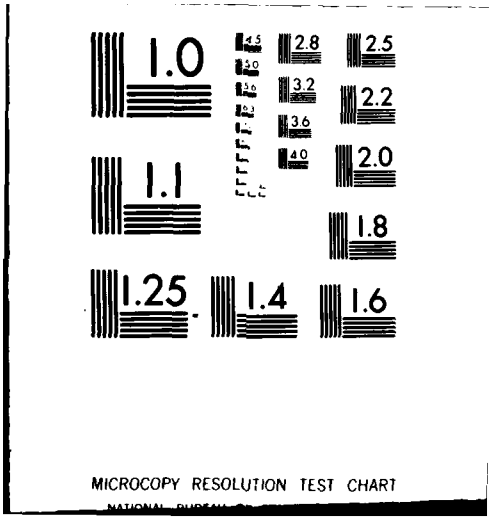
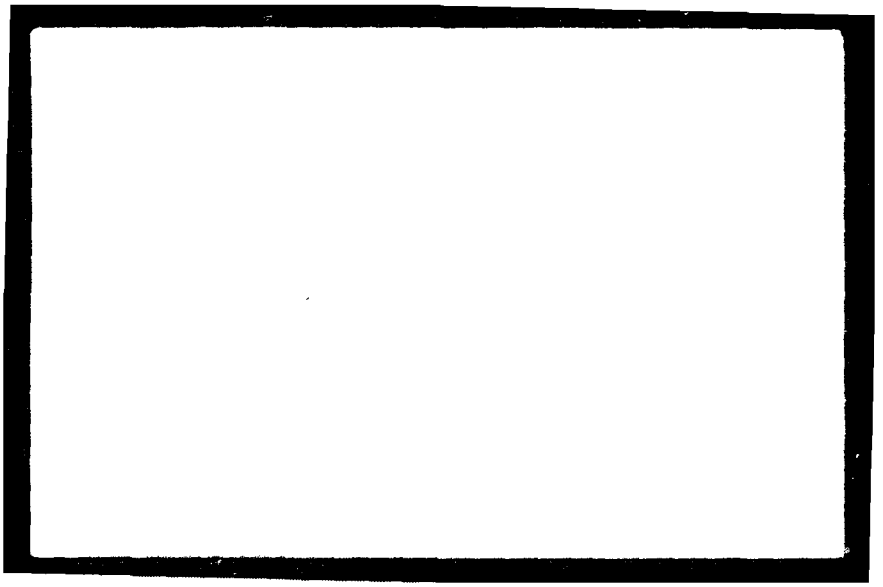
AD-A089 069 VIRGINIA POLYTECHNIC INST AND STATE UNIV BLACKSBURG --ETC F/G 1 /1
GENERALIZED INVERSES AND THEIR APPLICATION TO APPLIED PROBABILI-- TC(U)
MAY 80 J J HUNTER N00014-77-C-0743
UNCLASSIFIED VTR-80-06 NL

END
DATE FILMED
10-80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART



GENERALIZED INVERSES AND THEIR APPLICATION TO APPLIED PROBABILITY PROBLEMS

by

Jeffrey J. Hunter

Department of Industrial Engineering and Operations Research
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Tech Report
VTR-8006

May 1980

Written while on leave from the Department of Mathematics, University of Auckland, New Zealand and supported in part by a Visiting Professorship within the Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University. This paper was prepared partly under the support of ONR Contract N00014-77-C-0743 (NR042-296) and NSF Grant ENG77-22757. Distribution of this document is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report VTR-8006
2. GOVT ACCESSION NO.: AD-A089069
3. RECIPIENT'S CATALOG NUMBER:

(6) 4. TITLE (and Subtitle): GENERALIZED INVERSES AND THEIR APPLICATION TO APPLIED PROBABILITY PROBLEMS

5. TYPE OF REPORT & PERIOD COVERED: Technical Report 1980

6. PERFORMING ORG. REPORT NUMBER:

(15) 7. AUTHOR(s): (10) Jeffrey J. Hunter (11) May 80

8. CONTRACT OR GRANT NUMBER(s): N00014-77-C-0743, NSF-[illegible]-7[illegible]; NR042-296

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Dept. of Industrial Engineering & Operations Res. Virginia Polytechnic Inst. & State University Blacksburg, Virginia 24061

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (12) 59

11. CONTROLLING OFFICE NAME AND ADDRESS: Director, Statistics and Probability Program Mathematical & Information Sciences Division 800 N. Quincy St., Arlington, VA 22217

12. REPORT DATE: May 1980

13. NUMBER OF PAGES: 59

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): (14) VTR-80-06

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

(9) Technical Rept.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Generalized Inverses
Markov chains
Markov renewal processes
Stationary distributions
Passage time moments

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The main aim of this paper is to examine the applicability of generalized inverses to a wide variety of problems in applied probability where a Markov chain is present either directly or indirectly through some form of imbedding. By characterizing all generalized inverses of I-P, where P is the transition matrix of a finite irreducible discrete time Markov chain, we are able to obtain general procedures for finding stationary distributions, moments of the first passage time distributions and asymptotic forms for the moments of

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406 747

the occupation time random variables. It is shown that all known explicit methods for examining these problems can be expressed in this generalized inverse framework. More generally, in the context of a Markov renewal process setting the aforementioned problems are also examined using generalized inverses of I-P. As a special case Markov chains in continuous time are considered and we show that the generalized inverse technique can be applied direct to the infinitesimal generator of the process, instead of to I-P, where P is the transition matrix of the discrete time jump Markov chain.

ABSTRACT

The main aim of this paper is to examine the applicability of generalized inverses to a wide variety of problems in applied probability where a Markov chain is present either directly or indirectly through some form of imbedding. By characterizing all generalized inverses of I-P, where P is the transition matrix of a finite irreducible discrete time Markov chain, we are able to obtain general procedures for finding stationary distributions, moments of the first passage time distributions and asymptotic forms for the moments of the occupation time random variables. It is shown that all known explicit methods for examining these problems can be expressed in this generalized inverse framework. More generally, in the context of a Markov renewal process setting the aforementioned problems are also examined using generalized inverses of I-P. As a special case Markov chains in continuous time are considered and we show that the generalized inverse technique can be applied direct to the infinitesimal generator of the process, instead of to I-P, where P is the transition matrix of the discrete time jump Markov chain.

1. Introduction

In many stochastic modelling situations a Markov chain in discrete time is often present, either directly or indirectly through some form of imbedding. The most general framework that has been devised for handling such situations is the concept of a Markov renewal process together with its associated semi-Markov process. Invariably in an attempt to extract information concerning the behavior of such processes (either for the Markov chain in discrete time or for the more general continuous time process) we are interested in obtaining: a) stationary distributions, b) expressions for the moments of the first passage time distributions and c) asymptotic forms for the expected number of visits to each state (occupation time random variables in the Markov chain setting or, more generally, the Markov renewal counting random variables).

In section 2 we discuss these three general problem areas and show that they are all intimately related to the form and structure of P, the transition matrix of the imbedded discrete time Markov chain.

Quite often, once some preliminary analysis has been carried out, we find that the properties of interest can be obtained by solving a system of linear equations involving P, or more particularly I-P where I is the identity matrix. One condition generalized inverses (commonly called g-inverses) are shown, in section 3, to be extremely useful in handling such types of equations. Although g-inverses are not necessarily unique we are able to characterize all g-inverses of I-P when P is finite and irreducible.

In the remaining three sections of the paper the problems a), b) and c) alluded to earlier are investigated for Markov chains in discrete time, Markov renewal processes (or the associated semi-Markov process) and Markov chains in continuous time. In this latter process, a special Markov renewal process, its

stochastic behavior is governed by an infinitesimal generator Q and we show that the generalized inverse technique can be applied direct to Q instead of to I-P where P is the transition matrix of the imbedded discrete time "jump" chain.

No numerical or computational studies have been carried out using the methods presented but there are some compelling reasons why the procedures given in the paper should prove to be useful. Firstly, many computer subroutines for producing generalized inverses are now available and since many of the techniques developed herein require no special form for the generalized inverse any such package can be used without any modification. Secondly, if some additional information is known concerning the structure of the generalized inverse in the subroutine the solution of the problem at hand may be considerably simplified. Furthermore, any known explicit methods for solving these problems, that have already appeared in the literature, can be expressed in the framework presented in this paper. It is the universality of the results presented herein that should prove to be valuable.

2. Models and Problems of Interest

2.1. Markov chains in discrete time

Let $\{X_n\}$, $(n \geq 0)$, be a discrete time, finite, homogeneous, Markov chain with state space S taken to be $\{1,2,\cdots,m\}$. Let $P = [p_{ij}]$ be the transition matrix of the chain which, for the rest of the paper, will be assumed to be irreducible. We make no assumption whether the Markov chain is aperiodic or periodic and the analysis presented covers both of these cases.

It is well known (Feller [6], Kemeny and Snell [9]) that for such chains a stationary (or invariant) probability distribution $\{\pi_i\}$ $(i = 1,2,\cdots,m)$ exists,

is unique, and satisfies the equations

$$\pi_j = \sum_{i=1}^{m} \pi_i p_{ij} \quad \text{and} \quad \sum_{i=1}^{m} \pi_i = 1. \tag{2.1}$$

If $\underset{\sim}{\pi}' = (\pi_1, \pi_2, \cdots, \pi_m)$ and $\underset{\sim}{e}' = (1,1,\cdots,1)$, equation (2.1) can be expressed as

$$\underset{\sim}{\pi}'(I - P) = \underset{\sim}{0}' \text{ subject to } \underset{\sim}{\pi}'\underset{\sim}{e} = 1. \tag{2.2}$$

Thus the stationary probability vector $\underset{\sim}{\pi}'$ can be determined by solving a constrained system of linear equations involving I - P, a singular matrix. We examine general procedures for solving such a system in section 3.2 and apply these procedures to the problem at hand in section 4.1.

Let T_{ij} be the number of trials for a first passage from state i to state j (first return if i = j), i.e.,

$$T_{ij} = \min\{n: X_n = j | X_0 = i\}.$$

Under the assumption of irreducibility for finite chains the T_{ij} are proper random variables, the moments $m_{ij}^{(k)} \equiv E\, T_{ij}^k$ are well defined and finite for all $i,j \in S$ (cf., Theorem 13.4.1 of Neuts [14]). Furthermore, it can be shown (Kemeny and Snell [9]) that the mean first passage time from state i to state j, $m_{ij} \equiv m_{ij}^{(1)}$ satisfies the equation

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik}\, m_{kj}, \quad (i,j \in S). \tag{2.3}$$

If we use the notation $M = [m_{ij}]$, $M_d = [\delta_{ij}\, m_{ij}]$ and $E = [1]$ (i.e., all the elements of E are unity), equations (2.3) can be expressed as

$$(I - P)M = E - PM_d. \tag{2.4}$$

Once again we have a system of linear equations involving the matrix $I - P$. As we shall see, (section 3.2), when we solve such equations involving singular matrices there is some arbitrariness in the solution. However, the right hand side of equation (2.4) involves M_d which can be determined in terms of $\underset{\sim}{\pi}'$ and this turns out to be sufficient to enable us to obtain an explicit solution for M. In fact is is well known (Feller [6], Kemeny and Snell [9]) that if $\underset{\sim}{\pi}' = (\pi_1, \pi_2, \cdots, \pi_m)$ is the stationary probability vector then $m_{jj} = 1/\pi_j$ and thus if $\Pi = \underset{\sim\sim}{e}\underset{\sim}{\pi}'$,

$$M_d = (\Pi_d)^{-1}. \tag{2.5}$$

General solutions to equation (2.4) are discussed in section 5.1.

It is also possible to derive expressions for the matrices of higher moments $M^{(k)} = [m_{ij}^{(k)}]$. Since a Markov chain is a special Markov renewal process such problems can be considered in the more general Markov renewal setting.

Let $M_{ij}^{(n)}$ be the number of k ($0 \leq k \leq n$) and $N_{ij}^{(n)}$ be the number of k ($1 \leq k \leq n$) such that $X_k = j$ given $X_0 = i$. Thus $M_{ij}^{(n)} = N_{ij}^{(n)} + \delta_{ij}$ and these random variables record the number of visits the Markov chain makes to state j by trial n with $M_{ij}^{(n)}$ recording the initial occurrence at the zeroth time if $i = j$. There is no standard terminology for distinguishing between these random variables. We shall call $M_{ij}^{(n)}$ the occupancy time random variable and $N_{ij}^{(n)}$ the modified occupancy time random variable.

If $p_{ij}^{(k)}$ is the k step transition probability from state i to state j ($p_{ij}^{(0)} = \delta_{ij}$, $p_{ij}^{(1)} = p_{ij}$) and if $P^{(k)} = [p_{ij}^{(k)}]$ then $P^{(k)} = P^k$, ($k \geq 1$),

($P^{(0)} = I$). With this terminology it is easily seen that

$$(2.6) \qquad E M_{ij}^{(n)} = \sum_{k=0}^{n} p_{ij}^{(k)} \text{ and } E N_{ij}^{(n)} = \sum_{k=1}^{n} p_{ij}^{(k)},$$

and that

$$(2.7) \qquad [E M_{ij}^{(n)}] = \sum_{k=0}^{n} P^k \text{ and } [E N_{ij}^{(n)}] = \sum_{k=1}^{n} P^k.$$

We shall show in section 6.1 that we can utilize generalized inverses of $I - P$ to obtain a variety of expressions for $[E M_{ij}^{(n)}]$ and $[E N_{ij}^{(n)}]$.

