Matrix Subscripted by (i_1, i_2, i_3)

Determination of the Distribution of Higher Order Markov Chains at Various Stages by the Process of Matrix Multiplication

Let the first, the second, and the third stage transition probability matrices be represented by P, P^2, and P^3 respectively. The transition matrix P^2 is obtained by multiplying P by itself.

To find the value of $p_{aba}{}^2$ by the matrix method the following procedure is followed. Take the probability vector such that each of its elements has the first n of the given states as its first n subscripts in the proper order. The next step will be to find the column in P^{k-1} such that its first element's first n subscripts are the same as the last n subscripts of the first element of the row vector, these n subscripts in both the elements must be in the same order. The same relationship must exist for the rest of the elements of the row and column vectors. The n+1 st subscript of all the elements of the column vector is the state where the process is going to be in the n-1+k th step, where k=1, 2, 3 .... After selecting the row and column vectors they are dotted with each other to get the conditional probability. For $p_{aba}{}^2$, the row

and column vectors are (p_{aba}, p_{abb}) and $\begin{pmatrix} p_{baa} \\ p_{bba} \end{pmatrix}$ respectively. Therefore,

$$p_{aba}^{2} = (p_{aba}, p_{abb}) \cdot \begin{pmatrix} p_{baa} \\ p_{bba} \end{pmatrix}$$

$$= p_{aba}\, p_{baa} + p_{abb}\, p_{bba}$$

$$p_{aaa}^{2} = (p_{aaa}, p_{aab}) \cdot \begin{pmatrix} p_{aaa} \\ p_{aba} \end{pmatrix}$$

$$= (p_{aaa}\, p_{aaa} + p_{aab}\, p_{aba}) .$$

The same way as above the values of p_{bba}^{2}, p_{bbb}^{2}, p_{aab}^{2}, p_{baa}^{2}, and p_{bab}^{2} are calculated. These values constitute the second stage transition probability matrix P^2, i.e.

$$P^2 = \begin{bmatrix} \underline{p_{aaa}^{2}} & \underline{p_{aab}^{2}} \\ \underline{p_{aba}^{2}} & \underline{p_{abb}^{2}} \\ \underline{p_{baa}^{2}} & \underline{p_{bab}^{2}} \\ \underline{p_{bba}^{2}} & \underline{p_{bbb}^{2}} \end{bmatrix} .$$

For the numerical example, P^2 will be:

$$P^2 = \begin{bmatrix} \underline{0.76} & \underline{0.24} \\ \underline{0.42} & \underline{0.58} \\ \underline{0.70} & \underline{0.30} \\ \underline{0.36} & \underline{0.64} \end{bmatrix} .$$

The elements of P^2 completely describe the process at the second stage. Each vector gives the probability of the process for all possible states in the second step for the present and past states.

To obtain the third stage probability transition matrix P^3, P is multiplied by P^2, i.e., $P^3 = P \times P^2$.

$$P^3 = \begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{abb}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} \times \begin{bmatrix} \underline{p_{aaa}}^2 & \underline{p_{aab}}^2 \\ \underline{p_{aba}}^2 & \underline{p_{abb}}^2 \\ \underline{p_{baa}}^2 & \underline{p_{bab}}^2 \\ \underline{p_{bba}}^2 & \underline{p_{bbb}}^2 \end{bmatrix}$$

$$p_{aba}^{3} = (p_{aba}, p_{abb}) \cdot \begin{pmatrix} p_{baa}^2 \\ p_{bba}^2 \end{pmatrix}$$

$$= (p_{aba}\, p_{baa}^2 + p_{abb}\, p_{bba}^2).$$

$$p_{abb}^{3} = (p_{aba}, p_{abb}) \cdot \begin{pmatrix} p_{bab}^2 \\ p_{bbb}^2 \end{pmatrix}$$

$$= (p_{aba}\, p_{bab}^2 + p_{abb}\, p_{bbb}^2).$$

$$p_{aaa}^{3} = (p_{aaa}, p_{aab}) \cdot \begin{pmatrix} p_{aaa}^2 \\ p_{aba}^2 \end{pmatrix}$$

$$= (p_{aaa}\, p_{aaa}^2 + p_{aab}\, p_{aba}^2).$$

$$p_{aab}^{3} = (p_{aaa},\ p_{aab}) \cdot \begin{pmatrix} p_{aab}^{2} \\ p_{abb}^{2} \end{pmatrix}$$

$$= (p_{aaa}\ p_{aab}^{2} + p_{aab}\ p_{abb}^{2}).$$

$$p_{bba}^{3} = (p_{bba},\ p_{bbb}) \cdot \begin{pmatrix} p_{baa}^{2} \\ p_{bba}^{2} \end{pmatrix}$$

$$= (p_{bba}\ p_{baa}^{2} + p_{bbb}\ p_{bba}^{2}).$$

$$p_{bbb}^{3} = (p_{bba},\ p_{bbb}) \cdot \begin{pmatrix} p_{bab}^{2} \\ p_{bbb}^{2} \end{pmatrix}$$

$$= (p_{bba}\ p_{bab}^{2} + p_{bbb}\ p_{bbb}^{2}).$$

$$p_{baa}^{3} = (p_{baa},\ p_{bab}) \cdot \begin{pmatrix} p_{aaa}^{2} \\ p_{aba}^{2} \end{pmatrix}$$

$$= (p_{baa}\ p_{aaa}^{2} + p_{bab}\ p_{aba}^{2}).$$

$$p_{bab}^{3} = (p_{baa},\ p_{bab}) \cdot \begin{pmatrix} p_{aab}^{2} \\ p_{abb}^{2} \end{pmatrix}$$

$$= (p_{baa}\ p_{aab}^{2} + p_{bab}\ p_{abb}^{2}).$$

If P^3 is rewritten in terms of its elements it would appear as given below:

$$P^3 = \begin{bmatrix} \underline{p_{aaa}}^3 & \underline{p_{aab}}^3 \\ \underline{p_{aba}}^3 & \underline{p_{abb}}^3 \\ \underline{p_{baa}}^3 & \underline{p_{bab}}^3 \\ \underline{p_{bba}}^3 & \underline{p_{bbb}}^3 \end{bmatrix} .$$

The elements of P^3 obtained by the matrix method are the same as those of P^3 obtained by classical method. This can be verified for the value of p_{aba}^{3}.

$$p_{aba}^{3} = p_{aba}\, p_{baa}^{2} + p_{abb}\, p_{bba}^{2}. \qquad (2)$$

From P^2,

$$p_{baa}^{2} = p_{baa}\, p_{aaa} + p_{bab}\, p_{aba} \qquad (3)$$

and

$$p_{bba}^{2} = p_{bba}\, p_{baa} + p_{bbb}\, p_{bba}. \qquad (4)$$

Substituting the values of p_{baa}^{2}, and p_{bba}^{2} in (2),

$$\begin{aligned} p_{aba}^{3} &= p_{aba}\,(p_{baa}\, p_{aaa} + p_{bab}\, p_{aba}) + \\ &\quad p_{abb}\,(p_{bba}\, p_{baa} + p_{bbb}\, p_{bba}) \\ &= p_{aba}\, p_{baa}\, p_{aaa} + p_{aba}\, p_{bab}\, p_{aba} + \\ &\quad p_{abb}\, p_{bba}\, p_{baa} + p_{abb}\, p_{bbb}\, p_{bba} . \end{aligned} \qquad (5)$$

The same value for p_{aba}^{3} was obtained in the classical method (Equation 1).

For the numerical example, P^3 would be

$$P^3 = \begin{bmatrix} \underline{0.692} & \underline{0.308} \\ \underline{0.564} & \underline{0.436} \\ \underline{0.590} & \underline{0.410} \\ \underline{0.462} & \underline{0.538} \end{bmatrix} .$$

In multiplying higher order matrices, care must be taken to see that $P^k = P \times P^{k-1} \neq P^{k-1} \times P$. This can be proved by considering the above-mentioned second order Markov chain having the transition probability matrix P and the state space S= (a, b,). If the method of $P^2 \times P$ is used to determine the third stage transition probabilities, the element $p_{aba}{}^3$ would be as given below:

$$p_{aba}{}^3 = p_{aba}{}^2 \, p_{baa} + p_{abb}{}^2 \, p_{bba} .$$

Representing $p_{aba}{}^2$ and $p_{abb}{}^2$ in terms of one step probabilities

$$p_{aba}{}^3 = (p_{aba} \, p_{baa} + p_{abb} \, p_{bba}) \, p_{baa} + (p_{aba} \, p_{bab} + p_{abb} \, p_{bbb}) \, p_{bba}$$

$$= p_{aba} \, p_{baa} \, p_{baa} + p_{abb} \, p_{bba} \, p_{baa} + p_{aba} \, p_{bab} \, p_{bba} + p_{abb} \, p_{bbb} \, p_{bba} \qquad (6).$$

Equation (6) is not equal to (5). The first and the third factor in (6) do not follow the logic. The first factor must give the conditional probability of process being in a at (t=4) given that it was in a at t=0, is in b, a, and a at t=1, t=2, and t=3 respectively. The subscripts of the one step probabilities of the first factor are not

according to the required states at various steps. The third factor, in the same way, does not follow the required distribution of the states at various steps.
Therefore, $P^k = P \times P^{k-1} \neq P^{k-1} \times P$.

From the matrix method it is observed that P^k can be determined from P^{k-1} and P. In the classical method all the states in between the stages are analyzed exhaustively to determine P^k. For higher values of k, the classical method, therefore, does not have computational superiority over the matrix method.

Generalization of Matrix Method for n-Order Markov chains

Let $(X_t;\ t=0, 1, 2, \ldots)$ be an n-order Markov chain with state space $S= (a, b, c, \ldots, m)$ and the transition probability matrix P. The k step transition probability matrix P^k is the product of P and P^{k-1}. In determining the element of P^k the procedure mentioned in the description of the matrix method is followed. For example,

$$p^k_{a,b,..,i,j} = (p_{a,b,..,i,a}, p_{a,b,..,i,b}, .., p_{a,b,..i,m}) \cdot \begin{bmatrix} p_b^{k-1}{}_{..,i,a,j} \\ p_b^{k-1}{}_{..,i,b,j} \\ \vdots \\ p_b^{k-1}{}_{..,i,m,j} \end{bmatrix}.$$

It is noticed from the equation that the first n subscripts of all the elements of the row vector of P are the same as the first n subscripts of $p^{k}_{a,b,..,i,j}$ and these n subscripts have the same order. The last n subscripts of the elements of the row vector are the same as the first n subscripts of the corresponding elements of the column vector. The $(n+1)^{st}$ subscript of all the elements of column vector will be the state where the process is required to be at the $(n-1+k)^{th}$ step.

