Res. Lett. Inf. Math. Sci., (2000) **1**, 25-36
Available online at http://www.massey.ac.nz/~wwiims/rlims/

A Survey of Generalized Inverses and their Use in Stochastic Modelling

Jeffrey J. Hunter
I.I.M.S., Massey University Albany Campus, Auckland, N.Z.
j.hunter@massey.ac.nz

Abstract
In many stochastic models, in particular Markov chains in discrete or continuous time and Markov renewal processes, a Markov chain is present either directly or indirectly through some form of embedding. The analysis of many problems of interest associated with these models, eg. stationary distributions, moments of first passage time distributions and moments of occupation time random variables, often concerns the solution of a system of linear equations involving I – P, where P is the transition matrix of a finite, irreducible, discrete time Markov chain.

Generalized inverses play an important role in the solution of such singular sets of equations. In this paper we survey the application of generalized inverses to the aforementioned problems. The presentation will include results concerning the analysis of perturbed systems and the characterization of types of generalized inverses associated with Markovian kernels.

1. Introduction

A generalized inverse (g-inverse) of a matrix A is any matrix A^- such that

$$AA^-A = A$$

Such matrices are sometimes called 'one condition' g-inverses or 'equation solving' g-inverses because of their use in solving systems of linear equations. By imposing additional conditions we can obtain various types of multi-condition g-inverses. Consider real conformable matrices X such that:

(1) AXA = A.
(2) XAX = X.
(3) (AX)' = AX.
(4) (XA)' = XA.
(5) For square matrices AX = XA.

Let $A^{(i,j,.l)}$ be any matrix that satisfies conditions (i), (j), …, (*l*) of the above itemised conditions. Such matrices are termed (i, j, …, *l*) g-inverses of A.

Particular special cases include $A^{(1,2)}$, a 'reflexive' g-inverse; $A^{(1,3)}$, a 'least squares' g-inverse; $A^{(1,4)}$, a 'minimum norm' g-inverse; $A^{(1,2,3,4)}$, the 'Moore-Penrose' g-inverse; and $A^{(1,2,5)}$, the 'group inverse', which exists and is unique if rank (A) = rank (A^2). Note that, except for $A^{(1,2,3,4,)}$ and $A^{(1,2,5)}$, g-inverses are, in general, not unique.

2. Solving Systems of Linear Equations

Generalized inverses play a major role in solving systems of linear equations.

A necessary and sufficient condition for AXB = C *to have a solution is that* $AA^-CB^-B = C$.
If this consistency condition is satisfied the general solution is given by
$X = A^-CB^- + W – A^-AWBB^-$, *where W is an arbitrary matrix.*

Special cases of interest are the following:

(i) XB = C has a solution $X = CB^- + W(I – BB^-)$, where W is an arbitrary matrix, provided $C B^- B = C$.

(ii) $AX = C$ has a solution $X = A^- C + (I - A^- A)W$, where W is arbitrary, provided $AA^- C = C$.

(iii) $AXA = A$ has a solution $X = A^- A A^- + W - A^- AWAA^-$, where W is arbitrary.

Result (iii) provides a characterization of A{1}, the set of all g-inverses of A given any one g-inverse, A^-, ([4]): $A\{1\} = \{A^- + H - A^- AHAA^-, H \text{ arbitrary}\}$, or in equivalent form, $A\{1\} = \{A^- + (I - A^- A)F + G(I - AA^-), F, G, \text{arbitrary}\}$.

3. G-inverses of Markovian kernels, I – P

G-inverses arise in many applied probability problems where Markov chains are present. The key result, upon which our presentation is based, is the following, (see [4]).

Let P be the transition matrix of a finite irreducible Markov chain $\{X_n\}$ *with m states and stationary probability vector* π'. *Let* $\mathbf{e}' = (1, 1, \ldots, 1)$ *and* $\mathbf{t}$ *and* $\mathbf{u}$ *be any vectors.*

(a) $I - P + \mathbf{tu}'$ *is non-singular if and only if* $\pi'\mathbf{t} \neq 0$ and $\mathbf{u}'\mathbf{e} \neq 0$.

(b) *If* $\pi'\mathbf{t} \neq 0$ *and* $\mathbf{u}'\mathbf{e} \neq 0$ *then* $[I - P + \mathbf{tu}']^{-1}$ *is a g-inverse of* $I - P$.

From these results it can easily be shown that all one condition g-inverses of I – P are of the form $[I - P + \mathbf{tu}']^{-1} + \mathbf{ef}' + \mathbf{g}\pi'$ for arbitrary vectors $\mathbf{f}$ and $\mathbf{g}$.

Special cases include the following:

(a) Kemeny and Snell's 'fundamental matrix' $Z \equiv [I - P + \Pi]^{-1}$, where $\Pi = \mathbf{e}\pi'$, ([12]), was shown in [4] to be a one condition g-inverse of I – P.

(b) Styan, Paige, and Wachter [15] showed that the Moore-Penrose g-inverse of I – P can be expressed as $[I - P + \alpha\pi\mathbf{e}']^{-1} - \alpha\Pi$ where $\alpha = (m\pi'\pi)^{-1/2}$.