2.2. Markov renewal processes and semi-Markov processes

Let $\{(X_n, T_n)\}$ $(n \geq 0)$ be a Markov renewal process with state space $S = \{1, 2, \cdots, m\}$ and semi-Markov kernel $Q(t) = [Q_{ij}(t)]$ where

$$Q_{ij}(t) = P\{X_{n+1} = j,\ T_{n+1} - T_n \leq t | X_n = i\},\ i, j \in S.$$

The terminology used above is that which was introduced by Çinlar [4]. For a descriptive definition see Hunter [8]. One way of visualizing a Markov renewal process is to interpret X_n as the state the stochastic process is in following the nth transition and T_n is the time that the nth transition occurs, $(0 = T_0 \leq T_1 \leq T_2 \leq \cdots)$.

One pertinent observation is that the transitions are governed by a discrete time Markov chain $\{X_n\}$ with transition matrix $P = [p_{ij}] = [Q_{ij}(+\infty)]$. Given that the (n+1)th transition takes the process from state i to state j, the time between transitions is governed by a distribution function

$$F_{ij}(t) = P[T_{n+1} - T_n \leq t | X_n = i,\ X_{n+1} = j].$$

Thus, provided $p_{ij} > 0$, $Q_{ij}(t) = p_{ij} F_{ij}(t)$. (If $p_{ij} = 0$ then $Q_{ij}(t) = 0$ for all t and $F_{ij}(t)$ may be arbitrarily defined, say $F_{ij}(t) = 1$ for $t > 0$.)

Let $H_i(t) = P[T_{n+1} - T_n \leq t | X_n = i]$ be the distribution function of the time until the next transition given that the process has just entered state i, then

$$H_i(t) = \sum_{j=1}^{m} p_{ij} F_{ij}(t) = \sum_{j=1}^{m} Q_{ij}(t).$$

We shall use the following notation:

$$\mu_{ij}^{(k)} = \int_0^\infty t^k dQ_{ij}(t), \quad \eta_{ij}^{(k)} = \int_0^\infty t^k dF_{ij}(t), \quad \mu_i^{(k)} = \int_0^\infty t^k dH_i(t).$$

Thus

$$\mu_{ij}^{(k)} = p_{ij}\eta_{ij}^{(k)} \text{ and } \mu_i^{(k)} = \sum_{j=1}^{m} \mu_{ij}^{(k)}.$$

For convenience we write $\mu_{ij}^{(1)} = \mu_{ij}$, $\eta_{ij}^{(1)} = \eta_{ij}$ and $\mu_i^{(1)} = \mu_i$.

As in section 2.1 we shall assume that the imbedded Markov chain $\{X_n\}$ is irreducible and thus has a stationary probability vector $\underset{\sim}{\pi}'$.

Let T_{ij} be the time for a first passage from state i to state j to take place in the Markov renewal process and let $G_{ij}(t)$ be the distribution function of T_{ij}. If $m_{ij}^{(k)} = \int_0^\infty t^k dG_{ij}(t)$ with $m_{ij}^{(1)} = m_{ij}$ then it can be shown (Lemma 2.1, Hunter [8]) that for $r \geq 1$

$$(2.8) \qquad m_{ij}^{(r)} = \sum_{k \neq j} p_{ik} m_{kj}^{(r)} + \sum_{s=1}^{r-1} \binom{r}{s} \{ \sum_{k \neq j} \mu_{ik}^{(r-s)} m_{kj}^{(s)} \} + \mu_i^{(r)}.$$

so that if $M^{(r)} = [m_{ij}^{(r)}]$ and $P^{(r)} = [\mu_{ij}^{(r)}]$,

$$(2.9)\qquad (I - P)M^{(r)} = \sum_{s=1}^{r-1} \binom{r}{s} P^{(r-s)}[M^{(s)} - M_d^{(s)}] + P^{(r)}E - PM_d^{(r)}.$$

In particular when $r = 1$ if we take $M^{(1)} = M$ we have

$$(2.10)\qquad (I - P)M = P^{(1)}E - PM_d.$$

Observe that when the Markov renewal process $\{(X_n, T_n)\}$ degenerates to a Markov chain $\{X_n\}$ by taking $T_n = n$ for all n, $\mu_{ij} = p_{ij}$ and $\mu_i = 1$. For such a case equation (2.10) becomes equation (2.4). Consequently any general technique to solve equation (2.4) can be applied to the more general set of equation (2.10). Such a procedure is examined in section 5.2.

For each fixed $j \in S$ the instants T_n for which $X_n = j$ form a possibly delayed renewal process. In particular, if we let $N_t = \sup\{n \geq 0 \,|\, T_n \leq t\}$ and $N_j(t)$ = number of n $(0 < n < N_t)$ such that $X_n = j$ then the vector $\{N_1(t), N_2(t), \cdots, N_m(t)\}$ forms the Markov renewal counting process. If we are given that $X_0 = i$ then we can define $N_{ij}(t) = N_j(t)$. Consequently the random variables $N_{ij}(t)$ are analogous to the occupation time random variables in the Markov chain setting.

Let $M_{ij}(t) = E[N_{ij}(t)] = E[N_j(t) \,|\, X_0 = i]$ and $M(t) = [M_{ij}(t)]$. This matrix function is also called the Markov renewal kernel and can be expressed in terms of semi-Markov kernel, (Çinlar [4]). In particular

$$M(t) = [\sum_{n=1}^{\infty} Q_{ij}^{(n)}(t)],$$

where $Q_{ij}^{(n)}(t) = P\{X_n = j,\ T_n \leq t \,|\, X_0 = i\},\ (n \geq 1),$

$$= \sum_{k \in S} \int_0^t Q_{kj}^{(n-1)}(t-y)\, dQ_{ik}(y),\ (n \geq 2),$$

with $Q_{ij}^{(1)}(t) = Q_{ij}(t)$.

In Hunter [8] it was shown that

$$M_{ij}(t) = \frac{t}{m_{jj}} + \left(\frac{m_{jj}^{(2)}}{2m_{jj}} - \frac{m_{ij}}{m_{jj}}\right) + o(1), \tag{2.11}$$

or in terms of matrices

$$M(t) = tE(M_d)^{-1} + \frac{1}{2} E[(M_d)^{-1}]^2 M_d^{(2)} - M(M_d)^{-1} + o(1)E. \tag{2.12}$$

By utilizing general expressions for M and $M_d^{(2)}$ which we deduce in section 5.2, we derive a collection of simple expressions for equation (2.12) in section 6.2.

Associated with a Markov renewal process $\{(X_n,T_n)\}$ is its minimal semi-Markov process $\{Y_t\}$ $(t \geq 0)$ defined as

$$Y_t = X_n \text{ for } T_n \leq t < T_{n+1},$$

(where, since $\{X_n\}$ is assumed to be an irreducible Markov chain, $\sup_n T_n = +\infty$). Thus Y_t may be regarded as the state that the Markov renewal process is in at time t.

If $\{X_n\}$ is irreducible and $\{(X_n,T_n)\}$ is aperiodic, $\mu_i < \infty$ for all $i \in S$, then it can be shown (Çinlar [4]) if $\underset{\sim}{\nu}' = (\nu_1, \nu_2, \cdots, \nu_m)$ where

$$\nu_j = \lim_{t\to\infty} P\{Y_t = j \mid X_0 = i\},$$

then

$$\underset{\sim}{\nu}' = \frac{\underset{\sim}{\pi}'\Lambda}{\underset{\sim}{\pi}'\underset{\sim}{\mu}}, \tag{2.13}$$

where $\underset{\sim}{\pi}'$ is the stationary probability vector of the imbedded Markov chain, $\Lambda = \text{diag}(\mu_1, \mu_2, \cdots, \mu_m)$ and $\underset{\sim}{\mu}' = (\mu_1, \mu_2, \cdots, \mu_m)$. Thus the limiting distribution of the semi-Markov process can be found easily once $\underset{\sim}{\pi}'$ is known.

2.3. Markov chains in continuous time

When the semi-Markov kernel Q(t) has the form

$$Q_{ij}(t) = p_{ij}[1 - e^{-\lambda_i t}],\ i,j = 1,2,\cdots,m,\ t \geq 0; \tag{2.14}$$

with $p_{ii} = 0$ then the semi-Markov process Y_t associated with this kernel is a Markov process, or more specifically a Markov chain in continuous time. We shall also assume that $0 < \lambda_i < \infty$ so that the process is stable and regular.

In the terminology of section 2.2 we have that

$$F_{ij}(t) = 1 - e^{-\lambda_i t},\ i \neq j,\ t \geq 0,$$

implying that for $i,j \in S = \{1,2,\cdots,m\}$,

$$\eta_{ij} = \frac{1}{\lambda_i},\ (i \neq j);\ \mu_{ij} = \frac{p_{ij}}{\lambda_i};\ \mu_i = \frac{1}{\lambda_i}.$$

Although the properties of the process are determined by the semi-Markov kernel it is traditional to use the infinitesimal generator of the process when examining such processes. Çinlar [4] shows that knowledge of Q(t) as given by equations (2.14) is equivalent to knowledge of the infinitesimal generator $Q = [q_{ij}]$. In particular

$$q_{ij} = \begin{cases} -\lambda_i, & i = j, \\ \lambda_i p_{ij}, & i \neq j. \end{cases}$$

Conversely $\lambda_i = -q_{ii}$,

$$\text{and} \quad p_{ij} = \begin{cases} 0, & i = j, \\ -q_{ij}/q_{ii}, & i \neq j. \end{cases}$$

Observe that $q_{ij} \geq 0$ for $i \neq j$, $q_{ii} < 0$ and that $\sum_j q_{ij} = 0$. From these relationships it is easy to see that

(2.15) $$I - P = (Q_d)^{-1}Q \text{ and } \Lambda = -(Q_d)^{-1}.$$

An interesting by-product of these results is that instead of using equation (2.13) to find $\underset{\sim}{\nu}'$ via $\underset{\sim}{\pi}'$, the stationary probability vector of the irreducible imbedded Markov chain with transition matrix P, we can determine $\underset{\sim}{\nu}'$ from the infinitesimal generator of the process by solving the equations

(2.16) $$\underset{\sim}{\nu}'Q = \underset{\sim}{0}' \text{ subject to } \underset{\sim}{\nu}'\underset{\sim}{e} = 1.$$

Because of the singularity of Q, equation (2.15) lends itself to treatment using generalized inverses of Q. This approach is considered in section 4.3.

Expressions for the mean first passage times and the moments of the renewal counting random variables can be obtained from the relevant Markov renewal process results with an appropriate identification of parameters. Further discussion is presented in section 5.3 and 6.3.

3. Generalized Inverses

3.1. Definitions

In section 2 we presented a variety of applied probability situations where we have a set of linear equations involving a singular matrix that we wish to solve. The simplest class of "equation solving generalized inverses" are the one condition generalized inverses which we shall call g-inverses.

Definition 3.1.: A g-inverse of a matrix A is any matrix A^- such that

$$AA^-A = A. \tag{3.1}$$

□

Generalized inverses are in general not unique (unless A is non-singular in which case $A^- = A^{-1}$). By imposing additional conditions we can in fact however end up with a unique generalized inverse.

Definition 3.2.: The Moore-Penrose generalized inverse A^+ of a matrix A is the (unique) matrix satisfying the conditions

$$\text{(i)}\quad AA^+A = A \qquad \text{(ii)}\quad A^+AA^+ = A^+$$
$$\text{(iii)}\quad (AA^+)' = AA^+ \qquad \text{(iv)}\quad (A^+A)' = A^+A.$$

□

We have made no statement concerning the order of the matrices but if A is a rectangular m x n matrix then any g-inverse A^- must be of order n x m so that the matrix multiplications present in Definitions 3.1 and 3.2 are conformable. In the application areas we are concerned with, only square matrices of order m x m will arise and for such matrices there is another class of g-inverses of interest.

Definition 3.3.: If A is a square matrix then the group inverse for A, if it exists, is defined to be the matrix A* which satisfies the conditions

(i) $AA^*A = A$ (ii) $A^*AA^* = A^*$ (iii) $AA^* = A^*A$.

If rank (A) = rank (A^2) then a group inverse exists and it is unique. □

There is an extensive literature on generalized inverses and any reader interested in getting some impression as to the richness of the theory that has been developed should consult one of many books now available on this subject. For example Ben-Israel and Greville [1], Boullion and Odell [2] and Campbell and Meyer [3] (this latter reference contains an excellent chapter on applications of the group inverse to finite Markov chains).

3.2. Solving systems of linear equations

The following theorem and its corollaries provide us with tools for attacking systems of linear equations. In all cases we assume that matrix multiplications are well defined.

Theorem 3.1.: A necessary and sufficient condition for the equation A X B = C to have a solution is that $AA^-CB^-B = C$, (the condition for consistency), where A^- is any g-inverse of A and B^- is any g-inverse of B. If this consistency condition is satisfied the general solution is given by one of the two equivalent forms; either (a):

$$X = A^-CB^- + W - A^-AWBB^-, \tag{3.2}$$

where W is an arbitrary matrix, or (b):

$$X = A^-CB^- + (I - A^-A)U + V(I - BB^-), \tag{3.3}$$

where U and V are arbitrary matrices.

Proof: The necessity and sufficiency of the consistency conditions together

with equation (3.2) can be found in Rao [16] (or Theorem 6.3.2 of Campbell and Meyer [3]). The equivalence of equations (3.2) and (3.3) follows by taking $W = (I - A^{-}A)U + V(I - BB^{-})$, or conversely taking $U = V = \frac{1}{2}(W + A^{-}AW + WBB^{-})$. □

If I is an identity matrix $I^{-} = I$ and from Theorem 3.1 we deduce immediately the following two corollaries.