The result of the above equation is

$$p^{k}_{a,b,\ldots,i,j} = (p_{a,b,..,i,a})\,(p^{k-1}_{b,\ldots,i,a,j}) + (p_{a,b,..,i,b})\,(p^{k-1}_{b,\ldots,i,b,j}) + \ldots + (p_{a,b,..i,m})\,(p^{k-1}_{b,\ldots,i,m,j}).$$

CHAPTER III

STEADY STATE AND ABSORPTION PROPERTIES OF HIGHER ORDER MARKOV CHAINS

Determination of Steady State Conditions of Higher Order Ergodic Markov Chains

The existence of steady state conditions in a higher order ergodic Markov chain can be demonstrated by computing P^k for various values of k. For the example given in Chapter II the probability matrix at the steady state would be as given below:

$$P^k = \begin{bmatrix} .6 & .4 \\ .6 & .4 \\ .6 & .4 \\ .6 & .4 \end{bmatrix} .$$

Because of the tediousness involved in the above work, a computer program is given on page 31.

To determine the analytical method for finding the steady state probabilities recall what was said in Chapter I regarding the steady state probabilities of first order Markov chains. As in the case of the first order Markov chains, steady state probabilities for higher order

A FORTRAN IV Program for Determining the Steady State Probabilities of a Second Order Markov Chain with the State Space S=(a,b):

```
      DIMENSION A(8),P(8),R(8),S(8)
      DATAS(1),S(2),S(3),S(4),S(5),S(6),S(7),S(8)/0.0,0.0,0.0,0.0,0.0,0.0,
     10.0,0.0,0.0,0.0/
      READ(5,1)A
    1 FORMAT(8F4.2)
      DO2I=1,27
      IF(I.GT.1)GOTO3
      DO5J=1,8
    5 P(J)=A(J)
      GOTO4
    3 R(1)=A(1)*P(7)+A(2)*P(5)
      R(2)=A(1)*P(8)+A(2)*P(6)
      R(3)=A(3)*P(3)+A(4)*P(1)
      R(4)=A(3)*P(4)+A(4)*P(2)
      R(5)=A(5)*P(7)+A(6)*P(5)
      R(6)=A(5)*P(8)+A(6)*P(6)
      R(7)=A(7)*P(3)+A(8)*P(1)
      R(8)=A(7)*P(4)+A(8)*P(2)
      DO6K=1,8
    6 P(K)=R(K)
    4 WRITE(6,7)P
      DO8L=1,8
    8 S(L)=S(L)+P(L)
    2 CONTINUE
    7 FORMAT(5X,8(F8.6,2X))
      WRITE(6,7)S
      STOP
      END
```

```
$ENTRY
0.600000  0.400000  0.800000  0.200000  0.300000  0.700000  0.500000  0.500000
0.420000  0.580000  0.760000  0.240000  0.360000  0.640000  0.700000  0.300000
0.564000  0.436000  0.692000  0.308000  0.462000  0.538000  0.590000  0.410000
0.538800  0.461200  0.666400  0.333600  0.500400  0.499600  0.628000  0.372000
0.576960  0.423040  0.640880  0.359120  0.538680  0.461320  0.602600  0.397400
0.577032  0.422968  0.628096  0.371904  0.557856  0.442144  0.608920  0.391080
0.588494  0.411505  0.617883  0.382116  0.573175  0.426824  0.602564  0.397436
0.590808  0.409191  0.612005  0.387994  0.581991  0.418008  0.603188  0.396811
0.594710  0.405290  0.607766  0.392233  0.588350  0.411649  0.601407  0.398593
0.596184  0.403815  0.605154  0.394845  0.592267  0.407732  0.601238  0.398762
0.597649  0.402350  0.603360  0.396639  0.594958  0.405041  0.600669  0.399330
0.598385  0.401614  0.602218  0.397781  0.596671  0.403327  0.600505  0.399494
0.598971  0.401027  0.601451  0.398547  0.597821  0.402177  0.600301  0.399697
0.599309  0.400689  0.600955  0.399043  0.598565  0.401433  0.600211  0.399787
0.599553  0.400446  0.600626  0.399372  0.599059  0.400939  0.600132  0.399866
0.599703  0.400295  0.600411  0.399587  0.599381  0.400617  0.600089  0.399909
0.599806  0.400192  0.600270  0.399729  0.599593  0.400405  0.600057  0.399941
0.599872  0.400127  0.600177  0.399821  0.599732  0.400266  0.600038  0.399960
0.599916  0.400082  0.600116  0.399882  0.599824  0.400174  0.600024  0.399974
0.599944  0.400054  0.600076  0.399922  0.599884  0.400114  0.600016  0.399982
0.599963  0.400035  0.600049  0.399948  0.599923  0.400074  0.600010  0.399988
0.599975  0.400022  0.600032  0.399966  0.599949  0.400048  0.600006  0.399992
0.599983  0.400014  0.600021  0.399977  0.599966  0.400031  0.600003  0.399994
0.599989  0.400009  0.600013  0.399984  0.599977  0.400020  0.600002  0.399996
0.599992  0.400005  0.600008  0.399989  0.599985  0.400013  0.600001  0.399997
0.599994  0.400003  0.600005  0.399992  0.599989  0.400008  0.600000  0.399997
0.599996  0.400001  0.600003  0.399994  0.599993  0.400005  0.599999  0.399998
********  ********  ********  ********  ********  ********  ********  ********
```

Markov chains do not depend upon the present and past states of the process. Because of the complexity involved in the multiplication of higher order matrices, the analytical calculation of steady state probabilities for higher order Markov chains is much more difficult than in the case of first order Markov chains. The general procedure is to reduce the probability transition matrix P of the higher order Markov chain to an equivalent first order matrix. Once a first order matrix is determined, determination of steady state probabilities becomes easy. The concept of reducing an n-order matrix can be demonstrated with a second order matrix. Let P be the transition probability matrix of a second order Markov chain, then P would appear as given below:

$$P = \begin{bmatrix} \underline{P_{aaa}} & \underline{P_{aab}} \\ \underline{P_{aba}} & \underline{P_{abb}} \\ \underline{P_{baa}} & \underline{P_{bab}} \\ \underline{P_{bba}} & \underline{P_{bbb}} \end{bmatrix} .$$

P contains two 2 x 2 matrices, one for the state a immediately preceding the present state and the other for the state b. If the process is an ergodic one, these two matrices must be the same at the steady state independent of the states immediately preceding the present state. Since this steady state still depends upon the present state, it is called an intermediate steady state. The

intermediate steady state probabilities can be determined by treating the above matrices as <u>first order ones</u>.

$$p^{k}_{aba} = p^{k-1}_{aba}\, p_{baa} + p^{k-1}_{abb}\, p_{bba}$$

$$= p^{k-1}_{aba}\, p_{baa} + (1-p^{k-1}_{aba})\,(1-p_{bbb})$$

$$= p^{k-1}_{aba}\, p_{baa} + (1-p_{bbb}) - p^{k-1}_{aba} + p^{k-1}_{aba}\, p_{bbb}$$

$$= p^{k-1}_{aba}\,(p_{baa} + p_{bbb} - 1) + (1 - p_{bbb}) \quad .$$

If $X = p_{baa} + p_{bbb} - 1$

and $Y = (1 - p_{bbb})$

then $p^{k}_{aba} = p^{k-1}_{aba}\,(X) + Y$

$$= (p^{k-2}_{aba}\,(X) + Y)\, X + Y$$

$$= p^{k-2}_{aba}\,(X^2) + Y(1 + X)$$

$$= (p^{k-3}_{aba}\,(X) + Y)\, X^2 + Y\,(1 + X)$$

$$= p^{k-3}_{aba}\,(X^3) + Y\,(1 + X + X^2)$$

$$= p_{aba}\, X^{k-1} + Y\,(1 + X + X^2 + \ldots + X^{k-2})$$

$$\underset{k \Rightarrow \infty}{\text{Limit}}\; p^{k}_{aba} = \underset{k \Rightarrow \infty}{\text{Limit}}\; (p_{aba}\, X^{k-1} + Y\,(1 + X + X^2 + \ldots + X^{k-2}))$$

$$p^k_{aba} = Y(1-X)^{-1}$$

$$p^k_{aba} = \frac{1-p_{bbb}}{2-p_{baa}-p_{bbb}} \qquad (8).$$

$$p^k_{abb} = p^{k-1}_{aba}\, p_{bab} + p^{k-1}_{abb}\, p_{bbb}$$

$$= (1-p^{k-1}_{abb})\,(1-p_{baa}) + p^{k-1}_{abb}\, p_{bbb}$$

$$= 1-p_{baa} - p^{k-1}_{abb}\,(1-p_{baa}) + p^{k-1}_{abb}\, p_{bbb}$$

$$= p^{k-1}_{abb}\,(p_{bbb} + p_{baa} - 1) + (1 - p_{baa}) \quad .$$

If $X = p_{bbb} + p_{baa} - 1$

and $Y = 1 - p_{baa}$

then $p^k_{abb} = p^{k-1}_{abb}\, X + Y$

$$= (p^{k-2}_{abb}\, X + Y)\, X + Y$$

$$= p^{k-2}_{abb}\, X^2 + Y\,(1 + X)$$

$$= p_{abb}\, X^{k-1} + Y\,(1 + X + X^2 + \ldots + X^{k-2}) \quad .$$

$$\underset{k \Rightarrow \infty}{\text{Limit}}\ p^k_{abb} = \underset{k \Rightarrow \infty}{\text{Limit}}\ (p_{abb}\, X^{k-1} + Y\,(1 + X + X^2 + \ldots + X^{k-2}))$$

$$p^k_{abb} = Y\,(1-X)^{-1}$$

$$= \frac{1-p_{baa}}{2-p_{bbb}-p_{baa}}$$

$$p^{k}_{abb} = \frac{1-p_{baa}}{2-p_{baa}-p_{bbb}} \qquad (9).$$

$$p^{k}_{aaa} = p^{k-1}_{aaa}\, p_{aaa} + p^{k-1}_{aab}\, p_{aba}$$

$$= p^{k-1}_{aaa}\, p_{aaa} + (1-p^{k-1}_{aaa})\,(1-p_{abb})$$

$$= p^{k-1}_{aaa}\, p_{aaa} + (1-p_{abb}) - p^{k-1}_{aaa}\,(1-p_{abb})$$

$$= p^{k-1}_{aaa}\,(p_{aaa} + p_{abb} - 1) + (1-p_{abb})\,.$$

If $X = p_{aaa} + p_{abb} - 1$

and $Y = (1-p_{abb})$

then $p^{k}_{aaa} = p^{k-1}_{aaa}\, X + Y$

$$= (p^{k-2}_{aaa}\, X + Y)\, X + Y$$

$$= p^{k-2}_{aaa}\, X^2 + Y\,(1 + X)$$

$$= p_{aaa}\, X^{k-1} + Y\,(1 + X + \ldots + X^{k-2})$$

$$\underset{k \Rightarrow \infty}{\text{Limit}}\; p^{k}_{aaa} = \underset{k \Rightarrow \infty}{\text{Limit}}\,(p_{aaa}\, X^{k-1} + Y\,(1 + X + \ldots + X^{k-2}))$$