(c) Meyer [14] showed that $T \equiv [I - P + \Pi]^{-1} - \Pi$ is the group inverse of I – P.

(d) If $P = \begin{bmatrix} \mathbf{P}_{11} & \alpha \\ \beta' & p_{mm} \end{bmatrix}$ then ([5]), ([16]), I – P has a g-inverse of the form

$$\begin{bmatrix} (I - P_{11})^{-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} = [I - P + \mathbf{tu}']^{-1} + \mathbf{ef}',$$

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

4. Stationary Distributions

In many stochastic processes the determination of stationary distributions is an important problem as it leads, in the case of aperiodic irreducible processes, to knowledge of the limiting distribution. Since the determination involves solving systems of linear equations, we develop a collection of results for various types of stochastic processes using the generalized inverse approach.

4.1. Discrete Time Markov Chains

The stationary probability vector $\pi' = (\pi_1, \pi_2, \ldots, \pi_m)$ for the irreducible Markov chain, with transition matrix $P = [p_{ij}]$, is given by the solution of the equations

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

or, in matrix form, $\pi'(I - P) = \mathbf{0}'$ with $\pi'\mathbf{e} = 1$.

This is an equation of type XB = C with X = **π'**, B = I – P, C = **0'** and using, the technique of 2(i), it can be shown, (see [5], [6]), that:

If is any g-inverse of I – P *and* A ≡ I – (I – P)G,

then $\pi' = \dfrac{\mathbf{v}' A}{\mathbf{v}' A\mathbf{e}}$ *where* **v'** *is such that* **v'**A**e** ≠ 0.

[Note that A**e** ≠ 0 for all g-inverses and thus we can always find such a **v'**.]

Special cases:

(a) For any **u**, **t** such that **u'e** ≠ 0, **π't** ≠ 0, (Paige, Styan and Wachter [15], Kemeny [13]),

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

In particular: $\pi' = \mathbf{u}'[I - P + \mathbf{e}\mathbf{u}']^{-1}$.

[A recommendation choice fur **u'** is **u'** = $\mathbf{e}_j$'P for some j, ([15]).]

(b) If $(I – P)^-$ is a g-inverse such that $\mathbf{e}'[I -(I - P)(I - P)^-]\mathbf{e} \neq 0$, then (Decell and Odell [1]),

$$\pi' = \frac{\mathbf{e}'[I - (I - P)(I - P)^-]}{\mathbf{e}'[I - (I - P)(I - P)^-]\mathbf{e}}.$$

(c) (Rhode [16], Meyer [14]):

If $P = \begin{bmatrix} \mathbf{P}_{11} & \alpha \\ \beta' & p_{mm} \end{bmatrix}$ then $\pi' = \dfrac{(\beta'(I - P_{11})^{-1}, 1)}{(\beta'(I - P_{11})^{-1}, 1)\mathbf{e}}$.

(d) The simplest procedure for determining π' using generalised inverses is given in Hunter [11]. For any g-inverse G of I – P let A = I – (I – P)G = $[a_{ij}]$. Let r be the smallest integer i such that

$$\sum_{k=1}^{m} a_{ik} \neq 0. \text{ Then } \pi' = (\pi_1, \pi_2, \ldots \pi_m) \text{ where } \pi_j = \frac{a_{rj}}{\sum_{k=1}^{m} a_{rk}}, \; j = 1, 2, \ldots, m.$$

4.2. Semi-Markov Processes

Let $\{(X_n, T_n)\}$ be a Markov renewal process with state space S = {1, 2, …, m} and semi-Markov kernel $Q(x) = [Q_{ij}(x)]$, where $Q_{ij}(x) = P[X_{n+1} = j, T_{n+1} – T_n \le x | X_n = i]$.

Let $Q_{ij}(+\infty) = p_{ij}$ and $P = [p_{ij}]$ so that P is the transition matrix of the embedded Markov chain.

The following results are well known.

1. If $\{X_n\}$ is irreducible then the embedded Markov chain has a stationary probability vector **π'**.

2. If $Y_t = X_n$, $T_n \le t < T_{n+1}$, (under the assumption that sup $T_n = \infty$), then $\{Y_t\}$ is a semi-Markov process. If $\{X_n\}$ irreducible and $\{(X_n, T_n)\}$ is aperiodic then

$$\nu_j = \lim_{t \to \infty} P\{Y_t = j | X_0 = i\} \text{ exists.}$$

Let $\nu' = (\nu_1, \nu_2, \ldots, \nu_m)$ be the stationary probability vector of $\{Y_t\}$.

Let $\mu' = (\mu_1, \mu_2, \ldots, \mu_m)$ and $\Lambda = \text{diag}(\mu_1, \mu_2, \ldots, \mu_m)$ where

$$\mu_i = E[T_{n+1} - T_n \mid X_n = i] = \int_0^\infty \left[1 - \sum_j Q_{ij}(t)\right] dt \ (< \infty).$$

Then $\nu' = \dfrac{\pi'\Lambda}{\pi'\mu} = \dfrac{1}{\lambda_1}\pi'\Lambda$.

