AGGREGATION OF CONDITIONAL ABSORBING MARKOV CHAINS

C. Bernard Barfoot
AGGREGATION OF CONDITIONAL ABSORBING MARKOV CHAINS

C. Bernard Barfoot
Foreword

This paper was presented to the Sixth European Meeting on Cybernetics and Systems Research, which was held at the University of Vienna, April 13-16, 1982.
AGGREGATION OF CONDITIONAL ABSORBING MARKOV CHAINS

C. Bernard Barfoot
Center for Naval Analyses
Alexandria, Virginia U.S.A.

When modeling a process by means of a finite Markov chain, it is sometimes necessary or desirable to stratify the process into subprocesses and model each of these subprocesses. The resulting Markov chain for each subprocess becomes a conditional Markov chain in that its transition probabilities are relative to its associated subprocess. This paper derives the method for aggregating conditional absorbing Markov chains (each of which has the same state space) into a single (unconditional) chain that is representative of the total process and has the same state space as the conditional chains.

1. INTRODUCTION

When modeling a process by means of a finite Markov chain, it is sometimes necessary or desirable to stratify the process into subprocesses and model each of these individual subprocesses. For example, in a study of a distributed data base system [1], the flow of data was modeled as a Markov chain for several separate geographic locations. In a recruiting study [2], the movement of military-age men through the recruiting process and into the armed forces was modeled for separate racial and educational groups as a Markov chain with a single state space for each group—only the input data (transition probabilities) to the model were changed for each group. In [3], a Markov chain model was used to investigate the consequences of induced abortion for different groups of women by estimating transition probabilities separately for each group.

When the above procedure is used, the resulting Markov chain for each subprocess becomes a conditional Markov chain in that its transition probabilities are relative to its associated subprocess. This paper derives the method for aggregating these separate conditional chains (each of which has the same state space) into a single (unconditional) chain that is representative of the total process and has the same state space as the conditional chains.

2. ILLUSTRATIVE EXAMPLE

Figure 1 illustrates a simple four-state process that has been stratified into two subprocesses V₁ and V₂. Suppose data have been collected for each subprocess (which might represent different geographic regions or different groups of people, for example) and transition probabilities have been estimated as shown in the matrices of transition probabilities P₁ and P₂. Further,
suppose we know the fraction of time (i.e., the probability) that the process originates in each subprocess, say $f_1$ and $f_2$, and also know that the process always begins in state $S_1$. Given this information, how do we determine $P$? This example illustrates the general problem addressed in this paper. As we shall see later in the paper, what might be considered as two "obvious" methods of determining $P$ do not, in general, work:

1. aggregating the data from the subprocesses
2. defining $P = f_1P_1 + f_2P_2$.

3. DERIVATION OF $P$

Suppose the process $\mathcal{F}$ under study can be stratified into $m$ subprocesses $\mathcal{F}_k$, $(k = 1, \ldots, m)$, each with the same state space. We assume that $\mathcal{F}$ and all the $\mathcal{F}_k$ are being modeled as an irreducible, finite, absorbing Markov chain having $q$ transient states and $r$ absorbing states.

If we number the states of the chain so that the transient states "precede" the absorbing states, then the matrix $F_k$ of transition probabilities for subprocess $\mathcal{F}_k$ can be partitioned as follows:

$$F_k = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

where $1$ is the identity matrix of order $r$, $0$ is the $r \times q$ zero matrix, $F_k$ is the $q \times r$ matrix containing the transition probabilities from transient to absorbing states, and $Q_k$ is the matrix of order $q$ containing the transition probabilities among the transient states.

The probability that $\mathcal{F}_k$ is absorbed in state $S_j$ $(j = 1, \ldots, r)$, given that the process began in transient state $S_i$ $(i = 1, \ldots, q)$, is [4]

$$b_{ik} = \begin{bmatrix} b_{ik} \\ 0 \end{bmatrix} = (I - Q_k)^{-1}R_k = N_kR_k$$

where $N_k$ is the matrix that gives the expected number $N_{ik}$ of times that $\mathcal{F}_k$ is in each transient state $S_j$ given that it began in each transient state $S_i$.