$$= Y\,(1-X)^{-1}$$

$$p^k_{aaa} = \frac{1-p_{abb}}{2 - p_{aaa} - p_{abb}} \qquad (10).$$

$$p^k_{aab} = p^{k-1}_{aaa}\, p_{aab} + p^{k-1}_{aab}\, p_{abb}$$

$$= (1-p^{k-1}_{aab})\,(1-p_{aaa}) + p^{k-1}_{aab}\, p_{abb}$$

$$= (1-p_{aaa}) - p^{k-1}_{aab}\,(1-p_{aaa}) + p^{k-1}_{aab}\, p_{abb}$$

$$= p^{k-1}_{aab}\,(p_{aaa} + p_{abb} - 1) + (1-p_{aaa}) .$$

Let $X = p_{aaa} + p_{abb} - 1$

and $Y = (1-p_{aaa})$

then $p^k_{aab} = p^{k-1}_{aab}\, X + Y$

$$= (p^{k-2}_{aab}\, X + Y)\, X + Y$$

$$= p^{k-2}_{aab}\, X^2 + Y\,(1 + X)$$

$$\underset{k \Rightarrow \infty}{\text{Limit}}\ p^k_{aab} = \underset{k \Rightarrow \infty}{\text{Limit}}\ (p_{aab}\, X^{k-1} + Y\,(1 + X + \ldots + X^{k-2}))$$

$$p^k_{aab} = \frac{Y}{1 - X}$$

$$p^k_{aab} = \frac{1-p_{aaa}}{2 - p_{abb} - p_{aaa}} \qquad (11).$$

$$p^k_{bba} = p^{k-1}_{bba}\, p_{baa} + p^{k-1}_{bbb}\, p_{bba}$$

$$= p_{bba}^{k-1}\, p_{baa} + (1-p_{bba}^{k-1})\,(1-p_{bbb})$$

$$= p_{bba}^{k-1}\, p_{baa} + (1-p_{bbb}) - p_{bba}^{k-1}\,(1-p_{bbb})$$

$$= p_{bba}^{k-1}\,(p_{baa} + p_{bbb} - 1) + (1 - p_{bbb}) .$$

Let $X = p_{baa} + p_{bbb} - 1$

and $Y = (1-p_{bbb})$

then $p_{bba}^{k} = p_{bba}^{k-1}\, X + Y$

$$= (p_{bba}^{k-2}\, X + Y)\, X + Y$$

$$= p_{bba}^{k-2}\, X^2 + Y\,(1 + X)$$

$$= p_{bba}\, X^{k-1} + Y\,(1 + X + .. X^{k-2}) .$$

$$\underset{k \Rightarrow \infty}{\text{Limit}}\ p_{bba}^{k} = \underset{k \Rightarrow \infty}{\text{Limit}}\ (p_{bba}\, X^{k-1} + Y\,(1 + X + \ldots + X^{k-2}))$$

$$p_{bba}^{k} = \frac{Y}{1 - X}$$

$$p_{bba}^{k} = \frac{1-p_{bbb}}{2 - p_{baa} - p_{bbb}} \qquad (12).$$

$$p_{bbb}^{k} = p_{bba}^{k-1}\, p_{bab} + p_{bbb}^{k-1}\, p_{bbb}$$

$$= (1-p_{bbb}^{k-1})\,(1-p_{baa}) + p_{bbb}^{k-1}\, p_{bbb}$$

$$= (1-p_{baa}) - p_{bbb}^{k-1} (1-p_{baa}) + p_{bbb}^{k-1} p_{bbb}$$

$$= p_{bbb}^{k-1} (p_{bbb} + p_{baa} - 1) + (1-p_{baa}) .$$

If $X = p_{bbb} + p_{baa} - 1$

and $Y = 1-p_{baa}$

then $p_{bbb}^{k} = p_{bbb}^{k-1} X + Y$

$$= (p_{bbb}^{k-2} X + Y) X + Y$$

$$= p_{bbb}^{k-2} X^2 + Y (1 + X)$$

$$= p_{bbb} X^{k-1} + Y (1 + x + .. + X^{k-2}) .$$

$$\text{Limit}_{k \Rightarrow \infty} p_{bbb}^{k} = \text{Limit}_{k \Rightarrow \infty} (p_{bbb} X^{k-1} + Y (1 + X + \ldots + X^{k-2}))$$

$$p_{bbb}^{k} = \frac{Y}{1 - X}$$

$$p_{bbb}^{k} = \frac{1-p_{baa}}{2 - p_{bbb} - p_{baa}} \qquad (13).$$

$$p_{baa}^{k} = p_{baa}^{k-1} p_{aaa} + p_{bab}^{k-1} p_{aba}$$

$$= p_{baa}^{k-1} p_{aaa} + (1-p_{baa}^{k-1}) (1-p_{abb})$$

$$= p_{baa}^{k-1} p_{aaa} + (1-p_{abb}) - p_{baa}^{k-1} (1-p_{abb})$$

$$= p_{baa}^{k-1} (p_{aaa} + p_{abb} - 1) + (1 - p_{abb}) .$$

If $X = p_{aaa} + p_{abb} - 1$

and $Y = (1 - p_{abb})$

then $p_{baa}^{k} = p_{baa}^{k-1} X + Y$

$$= (p_{baa}^{k-2} X + Y) X + Y$$

$$= p_{baa}^{k-2} X^2 + Y (1 + X)$$

$$= p_{baa} X^{k-1} + Y (1 + X + \ldots + X^{k-2}) .$$

$$\text{Limit}_{k \Rightarrow \infty} p_{baa}^{k} = \text{Limit}_{k \Rightarrow \infty} (p_{baa} X^{k-1} + Y (1 + X + \ldots X^{k-2}))$$

$$p_{baa}^{k} = \frac{Y}{1 - X}$$

$$p_{baa}^{k} = \frac{1 - p_{abb}}{2 - p_{aaa} - p_{abb}} \qquad (14).$$

$$p_{bab}^{k} = p_{baa}^{k-1} p_{aab} + p_{bab}^{k-1} p_{abb}$$

$$= (1 - p_{bab}^{k-1}) (1 - p_{aaa}) + p_{bab}^{k-1} p_{abb}$$

$$= (1 - p_{aaa}) - p_{bab}^{k-1} (1 - p_{aaa}) + p_{bab}^{k-1} p_{abb}$$

$$= p_{bab}^{k-1} (p_{abb} + p_{aaa} - 1) + (1 - p_{aaa}) .$$

If $X = p_{abb} + p_{aaa} - 1$

and $Y = (1 - p_{aaa})$

then $p^k_{bab} = p^{k-1}_{bab} X + Y$

$= (p^{k-2}_{bab} X + Y) X + Y$

$= p^{k-2}_{bab} X^2 + Y (1 + X)$

$= p_{bab} X^{k-1} + Y (1 + X + \ldots + X^{k-2})$.

$\underset{k \Rightarrow \infty}{\text{Limit}}\ p^k_{bab} = \underset{k \Rightarrow \infty}{\text{Limit}}\ (p_{bab} X^{k-1} + Y (1 + X + \ldots + X^{k-2}))$

$$p^k_{bab} = \frac{Y}{1 - X}$$

$$p^k_{bab} = \frac{1 - p_{aaa}}{2 - p_{abb} - p_{aaa}} \qquad (15).$$

From equations (8) and (12)

$$p^k_{aba} = p^k_{bba} = \frac{1 - p_{bbb}}{2 - p_{baa} - p_{bbb}} \qquad (16).$$

From equations (9) and (13)

$$p^k_{abb} = p^k_{bbb} = \frac{1 - p_{baa}}{2 - p_{baa} - p_{bbb}} \qquad (17).$$

From equations (10) and (14)

$$p^k_{aaa} = p^k_{baa} = \frac{1-p_{abb}}{2 - p_{aaa} - p_{abb}} \qquad (18).$$

From equations (11) and (15)

$$p^k_{aab} = p^k_{bab} = \frac{1-p_{aaa}}{2 - p_{aaa} - p_{abb}} \qquad (19).$$

It is clear from (16), (17), (18), and (19),that at the intermediate steady state the effect upon the process by its state immediately preceding the present state is eliminated and the two 2 x 2 matrices are equal. They would appear as given below:

$$P_I = \begin{bmatrix} \underline{p^k_{aaa}} & \underline{p^k_{aab}} \\ \underline{p^k_{aba}} & \underline{p^k_{abb}} \\ \underline{p^k_{baa}} & \underline{p^k_{bab}} \\ \underline{p^k_{bba}} & \underline{p^k_{bbb}} \end{bmatrix} .$$

If the results of (16), (17), (18), and (19) are utilized for P_I, it would appear as given below:

$$P_I = \begin{bmatrix} \underline{p^k_{aa}} & \underline{p^k_{ab}} \\ \underline{p^k_{ba}} & \underline{p^k_{bb}} \\ \underline{p^k_{aa}} & \underline{p^k_{ab}} \\ \underline{p^k_{ba}} & \underline{p^k_{bb}} \end{bmatrix} .$$

From the above representation, it is observed that P_I consists of two identically equal 2 x 2 first order matrices. The steady state probabilities for these first order matrices can be computed using the same procedure for the determination of intermediate steady state probabilities. At the steady state all the probability vectors will be the same. This is due to the elimination of the effect of the process's present state. The steady state probability matrix P_s would appear as given below:

$$P_s = \begin{bmatrix} \underline{p_a^k} & \underline{p_b^k} \\ \underline{p_a^k} & \underline{p_b^k} \\ \underline{p_a^k} & \underline{p_b^k} \\ \underline{p_a^k} & \underline{p_b^k} \end{bmatrix}$$

where

$$p_a^k = \frac{1-p_{bb}^k}{2 - p_{aa}^k - p_{bb}^k} \qquad (20)$$

and

$$p_b^k = \frac{1-p_{aa}^k}{2 - p_{aa}^k - p_{bb}^k} \qquad (21).$$

Suffix $\underline{a}$ of p_a^k refers to the future outcome of the process. The future outcome of the process is completely independent of the process's present and past outcomes. The intermediate steady state probabilities for the numerical example having the P as given below can be

calculated using the above method.

$$P = \begin{bmatrix} p_{aaa} & p_{aab} \\ p_{aba} & p_{abb} \\ p_{baa} & p_{bab} \\ p_{bba} & p_{bbb} \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .6 & .4 \\ .5 & .5 \\ .3 & .7 \end{bmatrix} .$$

From equation (16)

$$p^k_{aba} = p^k_{bba} = p^k_{ba} = \frac{1 - p_{bbb}}{2 - p_{baa} - p_{bbb}}$$

$$= \frac{1 - .7}{2 - .5 - .7}$$

$$= 0.375 \quad .$$

From equation (17)

$$p^k_{abb} = p^k_{bbb} = p^k_{bb} = \frac{1 - p_{baa}}{2 - p_{baa} - p_{bbb}}$$

$$= \frac{1 - .5}{2 - .5 - .7}$$

$$= 0.625 \quad .$$

From equationa (18)

$$p^k_{aaa} = p^k_{baa} = p^k_{aa} = \frac{1-p_{abb}}{2 - p_{aaa} - p_{bbb}}$$

$$= \frac{1 - .4}{2 - .8 - .4}$$

$$= 0.75 .$$

and from equation (19)

$$p^k_{aab} = p^k_{bab} = p^k_{ab} = \frac{1-p_{aaa}}{2 - p_{aaa} - p_{abb}}$$

$$= \frac{1 - .8}{2 - .8 - .4}$$

$$= 0.25 .$$

Therefore the intermediate steady state probability matrix would be:

$$P_I = \begin{bmatrix} .75 & .25 \\ .375 & .625 \\ .75 & .25 \\ .375 & .625 \end{bmatrix} .$$

Finally, the steady state probabilities can be computed from equations (20) and (21).

$$p^k_a = \frac{1-p^k_{bb}}{2 - p^k_{aa} - p^k_{bb}}$$

$$= \frac{1 - .625}{2 - .75 - .625}$$

$$= 0.6 \quad .$$

$$p_b^k = \frac{1 - p_{aa}^k}{2 - p_{aa}^k - p_{bb}^k}$$

$$= \frac{1 - .75}{2 - .75 - .625}$$

$$= 0.4 \quad .$$

Therefore, the steady state probability matrix P_s would appear as given below:

$$P_s = \begin{bmatrix} .6 & .4 \\ .6 & .4 \\ .6 & .4 \\ .6 & .4 \end{bmatrix} .$$

The same value of P_s was obtained using the computer. The computer solution is given on page 31.