Using the results developed for discrete time chains (see [5]).

For any **t**, **u** *such* $\mathbf{u'e} \neq 0, \pi'\mathbf{t} \neq 0$, $\nu' = \dfrac{\mathbf{u}'[I - P + \mathbf{tu}']^{-1}\Lambda}{\mathbf{u}'[I - P + \mathbf{tu}']^{-1}\mu}$.

In particular, $\nu' = \mathbf{u}'[I - P + \mu\mathbf{u}']^{-1}\Lambda$.

4.3. Continuous Time Markov Chains

If the semi-Markov kernel Q has the form $Q_{ij}(x) = p_{ij}[1 - \exp(-\lambda_i x)]$, $x \geq 0$, with $p_{ii} = 0$, then the semi-Markov process $\{Y_t\}$ is a Markov chain in continuous time. It is traditional to examine such processes via their infinitesimal generator $Q = [q_{ij}]$, where

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

This implies that the embedded Markov chain has transition probabilities

$$p_{ij} = \begin{cases} 0, & i = j, \\ \dfrac{q_{ij}}{\sum_{k \neq i} q_{ik}} = \dfrac{q_{ij}}{q_{ii}}, & i \neq j, \end{cases}$$

so that $I - P = (Q_d)^{-1}Q$.

Furthermore, P is irreducible

if and only if ∃ a unique positive π' such that $\pi' = \pi'P$, with $\pi'e = 1$,

if and only if ∃ a unique positive ν' such that $\nu'Q = \mathbf{0}'$, with $\nu'\mathbf{e} = 1$.

As was the case for g-inverses for $I - P$, it can be verified that

(i) $Q + \mathbf{tu}'$ is non-singular ⇔ $\mathbf{u'e} \neq 0$ then $(Q + \mathbf{tu}')^{-1}$ is a g-inverse of Q.

There are two approaches that can be used to find ν - either via the embedded M.C. and the S.M.P. $\{Y_t\}$ or via the infinitesimal generator Q and a g-inverse of Q. Both approaches lead to equivalent results.

Typical results that can be obtained, ([5]), are the following:

$$\nu' = \mathbf{u}'\left[I - P + Q_d^{-1}\mathbf{eu}'\right]^{-1} Q_d^{-1},$$

$$\nu' = \frac{\mathbf{u}'\left[Q + \mathbf{tu}'\right]^{-1}}{\mathbf{u}'\left[Q + \mathbf{tu}'\right]^{-1}\mathbf{e}},$$

$$\nu' = \mathbf{u}'\left[Q + \mathbf{eu}'\right]^{-1}.$$

5. Moments of First Passage Time Distributions

5.1. Discrete Time Markov Chains

Let $\{X_n\}$ be a Markov chain with $S = \{1, 2, \ldots, m\}$.
Let $T_{ij} = \min\{n: X_n = j | X_0 = i\}$, the first passage time from i to j.
If $\{X_n\}$ is irreducible then the T_{ij} are proper r.v.'s and $m_{ij}^{(r)} \equiv ET_{ij}^{(r)} < \infty$.

We can deduce expressions from $[m_{ij}^{(r)}]$, (see [4], [5]), but in this paper we consider only expressions for $m_{ij} \equiv m_{ij}^{(1)}$, the mean first passage time from state i to state j.
It is well known that

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj},$$

so that if $M = [m_{ij}]$, $M_d = [\delta_{ij} m_{ij}]$, and $E = [1]$ then

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

This is an equation of the type $AX = C$ which can be solved by g-inverses.
The arbitrary nature of the solution suggested by the procedure 2(ii) disappears when the additional constraints $m_{jj} = 1/\pi_j$, $(j = 1, 2, \ldots, m)$, ie. $M_d = (\Pi_d)^{-1}$ where $\Pi = \mathbf{e}\boldsymbol{\pi}'$, are used, ([5]).

If G is any g-inverse of $I - P$ *then* $M = [G\Pi - E(G\Pi)_d + I - G + EG_d]D$, *where* $D = (\Pi_d)^{-1}$.

Special cases of interest are the following:
(a) If $G = [I - P + \mathbf{eu'}]^{-1} + \mathbf{ef'}$ then $M = [I - G + EG_d]D$. G-inverses of this form include Kemeny and Snell's Z and Meyer's group inverse T.

(b) If $G = [I - P + \mathbf{eu'}]^{-1}$ then $M = [I - G + EG_d][(\mathbf{eu'}G)_d]^{-1}$.