Corollary 3.1.1.: A necessary and sufficient condition for the equation $AX = C$ to be consistent is that $AA^{-}C = C$, where A^{-} is any g-inverse of A, in which case the general solution is given by

$$X = A^{-}C + (I - A^{-}A)U , \tag{3.4}$$

where U is an arbitrary matrix. □

Corollary 3.1.2.: A necessary and sufficient condition for the equation $XB = C$ to be consistent is that $CB^{-}B = C$, where B^{-} is any g-inverse of A, in which case the general solution is given by

$$X = CB^{-} + V(I - BB^{-}) , \tag{3.5}$$

where V is an arbitrary matrix. □

Since g-inverses are not unique let us write $A\{1\}$ for the set of one condition g-inverses A^{-} of A satisfying equation (3.1). Thus if $X \in A\{1\}$ then $A X A = A$. Suppose A^{-} is a particular g-inverse of A, then since $AA^{-}AA^{-}A = A$ from Theorem 3.1 we can deduce that the equation $A X A = A$ is consistent. The general solution of this equation then provides us with $A\{1\}$ given A^{-}. In fact we have a variety of characterizations of $A\{1\}$ given a particular g-inverse of A.

Theorem 3.2.: If A^- is any g-inverse of A then all g-inverses of A can be characterized as members of the following equivalent sets:

$$(3.6)\quad A\{1\} = \{A^-AA^- + W - A^-AWAA^- \mid W \text{ arbitrary}\},$$

$$(3.7)\quad A\{1\} = \{A^-AA^- + (I - A^-A)U + V(I - AA^-) \mid U, V \text{ arbitrary}\},$$

$$(3.8)\quad A\{1\} = \{A^- + H - A^-AHAA^- \mid H \text{ arbitrary}\},$$

$$(3.9)\quad A\{1\} = \{A^- + (I - A^-A)F + G(I - AA^-) \mid F,G \text{ arbitrary}\}.$$

Proof: The characterizations (3.6) and (3.7) follow from equations (3.2) and (3.3) with $C = B = A$ and the equivalence already established in Theorem 3.1. Furthermore by taking $W \equiv H + A^-$ (3.6) and (3.8) are equivalent as are (3.7) and (3.9) with $U \equiv F + \frac{1}{2}A^-$ and $V \equiv G + \frac{1}{2}A^-$. □

3.3. Generalized inverses of I - P

The following theorem provides us with a key result, hitherto unreported in the literature.

Theorem 3.3.: Let P be the transition matrix of a finite irreducible discrete time Markov chain. Let $\underset{\sim}{u}'$ be any vector such that $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$ and let $\underset{\sim}{t}$ be any non-zero vector. Then

(a) $I - P + \underset{\sim}{t}\underset{\sim}{u}'$ is non-singular.

(b) $[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}$ is a g-inverse of I - P.

Proof: (a) For any matrix X,

$$\det(X + \underset{\sim}{t}\underset{\sim}{u}') = \det(X) + \underset{\sim}{u}'(\text{adj } X)\underset{\sim}{t}.$$

By taking X = I - P, a singular matrix we have

$$\det(I - P + \underset{\sim}{t}\underset{\sim}{u}') = \underset{\sim}{u}'[\mathrm{adj}(I - P)]\underset{\sim}{t}. \tag{3.10}$$

Furthermore $[\mathrm{adj}(I - P)](I - P) = (I - P)[\mathrm{adj}(I - P)] = \det(I - P)I = 0$, so that if $A = \mathrm{adj}(I - P)$ then

$$AP = PA = A. \tag{3.11}$$

If P is irreducible then any matrix A satisfying equation (3.11) is a multiple of $\Pi = \underset{\sim}{e}\underset{\sim}{\pi}'$, where $\underset{\sim}{\pi}'$ is the stationary probability vector of the Markov chain. Thus $\mathrm{adj}(I - P) = k\underset{\sim}{e}\underset{\sim}{\pi}'$.

Now if A is an m x m matrix with eigenvalues $\mu_1, \mu_2, \cdots, \mu_m$ and if μ_1 is a zero eigenvalue then $\mathrm{tr}(\mathrm{adj}\ A) = \prod_{j=2}^{m} \mu_j$. Since P is irreducible its eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$ are such that $\lambda_1 = 1$ is the only eigenvalue equal to 1. Consequently $\mathrm{tr}(\mathrm{adj}\ (I - P)) = \prod_{j=2}^{m} (1 - \lambda_j) \neq 0$. But $\mathrm{tr}(\mathrm{adj}\ (I - P)) = k\ \mathrm{tr}(\underset{\sim}{e}\underset{\sim}{\pi}') = k$ and thus from equation (3.10)

$$\det(I - P + \underset{\sim}{t}\underset{\sim}{u}') = k(\underset{\sim}{u}'\underset{\sim}{e})(\underset{\sim}{\pi}'\underset{\sim}{t}) \neq 0,$$

establishing the required non-singularity.

(b) First observe that

$$(I - P + \underset{\sim}{t}\underset{\sim}{u}')(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1} = I,$$

so that

$$(I - P)(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1} = I - \underset{\sim}{t}\underset{\sim}{u}'(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}. \tag{3.12}$$

Now note, using equation (2.2), that

$$\underset{\sim}{\pi}'(I - P + \underset{\sim}{t}\underset{\sim}{u}') = \underset{\sim}{\pi}'(I - P) + \underset{\sim}{\pi}'\underset{\sim}{t}\underset{\sim}{u}' = (\underset{\sim}{\pi}'\underset{\sim}{t})\underset{\sim}{u}',$$

and thus

$$\frac{\underset{\sim}{\pi}'}{\underset{\sim}{\pi}'\underset{\sim}{t}} = \underset{\sim}{u}'(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}. \tag{3.13}$$

Substitution of the result of equation (3.13) into equation (3.12) gives

$$(I - P)(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1} = I - \frac{\underset{\sim}{t}\underset{\sim}{\pi}'}{\underset{\sim}{\pi}'\underset{\sim}{t}}, \tag{3.14}$$

and hence that

$$(I - P)(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}(I - P) = I - P - \frac{\underset{\sim}{t}\underset{\sim}{\pi}'}{\underset{\sim}{\pi}'\underset{\sim}{t}}(I - P) = I - P,$$

showing that $(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}$ is a g-inverse of $I - P$ by virtue of Definition 3.1. □

Paige, Styan and Wachter [15] developed a similar result in which they showed that $I - P + \underset{\sim}{e}\underset{\sim}{u}'$ is non-singular provided $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$ and that its inverse is a g-inverse of $I - P$. Of course, this is a special case of Theorem 3.3 with $\underset{\sim}{t} = \underset{\sim}{e}(\neq \underset{\sim}{0})$. However it is the generality of Theorem 3.3 that will prove useful in the applications to follow.

Corollary 3.3.1.: Under the conditions of the theorem, any g-inverse of $I - P$ can be expressed by the following equivalent forms:

$$G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \frac{\underset{\sim}{e}\underset{\sim}{u}'H}{\underset{\sim}{u}'\underset{\sim}{e}} + \frac{H\underset{\sim}{t}\underset{\sim}{\pi}'}{\underset{\sim}{\pi}'\underset{\sim}{t}} - \frac{\underset{\sim}{e}\underset{\sim}{u}'H\underset{\sim}{t}\underset{\sim}{\pi}'}{(\underset{\sim}{u}'\underset{\sim}{e})(\underset{\sim}{\pi}'\underset{\sim}{t})}, \tag{3.15}$$

where H is an arbitrary matrix.

$$G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \frac{\underset{\sim}{e}\underset{\sim}{u}'F}{\underset{\sim}{u}'\underset{\sim}{e}} + \frac{G\underset{\sim}{t}\underset{\sim}{\pi}'}{\underset{\sim}{\pi}'\underset{\sim}{t}}, \tag{3.16}$$

where F and G are arbitrary matrices.

Proof: Using an approach similar to that used above

$$(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}(I - P) = I - (I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}\underset{\sim}{t}\underset{\sim}{u}',$$

and

$$(I - P + \underset{\sim}{t}\underset{\sim}{u}')\underset{\sim}{e} = (I - P)\underset{\sim}{e} + \underset{\sim}{t}\underset{\sim}{u}'\underset{\sim}{e} = (\underset{\sim}{u}'\underset{\sim}{e})\underset{\sim}{t},$$

with $(I - P)\underset{\sim}{e} = \underset{\sim}{0}$ due to the stochastic nature of P, implying that

$$\frac{\underset{\sim}{e}}{\underset{\sim}{u}'\underset{\sim}{e}} = (I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}\underset{\sim}{t}, \tag{3.17}$$

and consequently that

$$(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}(I - P) = I - \frac{\underset{\sim}{e}\underset{\sim}{u}'}{\underset{\sim}{u}'\underset{\sim}{e}}. \tag{3.18}$$

If we use the characterizations (3.8) and (3.9) then equations (3.15) and (3.16) follow by taking $A = I - P$, $A^{-} = (I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}$ and using equations (3.14) and (3.18). □

Corollary 3.3.2.: Under the conditions of the theorem any g-inverse of I - P has the characterization

$$G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}', \tag{3.19}$$

for arbitrary $\underset{\sim}{f}$ and $\underset{\sim}{g}$.

Proof: Equations (3.16) and (3.19) are equivalent by taking $F = \underset{\sim}{e}\underset{\sim}{f}'$ and $G = \underset{\sim}{g}\underset{\sim}{\pi}'$ or conversely by taking

$$\underset{\sim}{f} = \frac{\underset{\sim}{u}'F}{\underset{\sim}{u}'\underset{\sim}{e}} \text{ and } \underset{\sim}{g} = \frac{G\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} .$$

Similarly equations (3.15) and (3.19) are equivalent by taking $H = \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'$ or conversely by taking, for example,

$$\underset{\sim}{f}' = \frac{\underset{\sim}{u}'H}{\underset{\sim}{u}'\underset{\sim}{e}} \text{ and } \underset{\sim}{g} = [I - \frac{\underset{\sim}{e}\underset{\sim}{u}'}{\underset{\sim}{u}'\underset{\sim}{e}}] \frac{H\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} . \qquad \square$$

A variety of generalized inverses of $I - P$ have appeared in the literature when their derivation was often obtained by 'ad hoc' techniques. In the following corollary we exhibit their representation in terms of the characterization given by (3.19).

Corollary 3.3.3.: Let P be the transition matrix of a finite irreducible Markov chain with stationary probability vector $\underset{\sim}{\pi}'$ and $\Pi \equiv \underset{\sim}{e}\underset{\sim}{\pi}'$. The following are all generalized inverses of $I - P$.

(a) $G_1 = [I - P + \Pi]^{-1}$.

(b) $G_2 = [I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1}$ provided $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$.

(c) $G_3 = [I - P + \Pi]^{-1} - \Pi$.

(d) $G_4 = [I - P + \alpha\underset{\sim}{\pi}\underset{\sim}{e}']^{-1} - \alpha\Pi$ where $\alpha = (m\underset{\sim}{\pi}'\underset{\sim}{\pi})^{-1/2}$.

(e) $G_5 = [I - P]^{-} + \beta \text{ adj}(I - P)$ for any $(I - P)^{-}$ and $\beta \neq 0$.

(f) If $P = \begin{bmatrix} P_{11} & \underset{\sim}{\alpha} \\ \underset{\sim}{\beta} & p_{mm} \end{bmatrix}$ then $I - P$ has a g-inverse of the form

$$(I - P)^{-} = \begin{bmatrix} (I - P_{11})^{-1} & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix} = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}',$$

where $\underset{\sim}{u}' = (\underset{\sim}{0}', 1)$, $\underset{\sim}{t}' = (\underset{\sim}{0}', 1)$, $\underset{\sim}{f}' = -(\underset{\sim}{\beta}'(I - P_{11})^{-1}, 1)$.

Proof: (a) G_1 follows by taking $\underset{\sim}{t} = \underset{\sim}{e}$, $\underset{\sim}{u} = \underset{\sim}{\pi}$, $\underset{\sim}{f} = \underset{\sim}{0}$, $\underset{\sim}{g} = \underset{\sim}{0}$. G_1 was first recognized as a g-inverse of $I - P$ by Hunter [8] and is known as the "fundamental matrix" of the irreducible transition matrix P (Kemeny and Snell [9]).

(b) As observed earlier, G_2 was identified as a g-inverse of $I - P$ by Paige, Styan and Wachter [15].

(c) G_3 follows by taking $\underset{\sim}{t} = \underset{\sim}{e}$, $\underset{\sim}{u} = \underset{\sim}{\pi}$, $\underset{\sim}{f} = -\underset{\sim}{\pi}$, $\underset{\sim}{g} = \underset{\sim}{0}$, and is, in fact, the group inverse $(I - P)^{\#}$ (see Definition 3.3) as first shown by Meyer [11]; (see also Campbell and Meyer [3]).

(d) G_4 follows by taking $\underset{\sim}{t} = \underset{\sim}{\pi}$, $\underset{\sim}{u} = \alpha\underset{\sim}{e}$, $\underset{\sim}{f} = -\alpha\underset{\sim}{\pi}$, $\underset{\sim}{g} = \underset{\sim}{0}$, and was shown by Paige, Styan and Wachter [15] to be the Moore-Penrose generalized inverse of $I - P$, $(I - P)^{+}$, (see Definition 3.2).

(e) From the proof of the non-singularity of $I - P + \underset{\sim}{t}\underset{\sim}{u}'$ we saw that $\text{adj}(I - P) = k\underset{\sim}{e}\underset{\sim}{\pi}'$ with $k \neq 0$.