Steady State Conditions of Higher Order Markov Chains

The procedure for finding the steady state probability distributions of higher order Markov chains can be demonstrated by using a third order Markov chain. Let $(X_t;\ t = 0, 1, 2, \ldots)$ be a third order Markov chain defined in the state space S= (a, b) having the transition probability matrix P. P will have four dimensions for the third order Markov chain and the number of elements of P

will be 16. After the first reduction the intermediate steady state probabilities will be as follows:

$$p^k_{aaaa} = p^k_{baaa} = p^k_{aaa} = \frac{1-p_{aabb}}{2 - p_{aaaa} - p_{aabb}}$$

$$p^k_{aaab} = p^k_{baab} = p^k_{aab} = \frac{1-p_{aaaa}}{2 - p_{aaaa} - p_{aabb}}$$

$$p^k_{aaba} = p^k_{baba} = p^k_{aba} = \frac{1-p_{abbb}}{2 - p_{abaa} - p_{abbb}}$$

$$p^k_{aabb} = p^k_{babb} = p^k_{abb} = \frac{1-p_{abaa}}{2 - p_{abaa} - p_{abbb}}$$

$$p^k_{abaa} = p^k_{bbaa} = p^k_{baa} = \frac{1-p_{babb}}{2 - p_{baaa} - p_{babb}}$$

$$p^k_{abab} = p^k_{bbab} = p^k_{bab} = \frac{1-p_{baaa}}{2- p_{baaa} - p_{baaa}}$$

$$p^k_{abba} = p^k_{bbba} = p^k_{bba} = \frac{1-p_{bbbb}}{2 - p_{bbaa} - p_{bbbb}}$$

$$p^k_{abbb} = p^k_{bbbb} = p^k_{bbb} = \frac{1-p_{bbaa}}{2 - p_{bbaa} - p_{bbbb}} .$$

From the above it is noticed that the process is reduced to a second order chain after the first reduction. The process now depends only upon the present state and the state immediately preceding the present state. Two more reductions will lead to the steady state depending only upon the future outcomes. This has been shown in the second

order Markov chain example.

The knowledge of steady state probability distributions of second and third order Markov chains can be utilized for developing a generalized procedure to determine the steady state probability distributions of n-order Markov chains. Let $(X_t;\ t=0,\ 1,\ 2,\ldots)$ be an n-order Markov chain with the state space $S=(a,\ b)$ and the one-step transition probability matrix P. P will have n + 1 dimensions. The number of reductions necessary to determine the steady state condition of this n-order Markov chain is n. In the first reduction the dependency of the process upon its past outcomes is reduced from n states to n-1 states. Let π represent the ordered sequence of the outcomes of the process from t=1 to t= n-1. Then, $p^{k}_{a,\pi,a}$ represents the probability of the process being in the state $\underline{a}$ at the $(n+k)^{th}$ step given that it was in $\underline{a}$, and $\underline{\pi}$. The process has to pass intermediate steady states to reach the final steady state and each intermediate steady state will be reached at every k steps. For determining the intermediate steady state the process is considered to be a first order one as far as the state for which the intermediate steady state is required. Therefore,

$$p^{k}_{a,\pi,a} = p^{k-1}_{a,\pi,a}\, p_{\pi,a,a} + p^{k-1}_{a,\pi,b}\, p_{\pi,b,a}$$

$$= p^{k-1}_{a,\pi,a}\, p_{\pi,a,a} + (1-p^{k-1}_{a,\pi,a})\,(1-p_{\pi,b,b})$$

$$= p^{k}_{a,\pi,a}\,(p_{\pi,a,a} + p_{\pi,b,b} - 1) + (1 - p_{\pi,b,b}) \,.$$

Let $(p_{\pi,a,a} + p_{\pi,b,b} - 1) = X$

and $(1 - p_{\pi,b,b}) = Y$.

$$p^{k}_{a,\pi,a} = p^{k-1}_{a,\pi,a}\,(X) + Y$$

$$= (X\, p^{k-2}_{a,\pi,a} + Y)\, X + Y$$

$$= X^2\, p^{k-2}_{a,\pi,a} + Y\,(1 + X)$$

$$= X^{k-1} p_{a,\pi,a} + Y\,(1 + X + X^2 + \ldots + X^{k-2}) \,.$$

$$\text{Limit}_{k \Rightarrow \infty}\; p^{k}_{a,\pi,a} = \frac{Y}{1 - X}$$

$$= \frac{1 - p_{\pi,b,b}}{2 - p_{\pi,a,a} - p_{\pi,b,b}} \,.$$

Using the above procedure

$$p^{k}_{b,\pi,a} = \frac{1 - p_{\pi,b,b}}{2 - p_{\pi,a,a} - p_{\pi,b,b}} \,.$$

Thus

$$p^{k}_{a,\pi,a} = p^{k}_{b,\pi,a} = p^{k}_{\pi,a} = \frac{1 - p_{\pi,b,b}}{2 - p_{\pi,a,a} - p_{\pi,b,b}} \,.$$

The values of other elements are determined in a similiar manner. From the above result it is observed that the effect of one past state on the process's future outcome

is eliminated and (n-1) more such reductions will completely eliminate the effect of all the past states and the future outcome of the process is independent of the past outcomes.

The procedure for determining the steady state distribution of higher order Markov chains is applicable only when the process has two states. A reduction technique for the case when the process has three or more states is discussed along with an example. Let a second order Markov chain have the transition matrix P as given below.

$$P=\begin{bmatrix} p_{aaa} & p_{aab} & p_{aac} \\ p_{aba} & p_{abb} & p_{abc} \\ p_{aca} & p_{acb} & p_{acc} \\ p_{baa} & p_{bab} & p_{bac} \\ p_{bba} & p_{bbb} & p_{bbc} \\ p_{bca} & p_{bcb} & p_{bcc} \\ p_{caa} & p_{cab} & p_{cac} \\ p_{cba} & p_{cbb} & p_{cbc} \\ p_{cca} & p_{ccb} & p_{ccc} \end{bmatrix} = \begin{bmatrix} .3 & .4 & .3 \\ .2 & .2 & .6 \\ .3 & .3 & .4 \\ .2 & .7 & .1 \\ .4 & .3 & .3 \\ .6 & .3 & .1 \\ .6 & .2 & .2 \\ .5 & .4 & .1 \\ .1 & .8 & .1 \end{bmatrix}.$$

Since the process is a second order Markov chain, it has to pass through an intermediate steady state before reaching final steady state. The intermediate steady state would be found as given below:

$$\begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix} \times \begin{bmatrix} .3 & .4 & .3 \\ .2 & .2 & .6 \\ .3 & .3 & .4 \end{bmatrix} = \begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix} \quad (22).$$

$$a+b+c=1 \quad (23).$$

$$\begin{bmatrix} d & e & f \\ d & e & f \\ d & e & f \end{bmatrix} \times \begin{bmatrix} .2 & .7 & .1 \\ .4 & .3 & .3 \\ .6 & .3 & .1 \end{bmatrix} = \begin{bmatrix} d & e & f \\ d & e & f \\ d & e & f \end{bmatrix} \quad (24).$$

$$d+e+f=1 \quad (25).$$

$$\begin{bmatrix} g & h & i \\ g & h & i \\ g & h & i \end{bmatrix} \times \begin{bmatrix} .6 & .2 & .2 \\ .5 & .4 & .1 \\ .1 & .8 & .1 \end{bmatrix} = \begin{bmatrix} g & h & i \\ g & h & i \\ g & h & i \end{bmatrix} \quad (26).$$

$$g+h+i=1 \quad (27).$$

From the matrix equations (22) and (23) a,b, and c are determined. Equations (24) and (25) are used to determine d,e, and f. The quantities g,h, and i are determined from equations (26) and (27). Therefore, the intermediate steady state distribution would be as given below:

$$P=\begin{bmatrix} \underline{a} & \underline{b} & \underline{c} \\ \underline{d} & \underline{e} & \underline{f} \\ \underline{g} & \underline{h} & \underline{i} \\ \underline{a} & \underline{b} & \underline{c} \\ \underline{d} & \underline{e} & \underline{f} \\ \underline{g} & \underline{h} & \underline{i} \\ \underline{a} & \underline{b} & \underline{c} \\ \underline{d} & \underline{e} & \underline{f} \\ \underline{g} & \underline{h} & \underline{i} \end{bmatrix}.$$

From the above it is observed that the effect of one past state is eliminated. Therefore the intermediate steady state transition matrix P_I appears as given below:

$$P_I=\begin{bmatrix} \underline{a} & \underline{b} & \underline{c} \\ \underline{d} & \underline{e} & \underline{f} \\ \underline{g} & \underline{h} & \underline{i} \end{bmatrix}=\begin{bmatrix} \underline{0.273} & \underline{0.300} & \underline{0.427} \\ \underline{0.365} & \underline{0.446} & \underline{0.189} \\ \underline{0.490} & \underline{0.362} & \underline{0.148} \end{bmatrix}.$$

P_I is equivalent to a first order matrix and the following matrix equations are used to determine the final steady state.

$$\begin{bmatrix} \underline{x} & \underline{y} & \underline{z} \\ \underline{x} & \underline{y} & \underline{z} \\ \underline{x} & \underline{y} & \underline{z} \end{bmatrix} X \begin{bmatrix} \underline{a} & \underline{b} & \underline{c} \\ \underline{d} & \underline{e} & \underline{f} \\ \underline{g} & \underline{h} & \underline{i} \end{bmatrix} = \begin{bmatrix} \underline{x} & \underline{y} & \underline{z} \\ \underline{x} & \underline{y} & \underline{z} \\ \underline{x} & \underline{y} & \underline{z} \end{bmatrix}$$

$$x+y+z=1 .$$

Therefore, the steady state distribution P_s will be

$$P_s = \begin{bmatrix} \underline{x} & \underline{y} & \underline{z} \\ \underline{x} & \underline{y} & \underline{z} \\ \underline{x} & \underline{y} & \underline{z} \end{bmatrix} = \begin{bmatrix} \underline{0.341} & \underline{0.372} & \underline{0.287} \\ \underline{0.341} & \underline{0.372} & \underline{0.287} \\ \underline{0.341} & \underline{0.372} & \underline{0.287} \end{bmatrix}.$$