(c) If $P = \begin{bmatrix} \mathbf{P}_{11} & \alpha \\ \beta' & p_{mm} \end{bmatrix}$ and $\mathbf{a} = (I - P_{11})^{-1}\mathbf{e}$, $\mathbf{b'} = \beta'(I - P_{11})^{-1}$, $\Delta = 1 + \mathbf{b'e}$, then ([5]),

$$M = \begin{bmatrix} [\mathbf{ab'} - E(\mathbf{ab'})_d + \Delta\{1 - (I - P_{11})^{-1} + E(I - P_{11})^{-1}\}_d]((\mathbf{eb'})_d)^{-1} & a \\ \Delta\mathbf{e'}((I - P_{11})^{-1})_d - \mathbf{e'}(\mathbf{ab'})_d]((\mathbf{eb'})_d)^{-1} & \Delta \end{bmatrix}$$

(d) Let G be any inverse of $I - P$ and define $C = I - G(1 - \Pi)$. Then $M = [C - ECd + E]D$, (Hunter, [10]).

(e) (Hunter [10]) If $A = I - (I - P)G = [a_{ij}]$ and $G = [g_{ij}]$. If r is the first index i $(1 \leq r \leq m)$ such that $\sum_{j=1}^{m} a_{ij} \neq 0$, then $m_{jj} = \left(\sum_{k=1}^{m} a_{rk}\right) / a_{rj}$ and for $i \neq j$

$$m_{ij} = [(g_{ii} - g_{ij})\left(\sum_{k=1}^{m} a_{rk}\right) + (a_{rj} - a_{ri})\left(\sum_{k=1}^{m} g_{ik}\right) / a_{rj}$$

(f) See also Heyman and O'Leary ([3]) for further computational procedures using group generalized inverses.

5.2. Markov Renewal Processes

Let $\{(X_n, T_n)\}$ be a Markov renewal process and let T_{ij} be the duration of a first passage from state i to state j, $(i, j \in S = \{1, 2, \ldots, m\})$.

Let $G_{ij}(t)$ be the distribution function of T_{ij}.

If $m_{ij} = \int_{0-}^{\infty} t dG_{ij}(t)$ and $\mu_{ij} = \int_{0-}^{\infty} t dQ_{ij}(t)$, then $M = [m_{ij}]$ satisfies the equation
$(I - P)M = P^{(1)}E - PD$ where $D = M_d = \lambda_1(\Pi_d)^{-1}$, $P^{(1)} = [\mu_{ij}]$ with $\lambda_1 = \pi'\mu$ and $\mu = P^{(1)}\mathbf{e}$.

This equation is of the form AX = C and with M_d specified its solution is ([5]):

If G is any g-inverse of I – P, $M = \left[\frac{1}{\lambda_1} GP^{(1)}P - \frac{1}{\lambda_1} E(GP^{(1)})_d + I - G + EG_d\right]D.$

An earlier derivation, ([4]), utilized Kemeny and Snell's Z. Other useful simplifications include:

(a) If $G = [I - P + \mu\mathbf{u'}]^{-1} + \mathbf{ef'}$ with $\mathbf{u'e} \neq 0$ and $\mu = P^{(1)}\mathbf{e}$, $M = [I - G + EG_d]D$.

(b) If $G = [I - P + \mu\mathbf{u'}]^{-1}$ with $\mathbf{u'e} \neq 0$ then $M = [I - G + EG_d][(\mathbf{eu'}G)_d]^{-1}$.

6. Occupation Time Random Variables

Another application arises in examining the asymptotic behaviour of the number of times particular states are entered.

6.1 Discrete Time Markov Chains

Given a Markov chain $\{X_n\}$, let $M_{ij}^{(n)}$ = Number of k ($0 \leq k \leq n$) such that $X_k = j$ given $X_0 = i$. Then

$$\left[EM_{ij}^{(n)}\right] = \left[\sum_{k=0}^{n} p_{ij}^{(k)}\right] = \sum_{k=0}^{n} P^k.$$

Let $A_n = \sum_{k=0}^{n-1} P^k$ then $(I - P)A_n = I - P^n$ or $A_n(I - P) = I - P_n$, subject to $A_n\Pi = \Pi A_n = n\Pi$.
For G, any g-inverse of I – P, ([5], [6]),

$$A_n = \sum_{k=0}^{n-1} P^k = \begin{cases} n\Pi + (I - P)G(I - P^n), \\ n\Pi + (I - P^n)G(I - P) \end{cases}$$

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

$$\text{Thus } \left[EM_{ij}^{(n)}\right] = (n + 1)\Pi + (I - \Pi)G(I - \Pi) + o(1)E.$$

A simple obsevation from this last result is that $(I - \Pi)G(I - \Pi)$ is invariant for all G. In fact it is equivalent to T, the group inverse of I – P, ([5]).

6.2. Markov Renewal Processes

Let $M(t) = \left[M_{ij}(t)\right] = \left[\sum_{n=1}^{\infty} Q_{ij}^{(n)}(t)\right] = \left[EN_{ij}(t)\right]$, be the 'Markov renewal matrix'. Note that $N_{ij}(t)$ is the number of times state j is entered, staring in i, in [0, t]. Expressions for the moments of these occupation times can be derived from renewal theoretic arguments:

$$M_{ij}(t) = \frac{t}{m_{ij}} + \left(\frac{m_{jj}^{(2)}}{2m_{jj}} - \frac{m_{ij}}{m_{jj}}\right) + o(1), \quad \text{where } m_{ij}^{(k)} = \int_0^\infty t^k dG_{ij}(t).$$

(a) For any g-inverse G of I – P, it can be shown, ([5]),

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

where $\lambda_2 = \sum_{i=1}^{m}\sum_{j=1}^{m}\pi_i \int_0^\infty t^2 dQ_{ij}(t)$.