(f) $$I - P + \underset{\sim}{t}\underset{\sim}{u}' = \begin{bmatrix} I - P_{11} & -\underset{\sim}{\alpha} \\ -\underset{\sim}{\beta}' & 2 - p_{mm} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $\underset{\sim}{t}' = (\underset{\sim}{0}', 1) \neq \underset{\sim}{0}'$ and $\underset{\sim}{u}'\underset{\sim}{e} = (\underset{\sim}{0}', 1)\underset{\sim}{e} = 1 \neq 0$. Thus $I - P + \underset{\sim}{t}\underset{\sim}{u}'$ is non-singular and from the results concerning inverses of partitioned matrices,

$$[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where $B_{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = 1$,

$$B_{21} = -B_{22}A_{21}A_{11}^{-1} = \underset{\sim}{\beta}'(I - P_{11})^{-1},$$

$$B_{12} = -A_{11}^{-1}A_{12}B_{22} = (I - P_{11})^{-1}\underset{\sim}{\alpha},$$

$$B_{11} = A_{11}^{-1} - A_{11}^{-1}A_{12}B_{21} = (I - P_{11})^{-1} + (I - P_{11})^{-1}\,\underset{\sim}{\alpha}\underset{\sim}{\beta}'(I - P_{11})^{-1}.$$

In determining these expressions we have made use of the fact that $(I - P)\underset{\sim}{e} = \underset{\sim}{0}$ so that $1 - p_{mm} = \underset{\sim}{\beta}'\underset{\sim}{e}$ and $(I - P_{11})\underset{\sim}{e} = \underset{\sim}{\alpha}$ implying that $1 - p_{mm} = \underset{\sim}{\beta}'(I - P_{11})^{-1}\underset{\sim}{\alpha}$. The required form now follows. This result, without the particular representation of the form (3.19), was reported by Rohde [17] as being provided by a personal communication from J. Hearon. (An expression for the group inverse of $I - P$ with P so partitioned has been given by Campbell and Meyer [3]). □

3.4. Generalized inverses of Q

If $Q = [q_{ij}]$ is the infinitesimal generator of the continuous time Markov chain and P is the transition matrix of the imbedded discrete time Markov chain then from equations (2.15) it is easy to see that

$$Q = -\Lambda^{-1}(I - P). \tag{3.20}$$

If P is irreducible we have that $I - P$ has rank $m - 1$ and that P has a simple unit eigenvalue. The equivalence of these two statements follows

since P is stochastic (pg. 133, Marcus and Minc [10]). However, even though Q is not stochastic we can establish analogous results.

Lemma 3.4.: If P is irreducible then for Q given by (3.20)

a) Q has rank m - 1,

b) Q has a simple zero eigenvalue.

Proof: Since I - P has rank m - 1 and Q is derived from I - P by a non-singular transformation, Q also has rank m - 1.

Since det(Q) = 0 it is easy to see that Q has at least one zero eigenvalue but to establish the simplicity of this eigenvalue requires more care. We make use of the result that zero is a simple eigenvalue of Q if and only if det(Q) = 0 and tr(adj Q) $\neq$ 0 (cf., p. 218, Mirsky [13]).

Now from equation (3.20),

$$\text{adj } Q = \text{adj}(I - P)\ \text{adj}(-\Lambda^{-1}),$$

$$= k\underset{\sim}{e}\underset{\sim}{\pi}'(-1)^{m-1}\left(\prod_{i=1}^{m}\lambda_i\right)\Lambda,$$

where, from the proof of Theorem 3.3 $k \neq 0$ and from the specifications in section 2.3, $\prod_{i=1}^{m}\lambda_i \neq 0$ with $\Lambda = \text{diag}(\frac{1}{\lambda_1},\frac{1}{\lambda_2},\cdots,\frac{1}{\lambda_m})$. Thus

$$\text{adj } Q = c\underset{\sim}{e}\underset{\sim}{\pi}'\Lambda \text{ with } c \neq 0, \tag{3.21}$$

implying $\text{tr}(\text{adj } Q) = c\underset{\sim}{\pi}'\Lambda\ \underset{\sim}{e} = c\sum_{i=1}^{m}\frac{\pi_i}{\lambda_i} \neq 0.$ □

Armed with these observations we can now generalize Theorem 3.3.

Theorem 3.5.: Let Q be the infinitesimal generator of a continuous time Markov chain with irreducible imbedded discrete time jump chain. Let $\underset{\sim}{u}'$ be any vector

such that $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$ and let $\underset{\sim}{t}$ be any non-zero vector. Then

(a) $Q + \underset{\sim\sim}{t}\underset{\sim\sim}{u}'$ is non-singular.

(b) $(Q + \underset{\sim\sim}{t}\underset{\sim\sim}{u}')^{-1}$ is a g-inverse of Q.

Proof: (a) Following the proof of Theorem 3.3 (a) we have that

$$\det (Q + \underset{\sim\sim}{t}\underset{\sim\sim}{u}') = \underset{\sim}{u}' \operatorname{adj}(Q)\underset{\sim}{t},$$

$$= c(\underset{\sim}{u}'\underset{\sim}{e})(\underset{\sim}{\pi}'\Lambda\underset{\sim}{t}) \neq 0, \text{ using (3.21).}$$

(b) With minor changes we can use the arguments given in the proof of Theorem 3.3 (b). If I - P is replaced by Q and $\underset{\sim}{\pi}'$ by $\underset{\sim}{\nu}'$ $(= \frac{\underset{\sim}{\pi}'\Lambda}{\underset{\sim}{\pi}'\underset{\sim}{\mu}})$, so that $\underset{\sim}{\nu}'Q = \underset{\sim}{0}'$, then the same procedure gives

$$Q(Q + \underset{\sim\sim}{t}\underset{\sim\sim}{u}')^{-1} Q = Q,$$

and hence the required conclusion. □

Without going into all the details we can work through the relevant corollaries of Theorem 3.3 to conclude that, under the conditions of Theorem 3.5, any g-inverse of Q has the characterization

$$H = (Q + \underset{\sim\sim}{t}\underset{\sim\sim}{u}')^{-1} + \underset{\sim\sim}{e}\underset{\sim\sim}{f}' + \underset{\sim\sim}{g}\underset{\sim\sim}{\nu}'. \tag{3.22}$$

The results of Theorem 3.5 and expression (3.22) are new. To my knowledge the theory generalized inverses has not previously been applied to infinitesimal generators of Markov processes.

4. Stationary Distributions

4.1. Markov chains in discrete time

Using the theory of generalized inverses we can now obtain a general

expression for the stationary probability vector $\underset{\sim}{\pi}'$.

Theorem 4.1.: Let P be the transition matrix of a finite irreducible Markov chain. If $(I - P)^-$ is any g-inverse of $I - P$ and if $A \equiv I - (I - P)(I - P)^-$ then

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{v}'A}{\underset{\sim}{v}'A\underset{\sim}{e}} \quad , \tag{4.1}$$

where $\underset{\sim}{v}'$ is any vector such that $\underset{\sim}{v}'A\underset{\sim}{e} \neq 0$.

Proof. Observe that equation (2.2), $\underset{\sim}{\pi}'(I - P) = \underset{\sim}{0}'$, has the form $XB = C$ where $X = \underset{\sim}{\pi}'$, $B = I - P$, $C = \underset{\sim}{0}'$. Using Corollary 3.1.2 we see that the consistency conditions are obviously satisfied and that equation (3.5) implies that the general solution is given by

$$\underset{\sim}{\pi}' = \underset{\sim}{z}'(I - (I - P)(I - P)^-) \equiv \underset{\sim}{z}'A,$$

where $\underset{\sim}{z}'$ must be chosen so that $\underset{\sim}{\pi}'\underset{\sim}{e} = 1 = \underset{\sim}{z}'A\underset{\sim}{e}$.

Let $\underset{\sim}{v}'$ be any vector such that $\underset{\sim}{v}'A\underset{\sim}{e} \neq 0$ and take $\underset{\sim}{z}' = \underset{\sim}{v}'/\underset{\sim}{v}'A\underset{\sim}{e}$ $(\neq \underset{\sim}{0}')$. Then, for such a choice, $\underset{\sim}{\pi}' = \underset{\sim}{v}'A/\underset{\sim}{v}'A\underset{\sim}{e}$ and $\underset{\sim}{\pi}'\underset{\sim}{e} = 1$. □

There are a variety of ways that we can use Theorem 4.1. Firstly, suppose we are given a computer subroutine for generating g-inverses. Can we use this package? In order that we can use equation (4.1) we have to be sure that we can in fact find a suitable $\underset{\sim}{v}'$. The following corollary establishes the required verification providing us with an affirmative answer to the query.

Corollary 4.1.1.: Under the conditions of the theorem, if $(I - P)^-$ is any g-inverse of $I - P$ then $A\underset{\sim}{e} \neq \underset{\sim}{0}$ and thus we can always find a $\underset{\sim}{v}'$ such that $\underset{\sim}{v}'A\underset{\sim}{e} \neq 0$.

Proof. Since any g-inverse of I - P can be characterized by the form given by equation (3.19) we see that

$$A = I - (I - P)[(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1} + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'],$$

$$= [\frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} - (I - P)\underset{\sim}{g}]\underset{\sim}{\pi}', \text{ using equation (3.14).}$$

Thus, since $\underset{\sim}{\pi}'\underset{\sim}{e} = 1$, $A\underset{\sim}{e} = \frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} - (I - P)\underset{\sim}{g}$. Now suppose that $A\underset{\sim}{e} = \underset{\sim}{0}$. This then implies that

$$\frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} = (I - P)\underset{\sim}{g},$$

so that $1 = \frac{\underset{\sim}{\pi}'\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} = \underset{\sim}{\pi}'(I - P)\underset{\sim}{g} = 0$, a contradiction, and thus $A\underset{\sim}{e} \neq \underset{\sim}{0}$. □

Another way to use Theorem 4.1 is to take a particular form of g-inverse of I - P that does not directly involve knowledge of $\underset{\sim}{\pi}$ and use such a g-inverse in the procedure described by equation (4.1).

Corollary 4.1.2.: If we take

$$(I - P)^{-} = (I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1} + \underset{\sim}{e}\underset{\sim}{f}',$$

where $\underset{\sim}{t} \neq \underset{\sim}{0}$, $\underset{\sim}{u}'$ such that $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$, and $\underset{\sim}{f}$ are arbitrarily chosen then

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{u}'[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}}{\underset{\sim}{u}'[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}\underset{\sim}{e}}. \tag{4.2}$$

Proof: With $(I - P)^{-}$ so chosen

$$A = I - (I - P)\{(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1} + \underset{\sim}{e}\underset{\sim}{f}'\},$$

$$= \underset{\sim}{t}\underset{\sim}{u}'(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}, \text{ using equation (3.12).}$$

Substitution in equation (4.1) yields

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{v}'\underset{\sim}{t}\underset{\sim}{u}'(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}}{\underset{\sim}{v}'\underset{\sim}{t}\underset{\sim}{u}'(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}\underset{\sim}{e}} .$$

Equation (4.2) now follows by dividing the numerator and denominator by $\underset{\sim}{v}'\underset{\sim}{t}$ (since we can obviously find a $\underset{\sim}{v}'$ so that $\underset{\sim}{v}'\underset{\sim}{t} \neq 0$). Note also that by equation (3.13) the denominator $\underset{\sim}{u}'(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}\underset{\sim}{e}$ is non-zero. □

Equation (4.2) is a new result and gives a very general procedure for finding $\underset{\sim}{\pi}'$. Further studies need to be carried out to determine efficient choices of $\underset{\sim}{t}$ and $\underset{\sim}{u}$ to simplify the computation.

The generality of the procedure outlined in Theorem 4.1 is such that all known explicit methods for finding stationary distributions of finite irreducible Markov chains can be expressed in this framework. We show how some of the more well known techniques can be deduced from this theorem.

Corollary 4.1.3.: Let P be the transition matrix of a finite m-state irreducible Markov chain.

(a) For any $\underset{\sim}{u}$ such that $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$,

$$\underset{\sim}{\pi}' = \underset{\sim}{u}'[I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1}. \tag{4.3}$$

(b) Let $(I - P)^-$ be a generalized inverse of $I - P$ such that $\underset{\sim}{e}'[I - (I - P)(I - P)^-]\underset{\sim}{e} \neq 0$, then

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{e}'[I - (I - P)(I - P)^-]}{\underset{\sim}{e}'[I - (I - P)(I - P)^-]\underset{\sim}{e}} . \tag{4.4}$$

(c) Let D_j be the determinant of the matrix formed by removing the jth row and jth column from I - P, then

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{e}'[\text{adj}(I - P)]_d}{\text{tr}(\text{adj}(I - P))},$$

$$= \frac{1}{\left(\sum_{j=1}^{m} D_j\right)} (D_1, D_2, \cdots, D_m). \tag{4.5}$$

(d) The matrix $(I - P)_j = I - P + \underset{\sim}{t}_j \underset{\sim}{e}_j'$ where $\underset{\sim}{e}_j$ is the jth elementary vector and $\underset{\sim}{t}_j = \underset{\sim}{e} - (I - P)\underset{\sim}{e}_j$, which can be formed from I - P by replacing its jth column by $\underset{\sim}{e}$ is non-singular; its inverse is a g-inverse of I - P and

$$\underset{\sim}{\pi}' = \underset{\sim}{e}_j'[(I - P)_j]^{-1}, \quad (j = 1,2,\cdots,m). \tag{4.6}$$

(e) If $P = \begin{bmatrix} P_{11} & \underset{\sim}{\alpha} \\ \underset{\sim}{\beta}' & P_{mm} \end{bmatrix}$ then

$$\underset{\sim}{\pi}' = \frac{(\underset{\sim}{\beta}'(I - P_{11})^{-1}, 1)}{(\underset{\sim}{\beta}'(I - P_{11})^{-1}, 1)\underset{\sim}{e}}. \tag{4.7}$$

Proof:

(a) In equation (4.2) take $\underset{\sim}{t} = \underset{\sim}{e}$ and (4.3) follows by observing that $\underset{\sim}{u}'(I - P + \underset{\sim}{e}\underset{\sim}{u}')^{-1}\underset{\sim}{e} = 1$, which can be established from equation (3.13) by postmultiplying by $\underset{\sim}{t} = \underset{\sim}{e}$. This particular result, with a direct proof, was given earlier by Paige, Styan and Wachter [15].