From the above P_s, it is observed that the effect of all past states is eliminated at the steady state. A reduction technique for n-order Markov chains will be developed after presenting an example for third order Markov chains with three states. Let $(X_t;\ t=0,1,2,\ldots)$ be a third order Markov chain with the state space $S=(a,b,c)$ and the transition matrix P. In interpreting the element $p_{aaba,}$ for example, of P given below, it is the conditional probability of the process being in $\underline{a}$ at t=3, given that it was in $\underline{a}$,$\underline{a}$, and $\underline{b}$ at t=0, 1, and 2 respectively.

$$P = \begin{vmatrix} p_{aaaa} & p_{aaab} & p_{aaac} \\ p_{aaba} & p_{aabb} & p_{aabc} \\ p_{aaca} & p_{aacb} & p_{aacc} \\ p_{baaa} & p_{baab} & p_{baac} \\ p_{baba} & p_{babb} & p_{babc} \\ p_{baca} & p_{bacb} & p_{bacc} \\ p_{caaa} & p_{caab} & p_{caac} \\ p_{caba} & p_{cabb} & p_{cabc} \\ p_{caca} & p_{cacb} & p_{cacc} \\ p_{abaa} & p_{abab} & p_{abac} \\ p_{abba} & p_{abbb} & p_{abbc} \\ p_{abca} & p_{abcb} & p_{abcc} \\ p_{bbaa} & p_{bbab} & p_{bbac} \\ p_{bbba} & p_{bbbb} & p_{bbbc} \\ p_{bbca} & p_{bbcb} & p_{bbcc} \\ p_{cbaa} & p_{cbab} & p_{cbac} \\ p_{cbba} & p_{cbbb} & p_{cbbc} \\ p_{cbca} & p_{cbcb} & p_{cbcc} \\ p_{acaa} & p_{acab} & p_{acac} \\ p_{acba} & p_{acbb} & p_{acbc} \\ p_{acca} & p_{accb} & p_{accc} \end{vmatrix} = \begin{vmatrix} 0.6 & 0.3 & 0.1 \\ 0.5 & 0.2 & 0.3 \\ 0.4 & 0.1 & 0.5 \\ 0.7 & 0.1 & 0.2 \\ 0.6 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.3 & 0.3 & 0.4 \\ 0.2 & 0.3 & 0.5 \\ 0.5 & 0.1 & 0.4 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.8 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.2 & 0.5 \\ 0.4 & 0.4 & 0.2 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.3 & 0.4 \\ 0.7 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.2 & 0.7 & 0.1 \end{vmatrix}.$$

$$\begin{bmatrix} p_{bcaa} & p_{bcab} & p_{bcac} \\ p_{bcba} & p_{bcbb} & p_{bcbc} \\ p_{bcca} & p_{bccb} & p_{bccc} \\ p_{caaa} & p_{caab} & p_{caac} \\ p_{caba} & p_{cabb} & p_{cabc} \\ p_{caca} & p_{cacb} & p_{cacc} \end{bmatrix} \quad \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.4 & 0.5 & 0.1 \\ 0.6 & 0.1 & 0.3 \\ 0.8 & 0.1 & 0.1 \\ 0.7 & 0.1 & 0.2 \\ 0.2 & 0.7 & 0.2 \end{bmatrix}$$

In P there are 3 submatrices for every step. For all the 3 steps, there are 9 submatrices. At the first intermediate steady state the three submatrices at every step will give one submatrix. This is because of the reduction of the effect of one past outcome on the process. The reduction of the 9 submatrices into 3 submatrices is as given below:

$$\begin{bmatrix} p_{aaaa} & p_{aaab} & p_{aaac} \\ p_{aaba} & p_{aabb} & p_{aabc} \\ p_{aaca} & p_{aacb} & p_{aacc} \\ p_{baaa} & p_{baab} & p_{baac} \\ p_{baba} & p_{babb} & p_{babc} \\ p_{baca} & p_{bacb} & p_{bacc} \\ p_{caaa} & p_{caab} & p_{caac} \\ p_{caba} & p_{cabb} & p_{cabc} \\ p_{caca} & p_{cacb} & p_{cacc} \end{bmatrix} \Rightarrow \begin{bmatrix} p_{aaa} & p_{aab} & p_{aac} \\ p_{aba} & p_{abb} & p_{abc} \\ p_{aca} & p_{acb} & p_{acc} \end{bmatrix} = \begin{bmatrix} .530 & .229 & .241 \\ .577 & .222 & .201 \\ .329 & .300 & .371 \end{bmatrix}$$

$$\begin{bmatrix} p_{abaa} & p_{abab} & p_{abac} \\ p_{abba} & p_{abbb} & p_{abbc} \\ p_{abca} & p_{abcb} & p_{abcc} \\ p_{bbaa} & p_{bbab} & p_{bbac} \\ p_{bbba} & p_{bbbb} & p_{bbbc} \\ p_{bbca} & p_{bbcb} & p_{bbcc} \\ p_{cbaa} & p_{cbab} & p_{cbac} \\ p_{cbba} & p_{cbbb} & p_{cbbc} \\ p_{cbca} & p_{cbcb} & p_{cbcc} \end{bmatrix} \Longrightarrow \begin{bmatrix} p_{baa} & p_{bab} & p_{bac} \\ p_{bba} & p_{bbb} & p_{bbc} \\ p_{bca} & p_{bcb} & p_{bcc} \end{bmatrix} = \begin{bmatrix} .370 & .293 & .337 \\ .205 & .538 & .257 \\ .437 & .312 & .251 \end{bmatrix}$$

$$\begin{bmatrix} p_{acaa} & p_{acab} & p_{acac} \\ p_{acba} & p_{acbb} & p_{acbc} \\ p_{acca} & p_{accb} & p_{accc} \\ p_{bcaa} & p_{bcab} & p_{bcac} \\ p_{bcba} & p_{bcbb} & p_{bcbc} \\ p_{bcca} & p_{bccb} & p_{bccc} \\ p_{ccaa} & p_{ccab} & p_{ccac} \\ p_{ccba} & p_{ccbb} & p_{ccbc} \\ p_{ccca} & p_{cccb} & p_{cccc} \end{bmatrix} \Rightarrow \begin{bmatrix} p_{caa} & p_{cab} & p_{cac} \\ p_{cba} & p_{cbb} & p_{cbc} \\ p_{cca} & p_{ccb} & p_{ccc} \end{bmatrix} = \begin{bmatrix} .288 & .556 & .156 \\ .472 & .403 & .125 \\ .712 & .170 & .118 \end{bmatrix}.$$

The reduced transition matrix P_1 after the first reduction is, therefore,

$$P_1 = \begin{bmatrix} p_{aaa} & p_{aab} & p_{aac} \\ p_{aba} & p_{abb} & p_{abc} \\ p_{aca} & p_{acb} & p_{acc} \\ p_{baa} & p_{bab} & p_{bac} \\ p_{bba} & p_{bbb} & p_{bbc} \\ p_{bca} & p_{bcb} & p_{bcc} \\ p_{caa} & p_{cab} & p_{cac} \\ p_{cba} & p_{cbb} & p_{cbc} \\ p_{cca} & p_{ccb} & p_{ccc} \end{bmatrix} = \begin{bmatrix} .530 & .229 & .241 \\ .577 & .222 & .201 \\ .329 & .300 & .371 \\ .370 & .293 & .337 \\ .205 & .538 & .257 \\ .437 & .312 & .251 \\ .288 & .556 & .156 \\ .472 & .403 & .125 \\ .712 & .170 & .118 \end{bmatrix}.$$

In the second reduction the above three submatrices will be reduced to a single 3 x 3 submatrix as given below:

$$P_1 = \begin{bmatrix} \underline{p_{aa}} & \underline{p_{ab}} & \underline{p_{ac}} \\ \underline{p_{ba}} & \underline{p_{bb}} & \underline{p_{bc}} \\ \underline{p_{ca}} & \underline{p_{cb}} & \underline{p_{cc}} \end{bmatrix} = \begin{bmatrix} \underline{.488} & \underline{.246} & \underline{.266} \\ \underline{.324} & \underline{.395} & \underline{.281} \\ \underline{.427} & \underline{.436} & \underline{.137} \end{bmatrix} .$$

After the second reduction the process becomes a first order Markov chain, i.e., its future outcome depends only upon its present outcome. In the third reduction the above matrix will have equal probability vectors. The final steady state probability matrix would be as givne below:

$$P_s = \begin{bmatrix} \underline{p_a} & \underline{p_b} & \underline{p_c} \\ \underline{p_a} & \underline{p_b} & \underline{p_c} \\ \underline{p_a} & \underline{p_b} & \underline{p_c} \end{bmatrix} = \begin{bmatrix} \underline{.315} & \underline{.361} & \underline{.324} \\ \underline{.315} & \underline{.361} & \underline{.324} \\ \underline{.315} & \underline{.361} & \underline{.324} \end{bmatrix} .$$

Interpreting P_s, the future outcome of the process does not depend upon its past outcomes.

The knowledge acquired from the above examples can be used to generalize the procedure for determining the steady state distribution of an n-order, m-state Markov chain. Let $(X_t; t=0, 1, 2 \ldots)$ be an n-order Markov chain with the state space S= (a,b, ..., m) and the transition probability matrix P. P will have n x m submatrices. This n-order Markov chain will need n reductions to reach

the steady state. The reduction ratio will be m: 1, i.e., if there are m submatrices now, then in the next intermediate steady state there will be only one submatrix. For all the n reductions the procedure discussed for the reduction of third order, 3-state Markov chains is used.

Analysis of Higher Order Absorbing Markov Chains

The following information may be obtained from the analysis of higher order absorbing Markov chains:

1. The expected number of steps the process is in any given non-absorbing state.
2. The probability of absorption by any given absorbing state.

The transition probability matrix P is rearranged into four submatrices as the first step in the analysis. After rearranging P, it will appear as indicated below:

$$P = \begin{array}{c|c} I & O \\ \hline A & N \end{array}$$

The submatrices I, O, A, and N are interpreted in the same way as in the first order case. Before attempting to obtain the above mentioned information about the absorption properties, the concept of identity matrix for higher order matrices is developed. An identity matrix for a second order matrix P for the state space S= (a,b) can be determined from the matrix concept $P \times I = P$. It must be remembered that $P \times I \neq I \times P$ in the case of higher

order matrices.