This was first established with G = Z, ([4]).

Other special cases include:

(1) If $G = [I - P + \mu\mathbf{u}']^{-1}$ with $\mathbf{u}'\mathbf{e} \neq 0$, then ([5]):

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

(2) If $G = [I - P + \mu\pi' P^{(1)}]^{-1}$ then ([5]):

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

(3) Expressions of the analogous results for continuous time Markov chains follow with

$P^{(1)} = -(Q_d)^{-1}P$, $\mu = -(Q_d)^{-1}\mathbf{e}$, $\lambda_1 = -\pi' Q^{-1}{}_d\mathbf{e}$, $\lambda_2 = 2\pi'(Q^{-1}{}_d)^2\mathbf{e}$.

7. Classification of g-inverses of I – P

7.1. Fully Efficient Characterizations

Let A = I – P and assume that $\pi'\mathbf{t} \neq 0$, $\mathbf{u}'\mathbf{e} \neq 0$. A{1}, the class of all one condition g-inverses of I – P, has various characterizations:

(a) $A\{1\} = \{[I - P + \mathbf{t}\mathbf{u}']^{-1} + \mathbf{e}\mathbf{f}' + \mathbf{g}\pi', \mathbf{f}, \mathbf{g} \text{ arbitrary}\}$.

(b) $$A\{1\} = \left\{[I - P + \mathbf{t}\mathbf{u}']^{-1} - \frac{\mathbf{e}\mathbf{u}'H}{\mathbf{u}'\mathbf{e}} - \frac{H\mathbf{t}\pi'}{\pi'\mathbf{t}} + \frac{\mathbf{e}\mathbf{u}'H\mathbf{t}\pi'}{(\mathbf{u}'\mathbf{e})(\pi'\mathbf{t})},\ H \text{ arbitrary}\right\}.$$

Observe that if P is m × m and **t** and **u** are given, (a) has 2m arbitrary elements and (b) has m^2 arbitrary elements. A fully efficient characterization, with specified **t** and **u**, requires 2m – 1 arbitrary elements, ([8]):

Let $\mathbf{u}' = (\mathbf{u}'_{m-1}, u_m)$ then

$A\{1\} = \{[I - P + \mathbf{t}\mathbf{u}']^{-1} + \mathbf{e}\mathbf{y}'_m + B\mathbf{z}_{m-1}\pi'$, where

$$B = \begin{bmatrix} I - \left(\frac{1}{\mathbf{u'e}}\right)\mathbf{eu'}_{m-1} \\ -\left(\frac{1}{\mathbf{u'e}}\mathbf{u'}_{m-1}\right) \end{bmatrix}$$, $\mathbf{y'}_m$, $\mathbf{z}_{m-1}$ arbitrary $1 \times m$, $(m-1) \times 1$ vectors}.

The representation is fully efficient and that given any $A^{(1)}$, **t**, **u**,

$$\mathbf{y'}_m = \frac{\mathbf{u'}}{\mathbf{u'e}}\left\{A^{(1)} - (I - P + \mathbf{tu'})^{-1}\right\},\ \mathbf{z'}_{m-1} = \left[I \ -e\right]\left\{A^{(1)} - (I - P + \mathbf{tu'})^{-1}\right\}\frac{\mathbf{t}}{\pi'\mathbf{t}}.$$

There are some disadvantages in that this representation requires preselected **t**, **u**.
Note, however, that we can convert from a form with given $\mathbf{t}_1$, $\mathbf{u}_1$ to a form with $\mathbf{t}_2$, $\mathbf{u}_2$, ($\mathbf{u'}_i e \neq 0$, $\pi'\mathbf{t}_i \neq 0$) using the result ([8]):

$$\left[I - P + \mathbf{t}_2\mathbf{u}_2'\right]^{-1} = \left[I - \frac{\mathbf{eu'}_2}{\mathbf{u'}_2\mathbf{e}}\right]\left[I - P + \mathbf{t}_1\mathbf{u'}_1\right]^{-1}\left[I - \frac{\mathbf{t}_2\pi'}{\pi'\mathbf{t}_2}\right] + \frac{\mathbf{e}\pi'}{(\pi'\mathbf{t}_2)(\mathbf{u'}_2\mathbf{e})}.$$

A further subsidiary observation is that given any g-inverse G of I – P and **t**, **u**, with $\pi'\mathbf{t} \neq 0$, $\mathbf{u'e} \neq 0$, we can computer $[I - P + \mathbf{tu'}]^{-1}$:

$$\left[I - P + \mathbf{tu'}\right]^{-1} = \left[I - \frac{\mathbf{eu'}}{\mathbf{u'e}}\right]G\left[I - \frac{\mathbf{t}\pi'}{\pi'\mathbf{t}}\right] + \frac{\mathbf{e}\pi'}{(\pi'\mathbf{t})(\mathbf{u'e})}.$$