We assume, without loss of generality, that each $\mathcal{F}_k$ always begins in a particular transient state, say $S_1$. Then the probability that $\mathcal{F}_k$ terminates in $S_j$ is

$$b_{1k} = e_{1}^{T}b_{k} = e_{1}^{T}N_{k}R_{k}$$

where $e_{1}$ is the unit column vector $e_{1} = (\delta_{11})$, $\delta_{11}$ is the Kronecker delta, and $e_{1}^{T}$ is the transpose of $e_{1}$.

If we let $f_k$ be the probability that $\mathcal{F}$ originates in $\mathcal{F}_k$ ($\mathcal{F}_k = 1$), then the probability that the process terminates in $S_j$ is

$$b_{1} = \sum_{k=1}^{m} f_{k}b_{1k} = e_{1}^{T}f_{k}(I - Q_k)^{-1}R_{k}.$$ 

Our criterion for determining $P$ is that the limiting probabilities of absorption obtained from $P$ must be the same as those obtained from $F_k$ and $f_k$. In other words, we want to determine the stochastic matrix $P$ of order $q + r$ such that

$$P = (P_{ij}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

and $e_{1}^{T}(I - Q)^{-1}R = e_{1}^{T}f_{k}(I - Q_k)^{-1}R_{k}$.

To obtain an expression for $P$, let

$$N_{ik} = \text{a diagonal matrix of order } q \text{ whose diagonal elements are from the first row of } N_k$$

$$A_k = \begin{bmatrix} e_{1}^{T}f_{k} \end{bmatrix} = \text{a diagonal matrix of order } q + r.$$ 

Then, as we shall prove in subsequent theorems, the matrix $P$ that satisfies our criterion is

$$P = (C_{k}f_{k}A_{k})^{-1}e_{1}^{T}f_{k}A_{k}P_{k}.$$ 

(1)

Also, if $P$ is given by (1), then the
submatrices $Q$ and $R$ are

$$Q = (z_k f_k Q k_{1k})^{-1} z_k f_k W_k Q_k$$
and $R = (z_k f_k W_k k_{1k})^{-1} z_k f_k W_k B_k$.

4. VERIFICATION OF $P$
To show that $P$ is in fact the desired matrix we need to show that:

1. $P$ is stochastic; i.e., $P_{ij} > 0$, all $i$ and $j$, and $e_1 p_{ij} = 1$, all $i$.
2. $e_1 (I - Q)^{-1}R = e_1 z_k f_k (I - Q)^{-1}B_k$.

We show that these conditions are met in the following two theorems.

Theorem 1. $P$ is stochastic.

Proof. From equation (1) it follows that $P_{ij} > 0$, all $i$ and $j$, since each term is nonnegative.

To show that $e_1 p_{ij} = 1$, all $i$, we need to show that $Pe = e$, where $e$ is the $q + r$ column vector all of whose elements are unity.

Writing $P$ and $e$ in partitioned form,

$$P = \begin{bmatrix} 0 & e_1 \\ e_2 & e_3 \\ e_4 & e_5 \end{bmatrix} = \begin{bmatrix} e_1 + Re_1 \\ e_2 \\ e_3 \end{bmatrix}$$

Now

$$Q e_1 + Re_1 = (z_k f_k W_k k_{1k})^{-1} z_k f_k W_k (Q e_1 + R e_1)$$

$$= (z_k f_k W_k k_{1k})^{-1} z_k f_k W_k e_1 = e_1$$

Hence the left hand side becomes

$$e_1 (I - Q)^{-1}R = e_1 z_k f_k (I - Q)^{-1}B_k$$

which completes the proof.

Another quantity of interest is the expected time to absorption in a given absorbing state
S_j, given that \( \mathbb{E}_k \) began in transient state \( S_k \). The matrix of transition probabilities for \( \mathbb{E}_k \) conditioned on the hypothesis that \( \mathbb{E}_k \) is absorbed in \( S_j \) is [4]

\[
P_{jk} = \begin{bmatrix}
Q_{jk} & e' (I - Q_{jk}) \\
0 & 1
\end{bmatrix}
\]

where \( Q_{jk} = D_j^{-1} Q_j D_{jk} \)

and \( D_{jk} = \begin{bmatrix}
D_{1jk} & 0 \\
0 & \cdots & 0
\end{bmatrix} \) is a diagonal matrix of order \( q \) formed from column \( j \) of \( D_k \). \( P_{jk} \) is of order \( q + 1 \), \( e' (I - Q_{jk}) \) is a \( q \)-component column vector, and \( 0 \) is the \( q \)-component zero row vector.