(b) This follows immediately from equation (4.1) with $\underset{\sim}{v} = \underset{\sim}{e}$. [Decell and Odell [5] derived this result under the additional assumption that $(I - P)(I - P)^-$ should be symmetric. This rather severe restriction establishes the non-zero nature of $k = \underset{\sim}{e}'A\underset{\sim}{e}$. Since, if $\underset{\sim}{\alpha} = A\underset{\sim}{e}$ then $\underset{\sim}{\alpha}' = \underset{\sim}{e}'A$ and it follows upon simplification that $\underset{\sim}{\alpha}'\underset{\sim}{\alpha} = k$. Now $\underset{\sim}{\alpha}'\underset{\sim}{\alpha} > 0$ iff $\underset{\sim}{\alpha} \neq 0$, so that $k = 0$ only if $\underset{\sim}{\alpha} = \underset{\sim}{0}$ but this would imply that $\underset{\sim}{e} = (I - P)(I - P)^-\underset{\sim}{e}$ and hence that $\underset{\sim}{\pi}'\underset{\sim}{e} = \underset{\sim}{\pi}'(I - P)(I - P)^-\underset{\sim}{e} = 0$, a contradiction.]

(c) In Theorem 4.1 take $\underset{\sim}{v}' = \underset{\sim}{e}_1'[\mathrm{adj}(I - P)]$. Then

$$\underset{\sim}{v}'A = \underset{\sim}{e}_1'[\mathrm{adj}(I - P)][I - (I - P)(I - P)^-],$$

$$= \underset{\sim}{e}_1'[\mathrm{adj}(I - P)], \text{ since } [\mathrm{adj}(I - P)](I - P) = 0,$$

$$= \underset{\sim}{v}',$$

so that

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{v}'A}{\underset{\sim}{v}'A\underset{\sim}{e}} = \frac{\underset{\sim}{e}_1'[\mathrm{adj}(I - P)]}{\underset{\sim}{e}_1'[\mathrm{adj}(I - P)]\underset{\sim}{e}} .$$

Now from the proof of Theorem 3.3 we have that $\mathrm{adj}(I - P) = k\Pi = [k\pi_j]$ where $k = \mathrm{tr}(\mathrm{adj}(I - P)) \neq 0$. Also $\mathrm{adj}(I - P) = [\alpha_{ji}]$ where α_{ji} is the cofactor of the (i,j)th element of $I - P$. Thus $\alpha_{ji} = k\pi_j = \alpha_{jj} = D_j$. Consequently,

$$\underset{\sim}{e}_1'[\mathrm{adj}(I - P)] = (\alpha_{11},\alpha_{21},\cdots,\alpha_{m1}) = (\alpha_{11},\alpha_{22},\cdots,\alpha_{mm}),$$

$$= \underset{\sim}{e}'[\mathrm{adj}(I - P)]_d = (D_1,D_2,\cdots,D_m),$$

and

$$\underset{\sim}{e}_1'[\mathrm{adj}(I - P)]\underset{\sim}{e} = \sum_{j=1}^{m} \alpha_{jj} = \mathrm{tr}(\mathrm{adj}(I - P)),$$

$$= \sum_{j=1}^{m} D_j \quad (= k \sum_{j=1}^{m} \pi_j = k \neq 0),$$

and equation (4.5) follows.

This representation of $\underset{\sim}{\pi}'$ in terms of D_j was originally due to Mihoc; (cf. Frechet [7], Hunter [8]).

(d) The elements of the jth column of a matrix can be reduced to zero by port multiplication by $I - \underset{\sim}{e}_j\underset{\sim}{e}_j'$ and thus

$$(I - P)_j = (I - \underset{\sim}{e}_j\underset{\sim}{e}_j') + \underset{\sim}{e}\underset{\sim}{e}_j',$$

$$= I - P + \underset{\sim}{t}_j\underset{\sim}{e}_j' .$$

Observe that $\underset{\sim}{t} \neq \underset{\sim}{0}$ and that $\underset{\sim}{e}_j'\underset{\sim}{e} = 1$ which, form Theorem 3.3 establishes the non-singularity and g-inverse properties of $(I - P)_j$. Expression (4.6) follows from equation (4.2) upon noting that equation (3.13) implies

$$\underset{\sim}{e}_j'[I - P + (P\underset{\sim}{e}_j - \underset{\sim}{e}_j + \underset{\sim}{e})\underset{\sim}{e}_j']^{-1}\underset{\sim}{e} = \underset{\sim}{\pi}'\underset{\sim}{e}/\underset{\sim}{\pi}'(P\underset{\sim}{e}_j - \underset{\sim}{e}_j + \underset{\sim}{e}) = 1.$$

This particular procedure was also suggested by Paige, Styan and Wachter [15].

(e) With $(I - P)^-$ as given by Corollary 3.3.3(f)

$$A = I - (I - P)(I - P)^- = \begin{bmatrix} 0 & \underset{\sim}{0} \\ \underset{\sim}{\beta}'(I - P_{11})^{-1} & 1 \end{bmatrix} .$$

Thus using equation (4.1) with $\underset{\sim}{y} = \underset{\sim}{e}$ we obtain

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{e}'A}{\underset{\sim}{e}'A\underset{\sim}{e}} = \frac{(\underset{\sim}{\beta}'(I - P_{11})^{-1}, 1)}{(\underset{\sim}{\beta}'(I - P_{11})^{-1}, 1)\underset{\sim}{e}} .$$

This result was reported by Rohde [17]. An alternative proof of this result via the group inverse $(I - P)^{\#}$ is given by Meyer [11].

An alternative derivation can also be given by using equation (4.6) with $j = m$, since with P so partitioned

$$(I - P)_m = \begin{bmatrix} I - P_{11} & \underset{\sim}{e} \\ -\underset{\sim}{\beta}' & 1 \end{bmatrix} , \text{ and (cf. proof of Corollary 3.3.3(f)),}$$

$$[(I - P)_m]^{-1} = \begin{bmatrix} (I - P_{11})^{-1} - \Delta(I - P_{11})^{-1}\underset{\sim}{e}\underset{\sim}{\beta}'(I - P_{11})^{-1} & -\Delta(I - P_{11})^{-1}\underset{\sim}{e} \\ \Delta\underset{\sim}{\beta}'(I - P_{11})^{-1} & \Delta \end{bmatrix}$$

where $\Delta = (1 + \underset{\sim}{\beta}'(I - P_{11})^{-1}\underset{\sim}{e})^{-1}$. Equation (4.7) now follows from equation (4.6). This technique was used by Meyer [12] and also by Snell [18]. □

If expression (4.2) is used as the basis for a procedure to determine $\underset{\sim}{\pi}'$ then there is still considerable flexibility in the choice of $\underset{\sim}{u}'$ and $\underset{\sim}{t}$. Paige, Styan and Wachter [15] carried out an error analysis and computational comparison between a variety of algorithms that included the techniques specified by equations (4.3) (4.4) and (4.6) together with rank reduction, limits of matrix powers and least squares procedures. Their study concluded with a recommendation for the method given by equation (4.3) with $\underset{\sim}{u}' = \underset{\sim}{e}_j'P$ ($j = m$ for convenience) using Gaussian elimination with pivoting to solve the equation $\underset{\sim}{\pi}'(I - P + \underset{\sim}{e}\underset{\sim}{u}') = \underset{\sim}{u}'$. This procedure gave the fastest computing times and the smallest average

residual errors.

4.2. Semi-Markov processes

From equation (2.13) we see that, under certain conditions, the limiting distribution of a semi-Markov process can be expressed in terms of $\underset{\sim}{\pi}'$, the stationary probability vector of the underlying Markov chain. Obviously we can use any of the techniques of Theorem 4.1 and its corollaries to find $\underset{\sim}{\pi}'$ and then compute $\underset{\sim}{\nu}'$. However are there any procedures that will give us some simple structure to the algorithm? The following theorem addresses itself to these questions.

Theorem 4.2.: Let $\{(X_n,T_n)\}$ be an aperiodic finite Markov renewal process with the Markov chain $\{X_n\}$ having an irreducible transition matrix P. Let $\{\mu_i\}$ be the mean holding times in the states with $\Lambda = \text{diag}(\mu_1,\mu_2,\cdots,\mu_m)$ and $\underset{\sim}{\mu} = \Lambda\underset{\sim}{e}$. Let $\underset{\sim}{\nu}' = (\nu_1,\nu_2,\cdots,\nu_m)$ be the vector of limiting probabilities for the minimal semi-Markov process then for any $\underset{\sim}{t} \neq \underset{\sim}{0}$ and $\underset{\sim}{u}$ such that $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$

$$\underset{\sim}{\nu}' = \frac{\underset{\sim}{u}'[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}\Lambda}{\underset{\sim}{u}'[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}\underset{\sim}{\mu}} . \tag{4.8}$$

In particular,

$$\underset{\sim}{\nu}' = \underset{\sim}{u}'[I - P + \underset{\sim}{\mu}\underset{\sim}{u}']^{-1}\Lambda . \tag{4.9}$$

Proof: Equation (4.8) follows from equation (2.13) and equation (4.2), with cancellation from numerator and denominator by $\underset{\sim}{u}'[I + P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}\underset{\sim}{e}$.

Equation (4.9) now follows by taking advantage of equation (3.13) with $\underset{\sim}{t} = \underset{\sim}{\mu}$. □

Observe that $[I - P - \underset{\sim}{\mu}\underset{\sim}{u}']^{-1}$ is a g-inverse of $I - P$ which reduces to $[I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1}$ when the Markov renewal process degenerates to a Markov chain with $\underset{\sim}{\mu} = \underset{\sim}{e}$. Thus equation (4.9) is the analog of equation (4.3) in this more general context.

Note also that there is a one to one correspondence between $\underset{\sim}{\nu}'$ and $\underset{\sim}{\pi}'$ so that in any application knowledge of either suffices. It is easily seen that

$$\underset{\sim}{\nu}' = \frac{\underset{\sim}{\pi}'\Lambda}{\underset{\sim}{\pi}'\Lambda\underset{\sim}{e}} \quad \text{is equivalent to} \quad \underset{\sim}{\pi}' = \frac{\underset{\sim}{\nu}'\Lambda^{-1}}{\underset{\sim}{\nu}'\Lambda^{-1}\underset{\sim}{e}} \ .$$

4.3. Markov chains in continuous time

Since Markov chains in continuous time are special semi-Markov processes we can use the results of the previous section with $\Lambda = -Q_d^{-1}$ and $\underset{\sim}{\mu} = -Q_d^{-1}\underset{\sim}{e}$. Furthermore, instead of expressing our algorithm in terms of P, the transition matrix of the imbedded jump chain we can give alternative expressions using Q, the infinitesimal generator of the process. This leads immediately to the following theorem.

Theorem 4.3.: Let $\underset{\sim}{t} \neq \underset{\sim}{0}$ and $\underset{\sim}{u}$ be vectors, with $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$ then the stationary probability vector $\underset{\sim}{\nu}'$ for the Markov chain in continuous time with imbedded jump chain transition matrix P can be expressed by

$$\underset{\sim}{\nu}' = \frac{\underset{\sim}{u}'[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}Q_d^{-1}}{\underset{\sim}{u}'[I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}Q_d^{-1}\underset{\sim}{e}} \ . \tag{4.10}$$

In particular,

(4.11) $$\underset{\sim}{\nu}' = \underset{\sim}{u}'[I - P + Q_d^{-1}\underset{\sim}{e}\underset{\sim}{u}']^{-1}Q_d^{-1} .$$

Alternatively, $\underset{\sim}{\nu}'$ can be expressed in terms of the infinitesimal generator as

(4.12) $$\underset{\sim}{\nu}' = \frac{\underset{\sim}{u}'[Q + \underset{\sim}{t}\underset{\sim}{u}']^{-1}}{\underset{\sim}{u}'[Q + \underset{\sim}{t}\underset{\sim}{u}']^{-1}\underset{\sim}{e}} .$$

In particular,

(4.13) $$\underset{\sim}{\nu}' = \underset{\sim}{u}'[Q + \underset{\sim}{e}\underset{\sim}{u}']^{-1}.$$

Proof: Equations (4.10) and (4.11) follow from equations (4.8) and (4.9) after substitution for Λ and $\underset{\sim}{\mu}$ and replacing both $\underset{\sim}{t}$ and $\underset{\sim}{u}$ by $-\underset{\sim}{t}$ and $-\underset{\sim}{u}$ respectively.

Equations (4.12) and (4.13) now follow from equation (4.10) and (4.11) with $I - P = Q_d^{-1}Q$ and $\underset{\sim}{t}$ replaced by $Q_d^{-1}\underset{\sim}{t} \neq 0$. □

An interesting observation is that equations (4.12) and (4.13) can also be obtained directly by solving the equations $\underset{\sim}{\nu}'Q = \underset{\sim}{0}$ subject to $\underset{\sim}{\nu}'\underset{\sim}{e} = 1$ using the g-inverse approach to solving systems of linear equations as outlined in section 4.1 but using the g-inverse of Q as found in section 3.4.

Although we won't restate the equivalent theorems in this section, every result derived in section 4.1 follows for Markov chains in continuous time but with $I - P$ replaced by Q and $\underset{\sim}{\pi}'$ replaced by $\underset{\sim}{\nu}'$. This, of course, is a direct consequence of the characterization of g-inverses for Q as given by equation (3.20).

5. Mean First Passage Time Matrices

5.1. Markov chains in discrete time

Let $M = [\mu_{ij}]$ be the mean first passage matrix for a Markov chain with irreducible transition matrix P and stationary probability vector $\underset{\sim}{\pi}'$. Let $\Pi = \underset{\sim}{e}\underset{\sim}{\pi}'$. Then from equation (2.4)

$$(I - P)M = E - PM_d ,$$

where, from equation (2.5) $M_d = (\Pi_d)^{-1}$.

Observe that equation (2.4) is of the form $AX = C$ where $A = I - P$, $X = M$ and $C = E - P(\Pi_d)^{-1}$. The general procedure for solving this system of linear equations is given by Corollary 3.1.1 and by taking a general form for the g-inverse of $I - P$ we are led to the following result.