7.2. Multi-condition g-inverses of I – P

From the representation for a typical one condition g-inverse $A^{(1)}$:
$A^{(1)} = [I - P + \mathbf{tu'}]^{-1} + \mathbf{e'f} + \mathbf{g'}\pi$,
We can derive conditions that result in particular multi-condition g-inverses of I – P,

(1) $A^{(1)} \in \{A^{(1,2)}\} \Leftrightarrow \frac{1}{(\mathbf{u'e})(\pi'\mathbf{t})} + \frac{\mathbf{f't}}{\pi'\mathbf{t}} + \frac{\mathbf{u't}}{\mathbf{u't}} = \mathbf{f'}(I - P)\mathbf{g}.$

(2) $A^{(1)} \in \{A^{1,3)}\} \Leftrightarrow \frac{\mathbf{t}}{\pi'\mathbf{t}} - (I - P)\mathbf{g} = \frac{\pi}{\pi'\pi}.$

(3) $A^{(1)} \in \{A^{(1,4)}\} \Leftrightarrow \frac{\mathbf{u'}}{\mathbf{u'e}} - \mathbf{f'}(I - P) = \frac{\mathbf{e'}}{\mathbf{e'e}}.$

(4) $A^{(1)} \in \{A^{(1,5)}\} \Leftrightarrow \frac{\mathbf{t}}{\pi'\mathbf{t}} - (I - P)\mathbf{g} = \mathbf{e}\ \textit{and}\ \frac{\mathbf{u'}}{\mathbf{u'e}} - \mathbf{f'}(I - P) = \frac{\mathbf{e'}}{\mathbf{e'e}}.$

Similarly we can write down conditions for $A^{(1,2,3)}$, $A^{(1,2,4)}$, $A^{(1,3,4)}$, $A^{(1,2,5)}$. In particular the Moore-Penrose g-inverse of I – P can be expressed as

$$A^{(1,2,3,4)} = [I - P + \pi\mathbf{e'}]^{-1} - \frac{\mathbf{e}\pi'}{m\pi'\pi}.$$

This is different form to that earlier reported. However, equivalence follows from the result, ([8]):

$$A_\delta \equiv [I - P + \delta \mathbf{tu'}]^{-1} - \frac{\mathbf{e}\pi'}{\delta(\pi'\mathbf{t})(\mathbf{u'e})} \text{ does not depend on } \delta \neq 0.$$

7.3. Parametric Forms of g-inverses of I – P

Given any g-inverse, G, *of* I – *P there exists unique parameters* α, β, γ *with the property that* $\pi'\alpha = 1$, $\beta'\mathbf{e} = 1$, *and* $G = [I - P + \alpha\beta']^{-1} + \gamma\mathbf{e}\pi'$, [9].

The construction is as follows:

Given G, let $A = I - (I - P)G$ $(= \alpha\pi')$ and $B = I - (I - P)G$ $(= \mathbf{e}\beta')$.

Then $\alpha = A\mathbf{e}$, $\beta' = \pi'B = e'_iB$ for all i, and $\delta + 1 = \pi'G\alpha = \beta'G\mathbf{e} = \beta'G\alpha$.

An important application is that if G has the representation $G(\alpha, \beta, \nu)$:

$G \in A\{1, 2\} \Leftrightarrow \gamma = -1$,

$G \in A\{1, 3\} \Leftrightarrow \alpha = \pi/\pi'\pi$,

$G \in A\{1, 4\} \Leftrightarrow \beta' = \mathbf{e}'/m$,

$G \in A\{1, 5\} \Leftrightarrow \alpha = \mathbf{e}, \beta' = \pi'$,

so that we have a much easier procedure for classifying the g-inverses of I – P. Knowledge of the parameters α, β, γ can provide useful information in determining the stationary distribution and moments of the first passage time distributions. Expressions for π' in terms of G for such special cases are given in [(11)].

8. Perturbed Markov Chains

The generalized inverse method for finding stationary distributions has also proved useful in examining the relationship between the stationary distributions of two Markov chains that differ by a perturbation, ([7]):

Let $P^{(1)}$, $P^{(2)}$ be the transition matrices of two finite, irreducible, m-state M.C's with stationary probability vectors $\pi^{(1)'}$, $\pi^{(2)'}$.

Suppose $P^{(2)} = P^{(1)} + \mathbf{ab'}$ with $\mathbf{b'e} = 0$.

From the earlier theory we have that, provided $\pi^{(i)'}\mathbf{t}_i \neq 0$ and $\mathbf{u'}_i\mathbf{e} \neq 0$,

$$\pi^{(1)'} = \frac{\mathbf{u'}_i[I - P^{(i)} + \mathbf{t}_i\mathbf{u}_i']^{-1}}{\mathbf{u'}_i[I - P^{(i)} + \mathbf{t}_i\mathbf{u}_i']^{-1}\mathbf{e}}.$$

The general idea is to try and use the same inverse in both computations.

ie. $I - P^{(2)} + \mathbf{t}_2\mathbf{u'}_2 \equiv I - P^{(1)} + \mathbf{t}_1\mathbf{u'}_1 \Leftrightarrow \mathbf{ab'} = \mathbf{t}_2\mathbf{u}_2' - \mathbf{t}_1\mathbf{u}_1'$.