If P and I are represented as:

$$P = \begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{abb}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} .$$

$$I = \begin{bmatrix} \underline{I_{aaa}} & \underline{I_{aab}} \\ \underline{I_{aba}} & \underline{I_{abb}} \\ \underline{I_{baa}} & \underline{I_{bab}} \\ \underline{I_{bba}} & \underline{I_{bbb}} \end{bmatrix} .$$

$$P \times I = P$$

i.e.,

$$\begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{abb}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} \times \begin{bmatrix} \underline{I_{aaa}} & \underline{I_{aab}} \\ \underline{I_{aba}} & \underline{I_{abb}} \\ \underline{I_{baa}} & \underline{I_{bab}} \\ \underline{I_{bba}} & \underline{I_{bbb}} \end{bmatrix} = \begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{abb}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} .$$

If the above equations are solved the identity matrix I is obtained as shown below:

$$I = \begin{bmatrix} \underline{1} & \underline{0} \\ \underline{0} & \underline{1} \\ \underline{1} & \underline{0} \\ \underline{0} & \underline{1} \end{bmatrix} .$$

The above procedure is used for matrices of any size.

The expected number of times the process will be in a non-absorbing state S_j is the sum of the following terms:

Expected number of times in S_j = (1) (probability of being in S_j at start) + (1) (probability of being in S_j after one step) + (1) (probability of being in S_j after 2 steps) + ...

$= I + N + N^2 + \ldots$

As k becomes large, N^k approaches zero. In the case of the first order chain the quantity $I + N + N^2 + \ldots$ is a geometric series and is equivalent to $(I-N)^{-1}$. In the higher order chains even though N^k approaches zero for large value of k and the quantity $I + N + N^2 + \ldots$ seems to be a geometric series and equals to $(I-N)^{-1}$, considerable difficulty is faced in determining the inverse of higher order matrix. Therefore, a method along with a numerical example is discussed for solving the series. Let $(X_t;\ t=0, 1, 2, \ldots)$ be a second order Markov chain with the state space S= (a,b,c) and the transition probability matrix P with c as the absorbing state as indicated below:

$$P = \begin{bmatrix} p_{aaa} & p_{aab} & p_{aac} \\ p_{aba} & p_{abb} & p_{abc} \\ p_{aca} & p_{acb} & p_{acc} \\ p_{baa} & p_{bab} & p_{bac} \\ p_{bba} & p_{bbb} & p_{bbc} \\ p_{bca} & p_{bcb} & p_{bcc} \\ p_{caa} & p_{cab} & p_{cac} \\ p_{cba} & p_{cbb} & p_{cbc} \\ p_{cca} & p_{ccb} & p_{ccc} \end{bmatrix} = \begin{bmatrix} .3 & .3 & .4 \\ .2 & .2 & .6 \\ 0 & 0 & 1 \\ .5 & .4 & .1 \\ .4 & .2 & .4 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

The submatrix N of P will be as given below:

$$N = \begin{bmatrix} p_{aaa} & p_{aab} \\ p_{aba} & p_{abb} \\ p_{baa} & p_{bab} \\ p_{bba} & p_{bbb} \end{bmatrix} = \begin{bmatrix} .3 & .3 \\ .2 & .2 \\ .5 & .4 \\ .4 & .2 \end{bmatrix}.$$

For this N, the expected number of times that the process being in every nonabsorbing state E is given by the series

$$E = I + N + N^2 + N^3 + \ldots$$

$$= \begin{bmatrix} \underline{1} & \underline{0} \\ \underline{0} & \underline{1} \\ \underline{1} & \underline{0} \\ \underline{0} & \underline{1} \end{bmatrix} + \begin{bmatrix} \underline{\sum_{i=1}^{n} p_{aaa}^{i}} & \underline{\sum_{i=1}^{n} p_{aab}^{i}} \\ \underline{\sum_{i=1}^{n} p_{aba}^{i}} & \underline{\sum_{i=1}^{n} p_{abb}^{i}} \\ \underline{\sum_{i=1}^{n} p_{baa}^{i}} & \underline{\sum_{i=1}^{n} p_{bab}^{i}} \\ \underline{\sum_{i=1}^{n} p_{bba}^{i}} & \underline{\sum_{i=1}^{n} p_{bbb}^{i}} \end{bmatrix} .$$

To find $\sum_{i=1}^{\infty} p_{aaa}^{i}$, $\sum_{i=1}^{\infty} p_{aab}^{i}$, $\sum_{i=1}^{\infty} p_{aba}^{i}$, $\sum_{i=1}^{\infty} p_{abb}^{i}$, $\sum_{i=1}^{\infty} p_{baa}^{i}$, $\sum_{i=1}^{\infty} p_{bab}^{i}$, $\sum_{i=1}^{\infty} p_{bba}^{i}$, and $\sum_{i=1}^{\infty} p_{bbb}^{i}$, the following procedure is used:

$$p_{aba} = p_{aba}$$

$$p_{aba}^{2} = p_{aba}\, p_{baa} + p_{abb}\, p_{bba}$$

$$p_{aba}^{3} = p_{aba}\, p_{baa}^{2} + p_{abb}\, p_{bba}^{2}$$

$$p_{aba}^{n} = p_{aba}\, p_{baa}^{n-1} + p_{abb}\, p_{bba}^{n-1}$$

$$\sum_{i=1}^{n} p_{aba}^{i} = p_{aba} \sum_{i=1}^{n-1} p_{baa}^{i} + p_{abb} \sum_{i=1}^{n-1} p_{bba}^{i} .$$

As $n \Rightarrow \infty$:

$$\sum_{i=1}^{\infty} p_{aba}^{i} = p_{aba} \sum_{i=1}^{\infty} p_{baa}^{i} + p_{abb} \sum_{i=1}^{\infty} p_{bba}^{i} \quad (28) .$$

Similarly:

$$\sum_{i=1}^{\infty} p_{abb}^{i} = p_{aba} \sum_{i=1}^{\infty} p_{bab}^{i} + p_{abb} (1+ \sum_{i=1}^{\infty} p_{bbb}^{i}) \quad (29)$$

$$\sum_{i=1}^{\infty} p_{aaa}^{i} = p_{aaa} (1 + \sum_{i=1}^{\infty} p_{aaa}^{i}) + p_{aab} \sum_{i=1}^{\infty} p_{aba}^{i} \quad (30)$$

$$\sum_{i=1}^{\infty} p^{i}{}_{aab} = p_{aaa} \sum_{i=1}^{\infty} p_{aab}^{i} + p_{aab} (1 + \sum_{i=1}^{\infty} p_{abb}^{i}) \quad (31)$$

$$\sum_{i=1}^{\infty} p_{bba}^{i} = p_{bba} (1 + \sum_{i=1}^{\infty} p_{baa}^{i}) + p_{bbb} \sum_{i=1}^{\infty} p_{bba}^{i} \quad (32)$$

$$\sum_{i=1}^{\infty} p_{bbb}^{i} = p_{bba} \sum_{i=1}^{\infty} p_{bab}^{i} + p_{bbb} (1 + \sum_{i=1}^{\infty} p_{bbb}^{i}) \quad (33)$$

$$\sum_{i=1}^{\infty} p_{baa}^{i} = p_{baa} (1 + \sum_{i=1}^{\infty} p_{aaa}^{i}) + p_{bab} \sum_{i=1}^{\infty} p_{aba}^{i} \quad (34)$$

and

$$\sum_{i=1}^{\infty} p_{bab}^{i} = p_{baa} \sum_{i=1}^{\infty} p_{aab}^{i} + p_{bab} (1 + \sum_{i=1}^{\infty} p_{abb}^{i}) \quad (35) .$$

The above eight simultaneous equations are solved for the eight unknowns.

$$(28) \times p_{bba} - (32) \times p_{aba} = p_{bba} \sum_{i=1}^{\infty} p_{aba}^{i} - p_{aab} \sum_{i=1}^{\infty} p_{bba}^{i}$$

$$= (p_{bba} p_{abb} - p_{aba} p_{bbb} \sum_{i=1}^{\infty} p_{bba}^{i}$$

therefore,

$$\sum_{i=1}^{\infty} p_{aba}^{i} = ((p_{aba} + p_{bba}\, p_{abb} - p_{abb}\, p_{bba})/p_{bba}) \sum_{i=1}^{\infty} p_{bba}^{i} \qquad (36).$$

$$(30) \times p_{bab} - (34) \times p_{aab} = p_{bab} \sum_{i=1}^{\infty} p_{aaa}^{i} - p_{aab} \sum_{i=1}^{\infty} p_{baa}^{i}$$

$$= (1 + \sum_{i=1}^{\infty} p_{aaa}^{i})(p_{aaa} p_{bab} - p_{baa}\, p_{aab}) .$$

Therefore,

$$\sum_{i=1}^{\infty} p_{aaa}^{i} = (p_{aaa}\, p_{bab} - p_{aaa}\, p_{aab})/(p_{baa}\, p_{aab} - p_{aaa}\, p_{bab} + p_{bab}) + ((p_{aab})/(p_{bab} + p_{baa}\, p_{aab} - p_{aaa}\, p_{bab})) \sum_{i=1}^{\infty} p_{baa}^{i} \qquad (37).$$

$$(30) \times p_{baa} - (34) \times p_{aaa} = p_{baa} \sum_{i=1}^{\infty} p_{aaa}^{i} - p_{aaa} \sum_{i=1}^{\infty} p_{baa}^{i}$$

$$= (p_{baa}\, p_{aab} - p_{aaa}\, p_{bab}) \times (\sum_{i=1}^{\infty} p_{aba}^{i}) \qquad (38).$$

Substituting (37) for $\sum_{i=1}^{\infty} p_{aaa}^{i}$ in (38)

$$p_{baa} (\sum_{i=1}^{\infty} p_{baa}^{i}\, p_{aab} + p_{aaa}\, p_{bab} - p_{baa}\, p_{aab})/(p_{bab} + p_{baa}\, p_{aab} - p_{aaa}\, p_{bab}) - p_{aa} \sum_{i=1}^{\infty} p_{baa}^{i} = (p_{baa}\, p_{aab} - p_{aaa}\, p_{bab}) \sum_{i=1}^{\infty} p_{aba}^{i} \qquad (39).$$

Substituting (32) for $\sum_{i=1}^{\infty} p_{bba}^{i}$ in (28)

$$\sum_{i=1}^{\infty} p_{aba}^{i} = (1 + \sum_{i=1}^{\infty} p_{baa}^{i})(p_{aba} - p_{aba}\, p_{bbb} + p_{abb}\, p_{bba})/(1 - p_{bbb}) \qquad (40).$$

Substituting (40) for $\sum_{i=1}^{\infty} p_{aba}^{i}$ in (39)

$$p_{baa}\, (\sum_{i=1}^{\infty} p_{baa}^{i}\, p_{aab} + p_{aaa}\, p_{bab} - p_{baa}\, p_{aab})/ (p_{bab} + p_{baa}\, p_{aab} - p_{aaa}\, p_{bab}) - p_{aaa} \sum_{i=1}^{\infty} p_{baa}^{i}$$

$$= (p_{baa}\, p_{aab} - p_{aaa}\, p_{bab})\, (1 + \sum_{i=1}^{\infty} p_{baa}^{i})\, (p_{aba} - p_{aba}\, p_{bbb} + p_{abb}\, p_{bba})/(1-p_{bbb}) \qquad (41).$$

In equation (41), $\sum_{i=1}^{\infty} p_{baa}^{i}$ is expressed in terms of one-step probabilities. Substituting the values of these probabilities for the numerical example in (41)

$$\sum_{i=1}^{\infty} p_{aaa}^{i} \simeq 1.100 \quad .$$

Substituting the values of $\sum_{i=1}^{\infty} p_{baa}^{i}$, and the one-step probabilities in (32),

$$\sum_{i=1}^{\infty} p_{bba}^{i} \simeq 1.05 \quad .$$

Substituting the value of $\sum_{i=1}^{\infty} p_{bba}^{i}$ in (36)

$$\sum_{i=1}^{\infty} p_{aba}^{i} \simeq 0.63 \quad .$$

From (37)

$$\sum_{i=1}^{\infty} p_{aaa}^{i} \simeq 0.70 \quad .$$

Similarly, $\sum_{i=1}^{\infty} p_{aab}^{i} \simeq 0.66$, $\sum_{i=1}^{\infty} p_{abb}^{i} \simeq 0.53$,

$\sum_{i=1}^{\infty} p_{bab}^{i} \simeq 0.94$, and $\sum_{i=1}^{\infty} p_{bbb}^{i} \simeq 0.72$.