Theorem 5.1.: If G is any g-inverse of $I - P$, then

$$M = [G\Pi - E(G\Pi)_d + I - G + EG_d]D , \tag{5.1}$$

where $D = (\Pi_d)^{-1}$.

Proof: By equation (3.4) the general solution to equation (2.4) is given by

$$M = G(E - PD) + \{I - G(I - P)\}U , \tag{5.2}$$

where U is an arbitrary matrix, provided the consistency condition

$$[I - (I - P)G](E - PD) = 0 ,$$

is satisfied. With G taken in the general form as given by equation (3.19), i.e., for suitable $\underset{\sim}{t}$, $\underset{\sim}{u}$, $\underset{\sim}{f}$ and $\underset{\sim}{g}$

$$G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}',$$

we have, using equation (3.14) that

$$[I - (I - P)G](E - PD) = [\frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} - (I - P)\underset{\sim}{g}]\underset{\sim}{\pi}'[\underset{\sim}{e}\underset{\sim}{e}' - PD]$$

$$= [\frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} - (I - P)\underset{\sim}{g}][\underset{\sim}{e}' - \underset{\sim}{\pi}'D]$$

$$= 0 \text{ since } \underset{\sim}{\pi}'D = \underset{\sim}{e}',$$

showing that equations (2.4) are in fact consistent. Now, using equation (3.18),

$$[I - G(I - P)]U = \underset{\sim}{e}[\frac{\underset{\sim}{u}'}{\underset{\sim}{u}'\underset{\sim}{e}} - \underset{\sim}{f}'(I - P)]U,$$

$$\equiv \underset{\sim}{e}\underset{\sim}{h}'U \equiv \underset{\sim}{e}\underset{\sim}{b}', \text{ say.}$$

Thus equation (5.2) can be written as

$$M = G(E - PD) + \underset{\sim}{e}\underset{\sim}{b}'. \tag{5.3}$$

Observe that the m^2 arbitrary elements of U have been reduced to only m, the elements of $\underset{\sim}{b}'$. These can be determined explicitly due to the restriction that $M_d = (\Pi_d)^{-1}$. Suppose $\underset{\sim}{b}' = (b_1, b_2, \cdots, b_m)$ and let $B = \text{diag}(b_1, b_2, \cdots, b_m)$. From equation (5.3) forming the matrices of diagonal elements gives

$$D = (G\Pi)_d D - (GP)_d D + B,$$

implying that

$$B = [I - (G\Pi)_d + (GP)_d]D. \tag{5.4}$$

Since $\underset{\sim}{e}\underset{\sim}{b}' = EB$ and $E = \Pi D$ substitution of equation (5.4) into equation (5.3) yields

$$M = [G\Pi - E(G\Pi)_d - GP + E(GP)_d + E]D. \tag{5.5}$$

Further simplification of equation (5.5) is possible since we have shown above that

$$I - G + GP = \underset{\sim}{e}\underset{\sim}{h}',$$

and thus

$$E(I - G + GP)_d = \underset{\sim}{e}\underset{\sim}{e}'(\underset{\sim}{e}\underset{\sim}{h}')_d = \underset{\sim}{e}\underset{\sim}{h}'.$$

Hence

$$E - EG_d + E(GP)_d = I - G + GP,$$

implying that $E + E(GP)_d - GP = I - G + EG_d$. Substitution into equation (5.5) gives the required result, equation (5.1). □

Theorem 5.1 presents a new result. It has the same desirable property alluded to earlier, following Theorem 4.1, namely that any computer package that generates g-inverses can be used to determine means of the first passage time distributions.

Because we have used an arbitrary g-inverse of $I - P$ in developing Theorem 5.1 we have considerable flexibility in choosing a particular g-inverse. Firstly the form of the solution for M given by equation (5.1) does not depend on the choice of $\underset{\sim}{f}$ and $\underset{\sim}{g}$.

Corollary 5.1.1.: If $G = G_0 + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'$ where

$$G_0 = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} \quad (\underset{\sim}{t} \neq \underset{\sim}{0},\ \underset{\sim}{u}'\underset{\sim}{e} \neq 0),$$

then

$$M = [G_0\Pi - E(G_0\Pi)_d + I - G_0 + E(G_0)_d]D, \tag{5.6}$$

where $D = (\Pi_d)^{-1}$.

Proof: Let $G = G_0 + H$ where $H = \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'$. Then

$$H\Pi = (\underset{\sim}{f}'\underset{\sim}{e})\Pi + \underset{\sim}{g}\underset{\sim}{\pi}' = (\underset{\sim}{f}'\underset{\sim}{e})\Pi + H - \underset{\sim}{e}\underset{\sim}{f}',$$

and

$$E(H\Pi)_d = (\underset{\sim}{f}'\underset{\sim}{e})E\Pi_d + EH_d - E(\underset{\sim}{e}\underset{\sim}{f}')_d = (\underset{\sim}{f}'\underset{\sim}{e})\Pi + EH_d - \underset{\sim}{e}\underset{\sim}{f}'.$$

Consequently $H\Pi - E(H\Pi)_d = H - EH_d$ and the result follows by substitution in equation (5.1). □

The advantage of equation (5.6) is that any computation with a more general g-inverse is effectively the same as that performed by taking a g-inverse of the form $G_0 = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}$.

However by placing some additional restrictions on the form of the g-inverse we can simplify the form of the expression for M.

Corollary 5.1.2.: (a) If $G = [I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}'$ then

$$M = [I - G + EG_d]D. \tag{5.7}$$

(b) If $G = [I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'$ then

$$M = [I - G_0 + E(G_0)_d]D, \tag{5.8}$$

where $G_0 = [I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1}$ and $D = (\Pi_d)^{-1}$.

Proof: (a) With G as specified

$$G\Pi = [I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1}\underset{\sim}{e}\underset{\sim}{\pi}' + \underset{\sim}{e}\underset{\sim}{f}'\underset{\sim}{e}\underset{\sim}{\pi}',$$

$$= \frac{\underset{\sim}{e}\underset{\sim}{\pi}'}{\underset{\sim}{u}'\underset{\sim}{e}} + (\underset{\sim}{f}'\underset{\sim}{e})\underset{\sim}{e}\underset{\sim}{\pi}', \text{ using equation (3.15)},$$

$$= \beta\Pi \text{ where } \beta = (1/\underset{\sim}{u}'\underset{\sim}{e}) + \underset{\sim}{f}'\underset{\sim}{e}.$$

Thus $E(G\Pi)_d = \beta E\Pi_d = \beta\Pi = G\Pi$ and the result follows from equation (5.1).

(b) With G as specified, part (a) with $\underset{\sim}{f} = 0$ implies that $E(G_0\Pi)_d = G_0$ and the result follows from equation (5.2). □

Special cases of Corollary 5.1.2 have appeared in the literature. In particular when $\underset{\sim}{f}' = \underset{\sim}{0}'$ and $\underset{\sim}{u}' = \underset{\sim}{\pi}'$ equation (5.7) becomes

$$M = [I - Z + EZ_d]D,$$

where $Z = [I - P + \Pi]^{-1}$, the fundamental matrix of the irreducible Markov chain, as obtained by Kemeny and Snell [9].

Similarly when $\underset{\sim}{f}' = -\underset{\sim}{\pi}'$ and $\underset{\sim}{u}' = \underset{\sim}{\pi}'$ equation (5.7) becomes

$$M = [I - T + ET_d]D,$$

where $T = [I - P - \Pi]^{-1} - \Pi$, the group inverse of I - P; as obtained by Meyer [11].

There are some computational considerations that need to be taken into account. Both Z and T effectively require the prior determination of $\underset{\sim}{\pi}'$ which of course, requires the use of some generalized inverse of I - P. Why not use the same g-inverse in one of the forms (5.1), (5.7) or (5.8)? In the corollary that follows we make use of some of the g-inverses used in establishing Corollary 4.1.3.

Corollary 5.1.3.:

(a) If $G = [I - P + \underset{\sim}{e}\underset{\sim}{u}']^{-1}$, then

$$M = [I - G + EG_d][\underset{\sim}{e}\underset{\sim}{u}'G)_d]^{-1}. \tag{5.9}$$

(b) If $G = [(I - P)_j]^{-1} \equiv [g_{ij}]$ then,

$$M = [I - G + EG_d]D + (\underset{\sim}{e}_j\underset{\sim}{e}' - \underset{\sim}{e}\underset{\sim}{e}_j'). \tag{5.10}$$

where $D = \text{diag}\left(\frac{1}{g_{j1}}, \frac{1}{g_{j2}}, \cdots, \frac{1}{g_{jm}}\right)$.

(c) If $P = \begin{bmatrix} P_{11} & \alpha \\ \underset{\sim}{\beta}' & p_{mm} \end{bmatrix}$,

$$\underset{\sim}{a} \equiv (I - P_{11})^{-1}\underset{\sim}{e},\ \underset{\sim}{b}' \equiv \underset{\sim}{\beta}'(I - P_{11})^{-1},\ \lambda \equiv 1 + \underset{\sim}{b}'\underset{\sim}{e},$$

then

$$M = \begin{bmatrix} [\underset{\sim}{a}\underset{\sim}{b}' - E(\underset{\sim}{a}\underset{\sim}{b}')_d + \lambda\{I - (I - P_{11})^{-1} + E((I - P_{11})^{-1})_d\}]((\underset{\sim}{e}\underset{\sim}{b}')_d)^{-1} & \underset{\sim}{a} \\ [\lambda\underset{\sim}{e}'((I - P_{11})^{-1})_d - \underset{\sim}{e}'(\underset{\sim}{a}\underset{\sim}{b}')_d]((\underset{\sim}{e}\underset{\sim}{b}')_d)^{-1} & \lambda \end{bmatrix}. \tag{5.11}$$

Proof:

(a) Equation (4.3) gives $\underset{\sim}{\pi}' = \underset{\sim}{u}'G$ and the result follows from equation (5.8) by noting that $\Pi_d = (\underset{\sim}{e}\underset{\sim}{\pi}')_d = (\underset{\sim}{e}\underset{\sim}{u}'G)_d$.

(b) From Corollary 4.1.3 we see that $G = [I - P + \underset{\sim}{t}_j\underset{\sim}{e}_j']^{-1}$ where $\underset{\sim}{t}_j = \underset{\sim}{e} - (I - P)\underset{\sim}{e}_j$. Furthermore, from equation (4.6), $\underset{\sim}{\pi}' = \underset{\sim}{e}_j'G = (g_{j1}, g_{j2}, \cdots, g_{jm})$ so that the expression for D follows.

Since $\underset{\sim}{t}_j \neq \underset{\sim}{e}$ we use the general expression for M as given by equation (5.1). First observe, however, that $G^{-1}\underset{\sim}{e}_j = (I - P + \underset{\sim}{t}_j\underset{\sim}{e}_j')\underset{\sim}{e}_j = (I - P)\underset{\sim}{e}_j + \underset{\sim}{e} - (I - P)\underset{\sim}{e}_j = \underset{\sim}{e}$ so that $G\underset{\sim}{e} = \underset{\sim}{e}_j$ and hence that $G\Pi = G\underset{\sim}{e}\underset{\sim}{\pi}' = \underset{\sim}{e}_j\underset{\sim}{\pi}_j'$.

Consequently,

$$[G\Pi - E(G\Pi)_d]D = [\underset{\sim}{e}_j\underset{\sim}{\pi}' - \underset{\sim}{e}\underset{\sim}{e}'(\underset{\sim}{e}_j\underset{\sim}{\pi}')_d]D,$$

$$= \underset{\sim}{e}_j\underset{\sim}{\pi}'D - \underset{\sim}{e}\underset{\sim}{e}_j'\underset{\sim}{\pi}_jD,$$

$$= \underset{\sim}{e}_j\underset{\sim}{e}' - \underset{\sim}{e}\underset{\sim}{e}_j',$$

and equation (5.10) follows from equation (5.1).

(c) If, in equation (5.1), we take $G = \begin{bmatrix} (I - P_{11})^{-1} & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix}$,

and $\underset{\sim}{\pi}' = \frac{1}{\lambda}(\underset{\sim}{b}', 1)$, as given by equation (4.7) and we carry out the requisite substitution, expression (5.11) is obtained. □

The procedure suggested by equation (5.9) has the advantage of simplicity but it does involve the inversion of an m x m matrix whereas the procedures suggested by equations (5.10) and (5.11) require the computation of an (m - 1) x (m - 1) matrix inverse. The expression given by equation (5.10) has appeared in literature for the case j = m (Meyer [12]).

5.2. Semi-Markov processes

In section 2.2 we saw that M, the matrix of mean first passage times in a semi-Markov process satisfies equation (2.10), i.e.,

$$(I - P)M = P^{(1)}E - PD,$$

where $P^{(1)} = [\mu_{ij}]$ and $D = M_d$.

This is a generalization of the relevant Markov chain result and can be solved in an analogous manner.

Theorem 5.2.: If G is any g-inverse of I - P then

$$M = [\frac{1}{\lambda_1} GP^{(1)}\Pi - \frac{1}{\lambda_1} E(GP^{(1)}\Pi)_d + I - G + EG_d]D, \tag{5.12}$$

where $\lambda_1 = \underset{\sim}{\pi}'\underset{\sim}{\mu}$, $\underset{\sim}{\mu} = P^{(1)}\underset{\sim}{e}$, $\Pi = \underset{\sim}{e}\underset{\sim}{\pi}'$, $D = \lambda_1(\Pi_d)^{-1}$.

Proof: First note that premultiplication of equation (2.10) by $\underset{\sim}{\pi}'$ yields $\underset{\sim}{\pi}'P^{(1)}E = \underset{\sim}{\pi}'D$, and further $\underset{\sim}{\pi}'P^{(1)}E = \underset{\sim}{\pi}'\underset{\sim}{\mu}\,\underset{\sim}{e}' = \lambda_1\underset{\sim}{e}'$.