This indicates that we should take $\mathbf{t}_1 = \mathbf{a}$, $\mathbf{t}_2 = \mathbf{a}$, $\mathbf{u}_1 = \mathbf{u}_1 = \mathbf{u}$, $\mathbf{u}_2 = \mathbf{b} + \mathbf{u}_1$.

If $\mathbf{u'e} \neq 0$, $\pi^{(1)'}\mathbf{a} \neq 0$, $\alpha' = \mathbf{u'}[I - P^{(1)} + \mathbf{au'}]^{-1}$, $\beta' = \mathbf{b'}[I - P^{(1)} + \mathbf{au'}]^{-1}$,

then $$\pi^{(1)'} = \frac{\alpha'}{\alpha'\mathbf{e}} \text{ and } \pi^{(2)'} = \frac{\alpha' + \beta'}{\alpha'\mathbf{e} + \beta'\mathbf{e}}.$$

If instead we set $I - P^{(2)} + \mathbf{t}_2\mathbf{u'}_2 \equiv I - P^{(1)} + \mathbf{t}_2\mathbf{u'}_1$ and use the result:

$$[I - P^{(1)} + \mathbf{t}_2\mathbf{u'}_1]^{-1} = [I - P^{(1)} + \mathbf{t}_1\mathbf{u'}_1]^{-1}\left[I - \frac{\mathbf{t}_2\pi^{(1)'}}{\pi^{(1)}\mathbf{t'}_2}\right] + \frac{\mathbf{e}\pi^{(1)'}}{(\pi^{(1)}\mathbf{t'}_2)(\mathbf{u'}_1\mathbf{e})},$$

we obtain the following more general result:

If $\mathbf{u'e} \neq 0$, $\pi^{(1)'}\mathbf{t} \neq 0$, $\alpha' = \mathbf{u'}[I - P^{(1)} + \mathbf{tu'}]^{-1}$ *and* $\beta' = \mathbf{b'}[I - P^{(1)} + \mathbf{tu'}]^{-1}$,

then $\pi^{(1)'} = \dfrac{\alpha'}{\alpha'\mathbf{e}}$ and $\pi^{(2)'} = \dfrac{(\alpha'\mathbf{a})\beta' + (1 - \beta'\alpha)\alpha'}{(\alpha'\mathbf{a})(\beta' e) + (1 - \beta'\alpha)\alpha'\mathbf{e}}$.

This implies that we can develop a recursive procedure that requires only one initial matrix inversion to generate stationary distributions for successive perturbations.

Some applications of interest include:

(i) Transition matrices differing in a single row, say the r^{th} row:
$\mathbf{a} = \mathbf{e}_r$, $\mathbf{b'} = \mathbf{p}_r^{(2)'} - \mathbf{p}_r^{(1)'}$.

(ii) Transition matrices differing only in two elements in same row:
$\mathbf{a'} = \mathbf{e}_r$, $\mathbf{b'} = \varepsilon(\mathbf{e'}_c - \mathbf{e'}_d)$ where $\varepsilon = p_{rc}^{(2)} - p_{rc}^{(1)} = p_{rd}^{(1)} - p_{rd}^{(2)}$.

(iv) Updating approximations to stationary distributions as observations are obtained.
Let n_{ij} be the number of transitions observed from i to j.

$$n_i = \sum_{j=1}^{m} n_{ij}, \quad n = \sum_{i=1}^{m} n_i.$$

Thus an initial estimate of the (i, j)th transition probability is given by

$$p_{ij}^{(1)} = \frac{n_{ij}}{n_i}, \quad i, j = 1, 2, \ldots, m.$$

Suppose at the next update a transition from state r to state s occurs yielding the revised approximations:

$$p_{ij}^{(2)} = \begin{cases} p_{ij}^{(1)}, & i \neq r,\ j = 1, \ldots, m, \\ n_{rj}/(n_r + 1), & i = r,\ j \neq s, \\ (n_{rs} + 1)/(n_r + 1), & i = r,\ j = s \end{cases}$$

$\Rightarrow \mathbf{a} = \mathbf{e}_r$, $\mathbf{b'} = \varepsilon(\mathbf{e}_s - p_r^{(1)})$ where $\varepsilon = 1/(n_r + 1)$.

9. Algorithm for Finding Stationary Distributions of Markov Chains

The perturbation procedure of the previous section can be adapted to finding the stationary distribution of any finite irreducible Markov chain without having to either solve a system of linear equations or invert a matrix.