Given that \( \mathbb{E}_k \) is absorbed in state \( S_j \), the expected time to absorption is [1]

\[
V_{jk} = (I - Q_{jk})^{-1} T_k = N_{jk} T_k
\]

where \( T_k = (E_{1k}) \) is a column vector whose elements are the expected times \( E_{1k} \) that \( \mathbb{E}_k \) spends in each transient state \( S_k \).

Since \( \mathbb{E}_k \) always begins in transient state \( S_1 \), the expected time for \( \mathbb{E}_k \) to be absorbed in state \( S_j \) is

\[
e_{1j} V_{jk} = e_{1j} (I - Q_{jk})^{-1} T_k
\]

and the expected time for the overall process \( f \) to be absorbed in state \( S_j \) is

\[
e_{1j} V_{jk} = e_{1j} (I - Q_{jk})^{-1} T_k
\]

To determine the matrix \( P_j \) and the time vector \( T \) for \( f \) under the hypothesis that \( f \) is absorbed in \( S_j \), we use the criterion that the expected time to absorption obtained by using \( P_j \) and \( T \) must be the same as the time obtained by using \( P_{jk} \), \( T_k \), and \( f_k \). In other words, we want to determine \( P_j \) and \( T \) such that

\[
P_j = \begin{bmatrix}
Q_j & e' (I - Q_j) \\
0 & 1
\end{bmatrix}
\]

and \( e_{1j} (I - Q_j)^{-1} T = e_{1j} f_k (I - Q_{jk})^{-1} T_k \).

To obtain expressions for \( P_j \) and \( T \), let

\[
N_{jk} = \begin{bmatrix}
N_{jk} & 0 \\
0 & 1
\end{bmatrix}
\]

and \( A_{jk} \) form \( q \)-component column vectors.

Then

\[
P_j = (f_k A_{jk})^{-1} f_k A_{jk} P_j ,
\]

\[
Q_j = (f_k N_{jk} A_{jk})^{-1} f_k N_{jk} A_{jk} P_j,
\]

and \( T = (f_k N_{jk} A_{jk})^{-1} f_k N_{jk} A_{jk} T_k \).

The proofs that \( P_j \) is stochastic and that

\[
e_{1j} (I - Q_j)^{-1} T = e_{1j} f_k (I - Q_{jk})^{-1} T_k
\]

are the same as those given in Theorems 1 and 2.

REFERENCES


PP 211  Mangel, Marc, "On Singular Characteristic Initial Value Problems with Unique Solution," 20 pp., Jun 1978, AD A054 539


PP 216  Mangel, Marc, "Diagonalization by Group Matrices," 26 pp., Apr 78, AD A054 443

PP 217  Mangel, Marc, "Bibliometric Studies of Scientific Productivity," 17 pp., Mar 78 (Presented at the Annual meeting of the American Society for Information Science held in San Francisco, California, October 1976), AD A054 442

PP 218 - Classified

PP 219  Hettinger, R. Later, "Market Analysis with Rational Expectations: Theory and Estimation," 60 pp., Apr 78, AD A054 422

PP 220  Hemen, Donald E., "Diagonalization of Group Matrices," 26 pp., Apr 78, AD A054 443


PP 223  Mangel, Marc, "Mathematical Problems in Feynman's Path Integrals," 30 pp., Jun 1978, AD A054 227


Ulguoff, Kathy, Clossen, and Broaching, Frank, "Taxes and Inflation," 25 pp., Nov 1979, AD A081 194


Thomason, James S., "Support Dependence and Inter-State Cooperation: The Case of Sub-Saharan Africa," 141 pp., Jan 1980, AD A081 193


Goldberg, Lawrence, "Recruiters Advertising and Navy Enlistments," 34 pp., Mar 1980, AD A082 221

Goldberg, Lawrence, "Delaying an Overhaul and Ship's Equipment," 40 pp., Mar 1980, AD A085 095


"The Graduate School of Mathematical Sciences, University of Rochester and the Center for Naval Analyses "The Graduate School of Mathematical Sciences, University of Rochester


Cape, Denis, "Limit Cycle Solutions of Reaction-Diffusion Equations," 35 pp., Jun 1980, AD A087 114


"University of Florida"