Now if we follow through the proof of Theorem 5.1 the consistency conditions are seen to be satisfied:

$$[I - (I - P)G](P^{(1)}E - PD) = [\frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} - (I - P)\underset{\sim}{g}][\underset{\sim}{\pi}'P^{(1)}E - \underset{\sim}{\pi}'D] = 0.$$

The general solution is then given by

$$\begin{aligned} M &= G(P^{(1)}E - PD) + [I - G(I - P)]U, \\ &= G(P^{(1)}E - PD) + \underset{\sim}{e}\underset{\sim}{b}', \end{aligned} \tag{5.13}$$

using equation (5.3). Replacing $\underset{\sim}{e}\underset{\sim}{b}'$ by EB, as in the proof of Theorem 5.1, and taking diagonal elements of equation (5.13) yields

$$D = (GP^{(1)}E)_d - (GP)_dD + B. \tag{5.14}$$

Since $GP^{(1)}E = \frac{1}{\lambda_1} GP^{(1)}\underset{\sim}{e}\underset{\sim}{\pi}'D = \frac{1}{\lambda_1} GP^{(1)}\Pi D$ substitution of the expression for B obtained from equation (5.14) into equation (5.13) yields

$$M = [\frac{1}{\lambda_1} GP^{(1)}\Pi - \frac{1}{\lambda_1} E(GP^{(1)}\Pi)_d - GP + E(GP)_d + E]D .$$

Further simplification, as in the proof of Theorem 5.1, yields the stated expression. □

An immediate observation is that if $P^{(1)} = P$ and hence $\underset{\sim}{\mu} = \underset{\sim}{e}$ and $\lambda_1 = 1$, as is the case for a discrete time Markov chain, equation (5.12) becomes equation (5.1).

Theorem 5.2 presents a new result. A special case when $G = Z = [I - P + \Pi]^{-1}$ appears in Hunter [8].

As was the case in section 5.1 it is of interest to know whether any simplification of expression (5.12) occurs when G has some additional structure.

Corollary 5.2.1.: If $G = [I - P + \underset{\sim}{\mu}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}'$ (with $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$) then

$$M = [I - G + EG_d]D. \tag{5.15}$$

Proof:

$$GP^{(1)}\Pi = G\underset{\sim}{\mu}\underset{\sim}{\pi}' = [I - P + \underset{\sim}{\mu}\underset{\sim}{u}']^{-1}\underset{\sim}{\mu}\underset{\sim}{\pi}' + \underset{\sim}{e}\underset{\sim}{f}'\underset{\sim}{\mu}\underset{\sim}{\pi}',$$

$$= [\frac{1}{\underset{\sim}{u}'\underset{\sim}{e}} + \underset{\sim}{f}'\underset{\sim}{\mu}]\Pi, \text{ using equation (3.17)},$$

$$= \beta\Pi, \text{ say.}$$

Thence $E(GP^{(1)}\Pi)_d = GP^{(1)}\Pi$, as in the proof of Corollary 5.1.2. □

An interesting observation is that this same form of g-inverse for I - P, with $\underset{\sim}{f}' = \underset{\sim}{0}'$, appeared in Theorem 4.2 when the limiting distribution of the semi-Markov process was considered. Further simplication occurs for D in this case and all properties of interest appear in a simple form.

Corollary 5.2.2.: Let $G = [I - P + \underset{\sim}{\mu}\underset{\sim}{u}']^{-1}$ (with $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$), then

$$\lambda_1 = (\underset{\sim}{u}'G\underset{\sim}{e})^{-1},\ \underset{\sim}{\pi}' = \lambda_1\underset{\sim}{u}'G,\ \underset{\sim}{\nu}' = \underset{\sim}{u}'G(\underset{\sim}{e}\underset{\sim}{u}')_d,$$

$$M = [I - G + EG_d][\underset{\sim}{e}\underset{\sim}{u}'G)_d]^{-1}. \tag{5.16}$$

Proof: Since $\underset{\sim}{\pi}' = \underset{\sim}{u}'G/\underset{\sim}{u}'G\underset{\sim}{e}$, (equation (4.21)) and $\underset{\sim}{u}'G\underset{\sim}{\mu} = 1$, (equation (3.17)) the results follow easily. □

The techniques used in this section can be extended to obtain expressions for $M^{(r)} = [m_{ij}^{(r)}]$ for $r \geq 1$. This problem was examined by Hunter [8] and explicit expressions were obtained for $M_d^{(2)}$ and $M^{(2)}$ in terms of Z, the fundamental matrix. Actually the methods used in this aforementioned paper extend naturally to give the following result for $M_d^{(2)}$. The proof follows as for Theorem 2.17, Hunter [8] but with Z replaced by G.

Theorem 5.3.: If G is any g-inverse of I - P then

$$M_d^{(2)} = \frac{1}{\lambda_1}[2\{\frac{1}{\lambda_1}(\Pi P^{(1)}GP^{(1)}\Pi)_d - (GP^{(1)}\Pi)_d - (\Pi P^{(1)}G)_d + \lambda_1 G_d\}D + \lambda_2 I]D, \tag{5.17}$$

where $\lambda_1 = \underset{\sim}{\pi}'P^{(1)}\underset{\sim}{e} = \underset{\sim}{\pi}'\underset{\sim}{\mu}$ and $\lambda_2 = \underset{\sim}{\pi}'P^{(2)}\underset{\sim}{e} = \underset{\sim}{\pi}'\underset{\sim}{\mu}^{(2)}$. □

Simplification of expression (5.17) can be effected by taking special forms for G. Firstly we use the form considered in Corollary 5.2.1.

Corollary 5.3.1.: If $G = [I - P + \underset{\sim}{\mu}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}'$ (with $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$) then

$$M_d^{(2)} = \frac{1}{\lambda_1}[2\{\lambda_1 G_d - (\Pi P^{(1)}G)_d\}D + \lambda_2 I]D. \tag{5.18}$$

Proof: From the proof of Corollary 5.2.1, $GP^{(1)}\Pi = \beta\Pi$ so that $\Pi P^{(1)}GP^{(1)}\Pi = \beta\underset{\sim}{e}\underset{\sim}{\pi}'P^{(1)}\underset{\sim}{e}\underset{\sim}{\pi}' = \beta\lambda_1\Pi$. Consequently the first two terms of

expression (5.17) cancel. □

Further simplification can be effected by taking a specific choice of $\underset{\sim}{u}$.

Corollary 5.3.2.: If $G = [I - P + \underset{\sim}{\mu}\underset{\sim}{\pi}'P^{(1)}]^{-1}$ then

$$M_d^{(2)} = [2G_dD + \frac{(\lambda_2 - 2)}{\lambda_1} I]D. \tag{5.19}$$

Proof: With G so chosen observe that from equation (3.13)

$$\Pi P^{(1)}G = \underset{\sim}{e}\underset{\sim}{\pi}'P^{(1)}[I - P + \underset{\sim}{\mu}\underset{\sim}{\pi}'P^{(1)}]^{-1} = \frac{1}{\lambda_1}\Pi ,$$

and equation (5.19) follows equation (5.18) upon noting $\Pi_d D = \lambda_1 I$. □

The g-inverse of I - P used in Corollary 5.3.2 has not appeared in the literature before. It appears to have the interpretation as the "fundamental matrix of Markov renewal process" since it is a generalization of Kemeny and Snell's Z matrix. When the process degenerates to a Markov chain $P^{(1)} = P$, $\underset{\sim}{\mu} = \underset{\sim}{e}$ and thus $\underset{\sim}{\mu}\underset{\sim}{\pi}'P^{(1)} = \underset{\sim}{e}\underset{\sim}{\pi}'$ and G becomes Z. Also in this case $P^{(2)} = P$, $\underset{\sim}{\mu}^{(2)} = \underset{\sim}{e}$, $\lambda_1 = \lambda_2 = 1$ and equation (5.19) becomes

$$M_d^{(2)} = [2Z_dD - I]D,$$

as obtained for Markov chains by Kemeny and Snell [9].

5.3. Markov chains in continuous time

In this case $\underset{\sim}{\mu} = -Q_d^{-1}\underset{\sim}{e}$ and the g-inverse of I - P used in Corollary 5.2.2 becomes $[I - P - Q_d^{-1}\underset{\sim}{e}\underset{\sim}{u}']^{-1}$. If we replace $\underset{\sim}{u}'$ by $-\underset{\sim}{u}'$ we can take a g-inverse of I - P as

$$G = [I - P + Q_d^{-1}\underset{\sim}{e}\underset{\sim}{u}']^{-1} = [Q + \underset{\sim}{e}\underset{\sim}{u}']^{-1}Q_d ,$$

using equation (2.15). However $[Q + \underset{\sim}{e}\underset{\sim}{u}']^{-1}$ is a g-inverse of the infinitesimal generator Q which we used in Theorem 4.3 to obtain simple results for $\underset{\sim}{\nu}'$. In this case we can express M in terms of properties of Q.

Theorem 5.4.: If $H = [Q + \underset{\sim}{e}\underset{\sim}{u}']^{-1}$ (with $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$) then

$$M = [H - EH_d - Q_d^{-1}][(\underset{\sim}{e}\underset{\sim}{u}'H)_d]^{-1}, \tag{5.20}$$

and $$\underset{\sim}{\nu}' = \underset{\sim}{u}'H, \ \lambda_1 = -(\underset{\sim}{\nu}'Q_d\underset{\sim}{e})^{-1}, \ \underset{\sim}{\pi}' = -\lambda_1\underset{\sim}{\nu}'Q_d.$$

Proof: From Corollary 5.5.1, since $G = HQ_d$

$$\begin{aligned} M &= [I - HQ_d + EH_dQ_d]D, \\ &= [Q_d^{-1} - H + EH_d]Q_dD. \end{aligned} \tag{5.21}$$

Now, from equation (4.13), $\underset{\sim}{\nu}' = \underset{\sim}{u}'H$. This then implies that, since $\Lambda = -Q_d^{-1}$,

$$\underset{\sim}{\pi}' = \frac{\underset{\sim}{\nu}'\Lambda^{-1}}{\underset{\sim}{\nu}'\Lambda^{-1}\underset{\sim}{e}} = \frac{\underset{\sim}{\nu}'Q_d}{\underset{\sim}{\nu}'Q_d\underset{\sim}{e}}.$$

Also, since $\underset{\sim}{\mu} = \Lambda\underset{\sim}{e}$,

$$\lambda_1 = \underset{\sim}{\pi}'\underset{\sim}{\mu} = \frac{\underset{\sim}{\nu}'\Lambda^{-1}\underset{\sim}{\mu}}{-\underset{\sim}{\nu}'Q_d\underset{\sim}{e}} = \frac{\underset{\sim}{\nu}'\underset{\sim}{e}}{-\underset{\sim}{\nu}'Q_d\underset{\sim}{e}} = \frac{1}{-\underset{\sim}{\nu}'Q_d\underset{\sim}{e}}.$$

Consequently, $\underset{\sim}{\pi}' = -\lambda_1\underset{\sim}{\nu}'Q_d$. Furthermore

$$\Pi_d = -\lambda_1(\underset{\sim}{e}\underset{\sim}{\nu}')_dQ_d = -\lambda_1(\underset{\sim}{e}\underset{\sim}{u}'H)_dQ_d,$$

and thus

$$Q_d D = \lambda_1 Q_d (\Pi_d)^{-1} = -(\underset{\sim}{e}\underset{\sim}{u}'H)_d \ , \tag{5.22}$$

and equation (5.20) follows from equations (5.21) and (5.22). □

6. Moments of the Occupation Time Random Variables

6.1. Markov chains in discrete time

In section 2.1 we saw that the behavior of the expectations of the random variables $M_{ij}^{(n)}$ and $N_{ij}^{(n)}$ was intimately connected with sums of power of the transition matrix. The following theorem shows that we can use g-inverses to obtain expressions for such sums.

Theorem 6.1.: If G is any g-inverse of I - P where P is an irreducible transition matrix, then

$$\sum_{r=0}^{n-1} P^r = \begin{cases} n\Pi + (I - \Pi)G(I - P^n), \\ n\Pi + (I - P^n)G(I - \Pi). \end{cases} \tag{6.1}$$

Proof: Let $A_n = \sum_{r=0}^{n-1} P^r$ then

$$(I - P)A_n = I - P^n, \tag{6.2}$$

and

$$A_n(I - P) = I - P^n. \tag{6.3}$$

Equations (6.2) and (6.3) are in a suitable form for applying Corollaries 3.1.1 and 3.1.2, respectively. Both equations are in fact consistent and the arbitrary constant matrix in each solution can be eliminated using the

observation that the A_n are constrained by the relationships

$$A_n\Pi = \Pi A_n = n\Pi. \tag{6.4}$$

With $G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'$ it is easily seen, as in the proof of Theorem 5.1, that

$$I - (I - P)G = [\frac{\underset{\sim}{t}}{\underset{\sim}{\pi}'\underset{\sim}{t}} - (I - P)\underset{\sim}{g}]\underset{\sim}{\pi}' \equiv \underset{\sim}{\alpha}\underset{\sim}{\pi}', \tag{6.5}$$

and

$$I - G(I - P) = \underset{\sim}{e}[\frac{\underset{\sim}{u}'}{\underset{\sim}{u}'\underset{\sim}{e}} - \underset{\sim}{f}'(I - P)] \equiv \underset{\sim}{e}\underset{\sim}{\beta}'. \tag{6.6}$$

We shall consider only equation (6.2). The procedure follows analogously for equation (6.3). The consistency condition is verified by using equation (6.5) while equation (6.6) shows that the general solution is given by

$$\begin{aligned} A_n &= G(I - P^n) + \underset{\sim}{e}\underset{\sim}{\beta}'U_n, \\ &= G(I - P^n) + \underset{\sim}{e}\underset{\sim}{u}_n', \text{ say.} \end{aligned} \tag{6.7}$$

Further, from equation (6.5)

$$n\Pi = \Pi G(I - P^n) + \underset{\sim}{e}\underset{\sim}{u}_n', \text{ (since } \Pi\underset{\sim}{e} = \underset{\sim}{e}). \tag{6.8}$$

Elimination of $\underset{\sim}{e}\underset{\sim}{u}_n'$ between equations (6.7) and (6.8) leads to the first form of equation (6.1) for A_n. □

If we restrict attention to regular Markov chains then $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$ for all $i,j = 1,2,\cdots,m$ and thus, in terms of matrices $\lim_{n\to\infty} P^n = \Pi$. With this observation we obtain the following information concerning the behavior

of $\sum_{r=0}^{n-1} P^r$ for large n.