Therefore,

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.70 & 0.66 \\ 0.63 & 0.53 \\ 1.10 & 0.94 \\ 1.05 & 0.72 \end{bmatrix} = \begin{bmatrix} 1.70 & 0.66 \\ 0.63 & 1.53 \\ 2.10 & 0.94 \\ 1.05 & 1.72 \end{bmatrix} .$$

A FORTRAN IV program for $I + N + N^2 + \ldots$ was run in IBM/360, the result obtained (page 73) was the same as the above result. From E, it is observed that the expected number of times that the process being in the non-absorbing state a given that it was in state a previously and is in a now, is 1.70. The other elements of E are identified in the same way.

The above method can be used to formulate a generalized procedure for determining the expected number of times that an n-order chain will be in the non-absorbing state before it is absorbed. Let $(X_t;\ t=0, 1, 2, \ldots)$ be an n-order Markov chain with the state space S= (a,b,...,m). If m is the absorbing state then the non-absorbing

submatrix N will have $(m-1)^n$ elements in it. This will lead to $(m-1)^n$ equations in $(m-1)^n$ unknowns, and these equations can be solved for the unknowns.

To find the probability of absorption by any given absorbing state, a similar logic is used in the analysis. Let S_j signify some given absorbing state; let S_i signify some specified non-absorbing state.

Probability of ending in S_j = (probability of going from S_i to S_j in 1 step) + (probability of going from S_i to S_j in 2 steps) + ... In the first order Markov chains this can be obtained from the series $A + NA + N^2A + \ldots$ without difficulty. In the case of the higher order chains, care must be taken in determing the factors of the series. Let $(X_t;\ t=0, 1, \ldots)$ be a second order Markov chain with S= (a,b,c). If c is the absorbing state then N and A would be as given below:

$$N = \begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{abb}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} \quad \text{and}$$

$$A = \begin{bmatrix} \underline{p_{aac}} \\ \underline{p_{abc}} \\ \underline{p_{bac}} \\ \underline{p_{bbc}} \end{bmatrix} .$$

In the series $A + NA + N^2A + \ldots$, the elements NA, $N^2A, \ldots$ are obtained by using n-dimensional matrix algebra. If the probability of the process ending in S_j can be obtained from $A + NA + N^2A + \ldots$, then N^2A, for example, must give the probability of the process being absorbed in exactly three steps for various given conditions. If the probability of the process being absorbed in exactly three steps given that it was in $\underline{a}$ at $t=0$ and is in $\underline{a}$ at $t=1$ is p^3_{aac} then, according to N^2A this would be equal to

$p_{aaa}\, p_{aaa}\, p_{aac} + p_{aab}\, p_{aba}\, p_{aac} + p_{aaa}\, p_{aab}\, p_{abc}$
$+ p_{aab}\, p_{abb}\, p_{abc}$. According to the classical method p^3_{aac} will be equal to

$p_{aaa}\, p_{aaa}\, p_{aac} + p_{aab}\, p_{aba}\, p_{bac} + p_{aaa}\, p_{aab}\, p_{abc}$
$+ p_{aab}\, p_{abb}\, p_{bbc}$. Therefore, in the higher order case, the series $A + NA + N^2A + \ldots$ will not give the required absorption probabilities if the above method of determining the factors of the series is used. According to the n-dimensional matrix algebra

$A + NA + N^2A + N^3A + \ldots = A + NA + N(NA) + N(N(NA)) + \ldots$

Therefore, to determine N^2A, for example, NA is determined first and then NA is multiplied by N, ie., the order of multiplication is from right to left. The value of N^2A, determined in this way, can be shown to be the same as that of the classical method.

$$N^2A = N(NA) = \begin{bmatrix} \underline{p_{aaa}} & \underline{p_{aab}} \\ \underline{p_{aba}} & \underline{p_{aba}} \\ \underline{p_{baa}} & \underline{p_{bab}} \\ \underline{p_{bba}} & \underline{p_{bbb}} \end{bmatrix} \times \begin{bmatrix} \underline{p^2_{aac}} & \underline{p^3_{aac}} \\ \underline{p^2_{abc}} & \underline{p^3_{abc}} \\ \underline{p^2_{bac}} & \underline{p^3_{bac}} \\ \underline{p^2_{bbc}} & \underline{p^3_{bbc}} \end{bmatrix} .$$

From the above p^3_{aac}, the probability of the process being absorbed given that it was in a at t=0 and is in a at t=1, will be equal to $p_{aaa}\, p^2_{aac} + p_{aab}\, p^2_{abc}$. Substituting for the values of p^2_{aac} and p^2_{abc}

$$p^3_{aac} = p_{aaa}\, p_{aaa}\, p_{aac} + p_{aaa}\, p_{aab}\, p_{abc} + p_{aab}\, p_{aba}\, p_{bac} + p_{aab}\, p_{abb}\, p_{bbc}.$$

One more example will be discussed before generalizing for n-order Markov chains. Let $(X_t;\ t=0, 1, 2, \ldots)$ be a second order Markov chain with the state space S= (a,b,c,d,e) and the transition matrix P as given below:

$$
P = \begin{vmatrix}
p_{aaa} & p_{aab} & p_{aac} & p_{aad} & p_{aae} \\
p_{aba} & p_{abb} & p_{abc} & p_{abd} & p_{abe} \\
p_{aca} & p_{acb} & p_{acc} & p_{acd} & p_{ace} \\
p_{ada} & p_{adb} & p_{adc} & p_{add} & p_{ade} \\
p_{aea} & p_{aeb} & p_{aec} & p_{aed} & p_{aee} \\
p_{baa} & p_{bab} & p_{bac} & p_{bad} & p_{bae} \\
p_{bba} & p_{bbb} & p_{bbc} & p_{bbd} & p_{bbe} \\
p_{bca} & p_{bcb} & p_{bcc} & p_{bcd} & p_{bce} \\
p_{bda} & p_{bdb} & p_{bdc} & p_{bdd} & p_{bde} \\
p_{bea} & p_{beb} & p_{bec} & p_{bed} & p_{bee} \\
p_{caa} & p_{cab} & p_{cac} & p_{cad} & p_{cae} \\
p_{cba} & p_{cbb} & p_{cbc} & p_{cbd} & p_{cbe} \\
p_{cca} & p_{ccb} & p_{ccc} & p_{ccd} & p_{cce} \\
p_{cda} & p_{cdb} & p_{cdc} & p_{cdd} & p_{cde} \\
p_{cea} & p_{ceb} & p_{cec} & p_{ced} & p_{cee} \\
p_{daa} & p_{dab} & p_{dac} & p_{dad} & p_{dae} \\
p_{dba} & p_{dbb} & p_{dbc} & p_{dbd} & p_{dbe} \\
p_{dca} & p_{dcb} & p_{dcc} & p_{dcd} & p_{dce} \\
p_{dda} & p_{ddb} & p_{ddc} & p_{ddd} & p_{dde} \\
p_{dea} & p_{deb} & p_{dec} & p_{ded} & p_{dee} \\
p_{eaa} & p_{eab} & p_{eac} & p_{ead} & p_{eae}
\end{vmatrix}
=
\begin{vmatrix}
0.2 & 0.2 & 0.3 & 0.1 & 0.2 \\
0.4 & 0.1 & 0.2 & 0.2 & 0.1 \\
0.3 & 0.3 & 0.1 & 0.2 & 0.1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0.2 & 0.1 & 0.3 & 0.1 & 0.3 \\
0.1 & 0.4 & 0.2 & 0.2 & 0.1 \\
0.1 & 0.3 & 0.1 & 0.3 & 0.2 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0.3 & 0.1 & 0.4 & 0.1 & 0.1 \\
0.1 & 0.3 & 0.3 & 0.1 & 0.2 \\
0.4 & 0.2 & 0.1 & 0.2 & 0.1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{vmatrix}.
$$

$$\begin{vmatrix} p_{eba} & p_{ebb} & p_{ebc} & p_{ebd} & p_{ebe} \\ p_{eca} & p_{ecb} & p_{ecc} & p_{ecd} & p_{ece} \\ p_{eda} & p_{edb} & p_{edc} & p_{edd} & p_{ede} \\ p_{eea} & p_{eeb} & p_{eec} & p_{eed} & p_{eee} \end{vmatrix} \quad \begin{vmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

For the above absorbing states d, and e of P, the submatrices N and A are as given below:

$$N = \begin{vmatrix} 0.2 & 0.2 & 0.3 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.3 & 0.1 \\ 0.2 & 0.1 & 0.3 \\ 0.1 & 0.4 & 0.2 \\ 0.1 & 0.3 & 0.1 \\ 0.3 & 0.1 & 0.4 \\ 0.1 & 0.3 & 0.3 \\ 0.4 & 0.2 & 0.1 \end{vmatrix} \qquad A = \begin{vmatrix} 0.1 & 0.2 \\ 0.2 & 0.1 \\ 0.2 & 0.1 \\ 0.1 & 0.3 \\ 0.2 & 0.1 \\ 0.3 & 0.2 \\ 0.1 & 0.1 \\ 0.1 & 0.2 \\ 0.2 & 0.1 \end{vmatrix} .$$

A FORTRAN IV program for A + NA + N (NA) + N (N(NA)) +... was run in IBM/360 and the result was found to be

(page 74).

$$A + NA + N(NA) + N(N(NA)) + \ldots = \begin{bmatrix} 0.466 & 0.533 \\ 0.534 & 0.465 \\ 0.555 & 0.444 \\ 0.413 & 0.586 \\ 0.586 & 0.413 \\ 0.552 & 0.447 \\ 0.515 & 0.484 \\ 0.482 & 0.517 \\ 0.558 & 0.441 \end{bmatrix} .$$

Each row in the solution matrix of A + NA + N (NA) + N (N (NA)) + .. is a probability vector. Therefore, theoretically, the elements of each row vector must sum to one. The difference between the row sum determined by the computer and 1 depends upon how many terms of the series A + NA + N (NA) + N (N (NA)) + .. are used by the computer to compute the row sum.