Let P_i be transition matrix of a Markov chain with stationary probability vector π'_i so that

$$\pi'_i = \frac{\alpha'_i}{\alpha'_i\,\mathbf{e}}, \text{ where } \alpha'_i = \mathbf{u'}_i A_i \text{ with } A_i = [I - P_i + \mathbf{t}_i\mathbf{u'}_i]^{-1}.$$

Now let $P_{i+1} = P_i + \mathbf{a}_{i+1}\mathbf{b'}_{i+1}$. If $\mathbf{u}_{i+1} = \mathbf{u}_i + \mathbf{b}_{i+1}$, $\mathbf{t}_{i+1} = \mathbf{a}_{i+1}$ then

$$A_{i+1} = A_i \left[I + (\mathbf{t}_i - \mathbf{t}_{i+1}) \frac{\pi'_i}{\pi'_i\,\mathbf{t}_{i+1}} \right].$$

Start with P_0, find α'_0, π'_0. Then perturb to P_1 and find α'_1, π'_1. Repeat by perturbing to P_2 and finding α'_2, π'_2. Carry out this procedure m times resulting in $P_m = P$ and finding α'_m, $\pi'_m = \pi'$.

This suggestion, implemented in [10], is to start with P_0, $\mathbf{t}_0$, $\mathbf{u}_0$ so that A_0 has a 'nice' form and make changes sequentially, row by row.

By starting with the simple 'doubly stochastic' matrix $P_0 = \frac{1}{m}\mathbf{ee}' = \left[\frac{1}{m}\right]$, with $\mathbf{t}_0 = \mathbf{e}$, $\mathbf{u}'_0 = \mathbf{e}'/m$, and $A_0 = I$ we have initially that $\pi'_0 = \mathbf{e}'/m$.

Now let $P_{i-1} = P_{i-1} + \mathbf{e}_i\mathbf{b}'_i$, $\mathbf{b}'_i = \mathbf{p}'_i - \mathbf{e}'/m$, where $\mathbf{p}'_i$ is the i-th row of P, ensuring that each P_i is irreducible. It can be shown that $A_{i+1} = A_i + B_i$, where all but the first (i+1) rows of B_i are zero, i = 1, …, m – 1. The algorithm is discussed in detail in [10]. See also Seneta [17].

Initial computer runs using the test matrices of Harrod & Plemmons, ([3]), with a program written in APL for a PC gave an accuracy to 14 d.p.s. The procedure appears to give very accurate results. The method is being explored in more detail including the starting with a "sparse" stochastic matrix P_0.

References:

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

2. Harrod, W.J. and Plemmons R.J., 'Comparison of some direct methods for computing stationary distributions of Markov chains', *SIAM J. Sci. Statist. Comput*., 5, 453-469, (1984)

3. Heyman, D.P. and O'Leary D.P. 'What is Fundamental for Markov chains? First passage times, Fundamental matrices, and group generalised Inverses' in *Computations with Markov chains*, Proceedings of the 2nd International Workshop on the Numerical Solution of Markov Chains, (Stewart W.J. Ed), Kluwer Academic Publishers, Dordrecht, pp151-159, (1995)

4. Hunter J.J. 'On the moments of Markov renewal processes', *Adv. Appl. Probab*., 1, 188-210, (1969)

5. Hunter J.J. 'Generalized inverses and their application to applied probability problems', *Linear Algebra Appl*., 45, 157-198, (1982)

6. Hunter, J.J. *Mathematical Techniques of Applied Probability, Volume 2, Discrete Time Models: Techniques and Applications*, Academic, New York, 1983

7. Hunter J.J. 'Stationary distributions of perturbed Markov chains', *Linear Algebra Appl.*, 82, 201-214, (1986)

8. Hunter J.J. 'Characterizations of generalized inverses associated with Markovian kernels', *Linear Algebra Appl.*,102, 121-142, (1988)

9. Hunter J.J. 'Parametric forms for generalized inverses of Markovian kernels and their applications', *Linear Algebra Appl.*, 127, 71-84, (1990)

10. Hunter J.J. 'The computation of stationary distributions of Markov chains through perturbations', *Journal of Applied Mathematics and Stochastic Analysis*, 4 (1), 29-46, (1991)

11. Hunter J.J. 'Stationary distributions and mean first passage times in Markov chains using generalised inverses', *Asia-Pacific Journal of Operational Research* 9, 145-153 (1992)

12. Kemeny J.G. and Snell, J.L. *Finite Markov Chains*, Van Nostrand, New York, 1960

13. Kemeny, J.G. ''Generalizations of a fundamental matrix', *Linear Algebra Appl*., 38, 193-206, (1981)

14. Meyer C.D. Jr. 'The role of the group generalized inverse in the theory of finite Markov chains', *SIAM Rev*.,17, 443-464, (1975)

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

16. Rhode, C.A., 'Special applications of the theory of generalized matrix inversion to statistics, in *Proceedings of the Symposium on Theory and Applications of Generalized Inverses of Matrices*, (Boullion T.L. and Odell P.L., Eds.), Texas Tech. Press, Lubbock, pp. 239-266, (1968)

17. Seneta, E. 'Sensitivity Analysis, Ergodicity Coefficients, and Rank-One Updates for finite Markov chains' in *Numerical Solution of Markov Chains*, (Stewart, W.J. Ed) Marcel Dekker Inc, New York, pp121-129, (1991)