Corollary 6.1.1.: If G is any g-inverse of I - P where P is the transition matrix of a regular Markov chain then

$$\sum_{r=0}^{n-1} P^r = n\Pi + (I - \Pi)G(I - \Pi) + o(1)E. \tag{6.9}$$

□

Corollary 6.1.2.: If $G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1} + \underset{\sim}{e}\underset{\sim}{f}' + \underset{\sim}{g}\underset{\sim}{\pi}'$ ($\underset{\sim}{t} \neq \underset{\sim}{0}$ and $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$) then if P is the transition matrix of a regular Markov chain

$$\sum_{r=0}^{n-1} P^r = n\Pi + (I - \Pi)(I - P + \underset{\sim}{t}\underset{\sim}{u}')^{-1}(I - \Pi) + o(1)E. \tag{6.10}$$

Proof: Equation (6.10) follows upon the required substitution and observing that $(I - \Pi)\underset{\sim}{e} = \underset{\sim}{0}$ and $\underset{\sim}{\pi}'(I - \Pi) = \underset{\sim}{0}'$.

□

In terms of the expectations of the random variables $M_{ij}^{(n)}$ and $N_{ij}^{(n)}$, the above corollaries, in conjunction with equation (2.7), give immediately:

Theorem 6.2.: For regular Markov chains

$$[EM_{ij}^{(n)}] = (n + 1)\Pi + (I - \Pi)G(I - \Pi) + o(1)E, \tag{6.11}$$

and

$$[EN_{ij}^{(n)}] = (n + 1)\Pi - I + (I - \Pi)G(I - \Pi) + o(1)E. \tag{6.12}$$

□

Although the general expression of Theorem 6.2 are new, if G is replaced by Z, the fundamental matrix, we obtain the results of Kemeny and Snell [9] while if G is replaced by $A^{\#}$, the group inverse of I - P, we obtain the results of Meyer [11]. In these special cases we get simpler forms. In particular we have the following corollary.

Corollary 6.2.1.: For a regular Markov chain

$$[EM_{ij}^{(n)}] = \begin{cases} n\Pi + Z + o(1)E, \\ (n+1)\Pi + A^{\#} + o(1)E. \end{cases} \tag{6.13}$$

Proof: Simplification of equation (6.11) gives the required forms after observing that $(I - \Pi)Z(I - \Pi) = Z - \Pi$ and $(I - \Pi)A^{\#}(I - \Pi) = A^{\#}$. □

6.2. Markov renewal processes

In section 2.2 an asymptotic form for the Markov renewal kernel, M(t) was given by equation (2.12) i.e.,

$$M(t) = tE(M_d)^{-1} + \frac{1}{2} E[(M_d)^{-1}]^2 M_d^{(2)} - M(M_d)^{-1} + o(1)E.$$

By utilizing the results of Theorems 5.2 and 5.3 and their corollaries we are able to obtain expression for M(t) in terms of g-inverses of I - P, where P is the transition matrix of the imbedded irreducible Markov chain.

Theorem 6.3.: If G is any g-inverse of I - P then

$$M(t) = \frac{t}{\lambda_1}\Pi + \frac{\lambda_2}{\lambda_1^2}\Pi + [I - \frac{1}{\lambda_1}\Pi P^{(1)}]G[I - \frac{1}{\lambda_1}P^{(1)}\Pi] - I + o(1)E, \tag{6.14}$$

where $\lambda_1 = \underset{\sim}{\pi}' P^{(1)} \underset{\sim}{e}$, $\lambda_2 = \underset{\sim}{\pi}' P^{(2)} \underset{\sim}{e}$.

Proof: As for the proof of Theorem 3.3 of Hunter [8] but with Z replaced by G. □

Equation (6.14) is identical in form to that obtained earlier with the special g-inverse, Z, the fundamental matrix of the imbedded Markov chain. However the generality of Theorem 6.13 enables us to use special forms of G to obtain simpler structural results.

Corollary 6.3.1.: (a) If $G = [I - P + \underset{\sim}{\mu}\underset{\sim}{u}']^{-1}$ (with $\underset{\sim}{u}'\underset{\sim}{e} \neq 0$) then

$$M(t) = \frac{t}{\lambda_1}\Pi + \frac{\lambda_2}{\lambda_1^2}\Pi + [I - \frac{1}{\lambda_1}\Pi P^{(1)}]G - I + o(1)E. \tag{6.15}$$

(b) If $G = [I - P + \underset{\sim}{t}\underset{\sim}{\pi}'P^{(1)}]^{-1}$ (with $\underset{\sim}{t} \neq \underset{\sim}{0}$) then

$$M(t) = \frac{t}{\lambda_1}\Pi + \frac{\lambda_2\Pi}{\lambda_1^2} + G[I - \frac{1}{\lambda_1}P^{(1)}\Pi] - I + o(1)E. \tag{6.16}$$

(c) If $G = [I - P + \underset{\sim}{\mu}\underset{\sim}{\pi}'P^{(1)}]^{-1}$ then

$$M(t) = \frac{t}{\lambda_1}\Pi + \frac{(\lambda_2 - 1)\Pi}{\lambda_1^2} + G - I + o(1)E. \tag{6.17}$$

Proof: If $G = [I - P + \underset{\sim}{t}\underset{\sim}{u}']^{-1}$ then

if $\underset{\sim}{t} = \underset{\sim}{\mu} = P^{(1)}\underset{\sim}{e}$, equation (3.17) implies $GP^{(1)}\underset{\sim}{e} = \dfrac{\underset{\sim}{e}}{\underset{\sim}{u}'\underset{\sim}{e}}$,

while if $\underset{\sim}{u}' = \underset{\sim}{\pi}'P^{(1)}$, equation (3.13) implies $\underset{\sim}{\pi}'P^{(1)}G = \dfrac{\underset{\sim}{\pi}'}{\underset{\sim}{\pi}'\underset{\sim}{t}}$.

With the appropriate substitution for $\underset{\sim}{t}$ and $\underset{\sim}{\mu}$ in equation (6.14) the stated results are obtained following simplification. □

The expressions given in Theorem 6.3 and its corollaries are new. Computationally equation (6.14) offers much more flexibility than previously thought possible whereas equation (6.17) offers a structurally simple result.

Observe that when the process degenerates to a Markov chain $\lambda_1 = \lambda_2 = 1$, $\underset{\sim}{\mu}\underset{\sim}{\pi}'P^{(1)} = \underset{\sim}{e}\underset{\sim}{\pi}'$ and equation (6.17) reduces immediately to the form

(6.18) $$M(n) = [EN_{ij}^{(n)}] = n\Pi + Z - I + o(1)E\,,$$

which follows, alternatively, from equations (2.7) and (6.13).

We have not examined expressions for the higher moments of Markov renewal counting processes. A discussion concerning the computation of variances is given in Hunter [8].

6.3. Markov chains in continuous time

In this case the results of Theorem 6.3 and its corollary apply with $P^{(1)} = -Q_d^{-1}P$ and $P^{(2)} = 2(Q_d^{-1})^2P$ implying that $\lambda_1 = -\underset{\sim}{\pi}'Q_d^{-1}\underset{\sim}{e} = -(\underset{\sim}{\nu}'Q_d\underset{\sim}{e})^{-1}$ and $\lambda_2 = 2\underset{\sim}{\pi}'(Q_d^{-1})^2\underset{\sim}{e} = -2\lambda_1\underset{\sim}{\nu}'Q_d^{-1}\underset{\sim}{e}$.

Consequently we can derive expressions for M(t) involving the stationary probability vector $\underset{\sim}{\nu}'$ and generalized inverses of Q for irreducible Markov chains in continuous time.

Theorem 6.4.: If H is any g-inverse of Q then

(6.19) $$M(t) = -tLQ_d + 2\tau LQ_d + [I - LP]H[I - L]Q_d - I + o(1)E,$$

where $L = \underset{\sim}{e}\underset{\sim}{\nu}'$ and $\tau = \underset{\sim}{\nu}'Q_d^{-1}\underset{\sim}{e}$

Proof: If H is any g-inverse of Q then, using equation (3.20) and the observations that $I - P = Q_d^{-1}Q$ and $\underset{\sim}{\pi}' = -\lambda_1\underset{\sim}{\nu}'Q_d$, it is easily seen that $H = GQ_d^{-1}$ where G is a g-inverse of I - P.

With $L = \underset{\sim}{e}\underset{\sim}{\nu}'$ it is easily seen that $\Pi = -\lambda_1 LQ_d$, PL = L and thus equation (6.19) follows from equation (6.14). □

Analogous results for the special cases of Corollary 6.3.1 can be examined by taking H as $(Q + \underset{\sim}{t}\underset{\sim}{\nu}'P)^{-1}$, $(Q + \underset{\sim}{e}\underset{\sim}{u}')^{-1}$ and $(Q + LP)^{-1}$.

7. Concluding remarks

In a paper examining the role of group generalized inverses in the theory of finite Markov chains, Meyer [11] stated that the "correct" generalized inverse to use is the group generalized inverse. He also claimed that the Moore-Penrose inverse or any other g-inverse that satisfies a subset of the Penrose conditions of Definition 3.2 are the wrong types since they are "forced" into the theory because of their ability to produce solutions of consistent systems of linear equations. However, as we have exposed in this paper, by taking the class of all g-inverses as our starting point we are able to present a unified approach to a large class of different but related problems for a wide variety of Markovian and semi-Markovian models.

If any one g-inverse stands out it is the matrix $[I - P + \underset{\sim}{\mu}\underset{\sim}{\pi}'P^{(1)}]^{-1}$ as presented for the first time in Corollary 5.3.2 and later in Corollary 6.3.1(c). As we saw, this matrix is a generalization for semi-Markovian processes of the fundamental matrix of irreducible Markov chains as first presented by Kemeny and Snell [9].

Of course, the group inverse has other advantages when one is interested in classifying states of the Markov chains but Campbell and Meyer's claim [3] that some types of g-inverses lead to cumbersome expressions which do little to enhance or unify the theory and provide no practical or computational advantage is disputed. The observation that all known explicit methods for finding stationary distributions can be put in a generalized inverse framework together with the result that <u>any</u> g-inverse can be used in examining the problems presented in this paper is a compelling reason that g-inverses have an important role to play with a wide variety of computer subroutines available for g-inverses,

many included in statistical packages, the applied probabilist now has a large arsenal of techniques available for his use in tackling some of the computational aspects of this study.

References

[1] A. Ben-Israel and T. N. E. Greville, Generalized Inverses: Theory and Applications, Wiley-Interscience, New York, (1974).

[2] T. L. Boullion and P. L. Odell, Generalized Inverse Matrices, Wiley-Interscience, New York, (1971).

[3] S. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, (1979).

[4] E. Çinlar, Introduction to Stochastic Processes, Prentice-Hall, Englewood Cliffs, N.J., (1975).

[5] H. P. Decell, Jr. and P. L. Odell, On the fixed point probability vector of regular or ergodic transition matrices, J. Amer. Statist. Assoc., 62, (1967), pp. 600-602.

[6] W. Feller, An Introduction to Probability Theory and its Applications, 1, Wiley, New York, (1950). (Second edition, 1957; third edition, 1967).

[7] M. Frechet, Recherches Théoriques Modernes sur le Calcul des Probabilités, 2, (second edition), Gauthier-Villars, Paris, (1950).

[8] J. J. Hunter, On the moments of Markov renewal processes, Advances in Appl. Probability, 1, (1969), pp. 188-210.

[9] J. G. Kemeny and J. L. Snell, Finite Markov Chains, Van Nostrand, New York, (1960).

[10] M. Marcus and H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Allyn and Bacon, and Prindle Weber and Schmidt, Boston, (1964).

[11] C. D. Meyer, Jr., The role of the group generalized inverse in the theory of finite Markov chains, SIAM Review, 17, (1975), pp. 443-464.

[12] C. D. Meyer, Jr., An alternative expression for the mean first passage matrix, Linear Algebra and Appl., 22, (1978), pp. 41-47.

[13] L. Mirsky, An Introduction to Linear Algebra, Oxford at the Clarendon Press, London, (1955).

[14] M. F. Neuts, Probability, Allyn and Bacon, Boston, (1973).

[15] C. C. Paige, G. P. H. Styan and P. G. Wachter, Computation of the stationary distribution of a Markov chain, J. Statist. Comput. Simul., 4, (1975), pp. 173-186.

[16] C. R. Rao, Generalized inverse for matrices and its applications in mathematical statistics, Research papers in Statistics, Festschrift for J. Neyman, Wiley, New York, (1966).

[17] C. A. Rohde, Special applications of the theory of generalized matrix inversion to statistics, Proceedings of the Symposium on Theory and Applications of Generalized Inverses of Matrics, (Ed. T. L. Boullion and P. L. Odell), Texas Tech. Press, Lubbock, (1968), pp. 239-266.

[18] J. L. Snell, Introduction to Probability Theory with Computing, Prentice-Hall, Englewood Cliffs, N.J., (1975).