From the above two numerical examples a procedure can be developed for determining the absorption probabilities for n-order Markov chains. Let $(X_t; t=0, 1, 2, ..)$ be an n-order Markov chain with the state space S= (a,b,.., j, j + 1, .. m). For the absorbing states j + 1, .., m, the submatrix N will have n dimensions and j number of j x j submatrices and the submatrix A will have n dimensions and j number of (m - j) x (m - j) submatrices. The absorption probabilities are obtained from the series A + NA + N (NA) + N (N(NA)) +....

A FORTRAN IV Program for Determining the Expected Number of Times that a Second Order Markov Chain is in any Non-Absorbing State:

```
      DIMENSION A(8),P(8),R(8),S(8)
      DATAS(1),S(2),S(3),S(4),S(5),S(6),S(7),S(8)/0.0,0.0,0.0,0.0,0.0,
     10.0,0.0,0.0/
      READ(5,1)A
    1 FORMAT(8F4.2)
      DO2I=1,27
      IF(I.GT.1)GOTO3
      DO5J=1,8
    5 P(J)=A(J)
      GOTO4
    3 R(1)=A(1)*P(7)+A(2)*P(5)
      R(2)=A(1)*P(8)+A(2)*P(6)
      R(3)=A(3)*P(3)+A(4)*P(1)
      R(4)=A(3)*P(4)+A(4)*P(2)
      R(5)=A(5)*P(7)+A(6)*P(5)
      R(6)=A(5)*P(8)+A(6)*P(6)
      R(7)=A(7)*P(3)+A(8)*P(1)
      R(8)=A(7)*P(4)+A(8)*P(2)
      DO6K=1,8
    6 P(K)=R(K)
    4 WRITE(6,7)P
      DO8L=1,8
    8 S(L)=S(L)+P(L)
    2 CONTINUE
    7 FORMAT(5X,8(F8.6,2X))
      WRITE(6,7)S
      STOP
      END
```

```
$ENTRY
0.200000  0.200000  0.300000  0.300000  0.400000  0.200000  0.500000  0.400000
0.180000  0.120000  0.150000  0.150000  0.280000  0.200000  0.230000  0.230000
0.102000  0.086000  0.099000  0.081000  0.148000  0.132000  0.147000  0.123000
0.059000  0.051000  0.060300  0.050100  0.088400  0.075600  0.090300  0.074900
0.035740  0.030100  0.035790  0.030330  0.053800  0.045080  0.053750  0.045450
0.021510  0.018106  0.021459  0.018129  0.032260  0.027196  0.032191  0.027205
0.012890  0.010880  0.012891  0.010870  0.019328  0.016321  0.019333  0.016307
0.007732  0.006526  0.007734  0.006525  0.011599  0.009787  0.011601  0.009787
0.004640  0.003915  0.004640  0.003915  0.006960  0.005872  0.006960  0.005873
0.002784  0.002349  0.002784  0.002349  0.004176  0.003524  0.004176  0.003524
0.001670  0.001409  0.001670  0.001409  0.002506  0.002114  0.002506  0.002114
0.001002  0.000846  0.001002  0.000846  0.001503  0.001268  0.001503  0.001268
0.000601  0.000507  0.000601  0.000507  0.000902  0.000761  0.000902  0.000761
0.000361  0.000304  0.000361  0.000304  0.000541  0.000457  0.000541  0.000457
0.000216  0.000183  0.000216  0.000183  0.000325  0.000274  0.000325  0.000274
0.000130  0.000110  0.000130  0.000110  0.000195  0.000164  0.000195  0.000164
0.000078  0.000066  0.000078  0.000066  0.000117  0.000099  0.000117  0.000099
0.000047  0.000039  0.000047  0.000039  0.000070  0.000059  0.000070  0.000059
0.000028  0.000024  0.000028  0.000024  0.000042  0.000036  0.000042  0.000036
0.000017  0.000014  0.000017  0.000014  0.000025  0.000021  0.000025  0.000021
0.000010  0.000009  0.000010  0.000009  0.000015  0.000013  0.000015  0.000013
0.000006  0.000005  0.000006  0.000005  0.000009  0.000008  0.000009  0.000008
0.000004  0.000003  0.000004  0.000003  0.000005  0.000005  0.000005  0.000005
0.000002  0.000002  0.000002  0.000002  0.000003  0.000003  0.000003  0.000003
0.000001  0.000001  0.000001  0.000001  0.000002  0.000002  0.000002  0.000002
0.000001  0.000001  0.000001  0.000001  0.000001  0.000001  0.000001  0.000001
0.000000  0.000000  0.000000  0.000000  0.000001  0.000001  0.000001  0.000001
0.630471  0.532398  0.698772  0.656741  1.050776  0.720663  1.101563  0.941329
```

A FORTRAN IV Program for the Determination of the Absorption Probabilities of a Second Order Markov Chain:

```
      DIMENSIONA(27),P(18),R(18),S(18)
      DO3M=1,18
    3 S(M)=0.0
      READ(5,4)P
    4 FORMAT(18F4.2)
      READ(5,1)A
    1 FORMAT(20F4.2)
      DO2I=1,27
      IF(I.EQ.1)GOTO5
      R(1)=A(1)*P(1)+A(2)*P(3)+A(3)*P(5)
      R(2)=A(1)*P(2)+A(2)*P(4)+A(3)*P(6)
      R(3)=A(4)*P(7)+A(5)*P(9)+A(6)*P(11)
      R(4)=A(4)*P(8)+A(5)*P(10)+A(6)*P(12)
      R(5)=A(7)*P(13)+A(8)*P(15)+A(9)*P(17)
      R(6)=A(7)*P(14)+A(8)*P(16)+A(9)*P(18)
      R(7)=A(10)*P(1)+A(11)*P(3)+A(12)*P(5)
      R(8)=A(10)*P(2)+A(11)*P(4)+A(12)*P(6)
      R(9)=A(13)*P(7)+A(14)*P(9)+A(15)*P(11)
      R(10)=A(13)*P(8)+A(14)*P(10)+A(15)*P(12)
      R(11)=A(16)*P(13)+A(17)*P(15)+A(18)*P(17)
      R(12)=A(16)*P(14)+A(17)*P(16)+A(18)*P(18)
      R(13)=A(19)*P(1)+A(20)*P(3)+A(21)*P(5)
      R(14)=A(19)*P(2)+A(20)*P(4)+A(21)*P(6)
      R(15)=A(22)*P(7)+A(23)*P(9)+A(24)*P(11)
      R(16)=A(22)*P(8)+A(23)*P(10)+A(24)*P(12)
      R(17)=A(25)*P(13)+A(26)*P(15)+A(27)*P(17)
      R(18)=A(25)*P(14)+A(26)*P(16)+A(27)*P(18)
      DO6K=1,18
    6 P(K)=R(K)
    5 DO2L=1,18
    2 S(L)=S(L)+P(L)
    7 FORMAT(5X,9(F8.6,2X))
      WRITE(6,7)S
      STOP
      END
```

```
     $ENTRY
  0.466911  0.533057  0.534480  0.465493  0.555460  0.444507  0.413464  0.586508  0.586344
 .413630  0.552322  0.447655  0.515699  0.484263  0.482942  0.517033  0.558736  0.441230
```

CHAPTER IV

CONCLUSION

This research was primarily concerned with the presentation of quantitative information about the solution of higher order Markov chains. Part one reviewed the work done in the first order Markov chains. Part two developed a model for n-order Markov chains along with the concept of n-dimensional matrix algebra. Part three analyzed the steady state and absorption characteristics of higher order chains. Since the study was broken into three parts, the summary will follow the same general plan.

n-Dimensional Matrix Algebra and Solution of Higher Order Markov Chains

Three dimensional matrices were analyzed with respect to how they can be multiplied. The concept of three dimensional matrix algebra was extended to n-dimensional matrices. Second order Markov chains were solved by the matrix method and its computational superiority over the classical probability theory was proven. When the concept of n-dimensional matrix algebra was used to solve higher

order Markov chains, it was observed that $P^k = P \times P^{k-1} \neq P^{k-1} \times P$.

Steady State and Absorptions Characteristics

It was observed that at sufficiently large number of steps the probability distribution of higher order ergodic Markov chains is independent of its past and present outcomes. A reduction technique was used as an analytical method in determining the steady probabilities. For an n-order Markov chain n reductions were required to reach the steady state. It was observed that for every reduction the number of states the process depends upon for its future outcomes was reduced by one, and at the steady state the process depends only upon its future outcome.

Difficulty was experienced in determining the inverse of higher order matrices, and therefore a new method called the simultaneous equations method was developed to analyze the absorption characteristics. In determining the absorption probabilities it was observed that the series $A + NA + N^2A + \ldots$ did not give the required result if the order of multiplication for every factor of the series was from left to right. The logical way to determine the value of N^2A, for example, would be to determine NA first and then multiply N by NA, ie., $N^2A = N\,(NA)$. According to the order of multiplication the series $A + NA + N^2A + N^3A + \ldots$ would be equal to $A + NA + N\,(NA) + N\,(N\,(NA)) + \ldots$

BIBLIOGRAPHY

1. Barlow, R.E., and F. Proschan. Mathematical Theory of Reliability, Wiley, New York, 1965.

2. Breiman, Leo. Probability and Stochastic Processes with a View Toward Applications, Houghton Mifflin Company, Boston, 1969.

3. Chung, K.L. Markov Chains with Stationary Transition Probabilities, Springer-Verlag, Berlin, 1960.

4. Feller, W. An Introduction to Probability Theory and its Applications, Wiley, New York, 2nd ed., 1957.

5. Hillier, F.S., and G.J. Lieberman. Introduction to Operations Research, Holden-Day, San Francisco, 1967.

6. Howard, R. Dynamic Programming and Markov Processes, M.I.T. Press, Cambridge, Mass., 1960.

7. Karlin, Samuel. A First Course in Stochastic Processes, Academic Press, New York, 1966.

8. Kemeny, J.G. and J.L. Snell, Finite Markov Chains, Van Nostrand, Princeton, N.J., 1959.

9. Parzen, E. Stochastic Processes, Holden-Day, San Francisco, 1962.

VITA

3

Sengoda Gounder Ganesan

Candidate for the Degree of

Doctor of Philosophy

Thesis: ANALYSIS OF HIGHER ORDER MARKOV CHAINS

Major Field: Industrial Engineering and Management

Biographical:

Personal Data: Born in Sappaiyapuram, Salem, Tamil Nadu, Idia, June 22, 1942, the son of Sengoda Gounder and Pappa.

Education: Graduated from Board High School, Attayampatti, Salem, Tamil Nadu, India, in March, 1959; received a Bachelor of Engineering degree from Government College of Technology, Coimbatore of Madras University, India, with a major in Mechanical Engineering in January, 1965; received a Master of Science degree from P.S.G. College of Technology, Coimbatore of Madras University, India, with a major in Mechanical Engineering, in December, 1966; completed requirements for the Doctor of Philosophy degree in May, 1970.

Professional Experience: Graduate Teaching Assistant, Oklahoma State University, December 1968 to May 1970.

Professional Memberships: American Institute of Industrial Engineers, and Alpha Pi Mu.