The spectral approach to first passage time distributions for Markov processes requires knowledge of the eigenvalues and eigenvectors of the infinitesimal generator matrix. We demonstrate that in many cases knowledge of the eigenvalues alone is sufficient to compute the first passage time distribution.
IDENTIFYING COEFFICIENTS IN THE SPECTRAL REPRESENTATION FOR FIRST PASSAGE TIMES

By

Mark Brown
and Yi-Shi Shao

The City College, CUNY

Approved for public release;
distribution unlimited.

May 1986

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
OFFICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is approved for public release IAT AFOSR 84-12.
Distribution is unlimited.

"TH" L. KEEPER
Chief, Technical Information Division

City College, CUNY Report No. MB84-02
AFOSR Technical Report No. 84-02
AFOSR Grant No. 84-0095
Abstract. The spectral approach to first passage time distributions for Markov processes requires knowledge of the eigenvalues and eigenvectors of the infinitesimal generator matrix. We demonstrate that in many cases knowledge of the eigenvalues alone is sufficient to compute the first passage time distribution.

1. Introduction and Summary. Consider a continuous time Markov process with state space $N = \{0,1,2,\ldots\}$. For $j < n$ define $T_{j,n}$ to be first passage time from $j$ to $\{k: k \geq n\}$. Define $A(n)$ to be the $n \times n$ matrix obtained from the first $n$ rows and columns of the infinitesimal matrix $A$, i.e. the matrix of transition rates among the states $0, 1, \ldots, n-1$. If $A(n)$ is similar to a diagonal matrix then spectral analysis of $A(n)$ leads to the representation:

$$\Pr(T_{j,n} > t) = \sum_{i=1}^{n} \gamma_i e^{-\lambda_i t}.$$  

In (1.1), $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $A(n)$ and $\gamma_1, \ldots, \gamma_n$ depend on the right and left eigenvectors of $A(n)$.

Convenient references for this algebraic approach are Cinlar (1975), Karlin (1966), Keilson (1975) and Kemeny and Snell (1960).

In general in order to apply (1.1) a complete spectral analysis of $A(n)$ is required. The computational difficulties in performing this analysis limit the applicability of the spectral approach.
We consider a class of Markov processes in which upward jumps may only be of size 1 \((A_{i,i+1} > 0, A_{i,j} = 0, j \geq i+2)\) with no restriction on downward jumps. This class includes birth and death processes. For these Markov processes the coefficients \(y_1, \ldots, y_n\) can be computed from the eigenvalues of \(A(n)\), without computation of right or left eigenvectors. This result should simplify computation and render the spectral approach more attractive. A summary of the results now follows.

If \(\lambda_1, \ldots, \lambda_n\) are distinct then \(A(n)\) is similar to a diagonal matrix and (1.1) holds. In section 2 it is shown that:

\[
\Pr(T_{0,n} > t) = \prod_{i=1}^{n} \left( \frac{\lambda_i}{\lambda_{i+1}} \right)^{-\lambda_i t}.
\]

From (1.2) it follows that if \(\lambda_1, \ldots, \lambda_n\) are real then \(T_{0,n}\) is the convolution of exponentially distributed random variables with parameters \(\lambda_1, \ldots, \lambda_n\). If the \(\lambda_i\)'s are real but not distinct, say \(\lambda_1, \ldots, \lambda_r\) are distinct with multiplicities \(m_1, \ldots, m_r\), then \(T_{0,n}\) is distributed as the convolution of \(\Gamma(m_i, \lambda_i)\), \(i = 1, \ldots, r\), where \(\Gamma(m, \lambda)\) is the gamma distribution with parameters \(m\) and \(\lambda\). For convolutions of gammas we know of no nice analogue of (1.2) for representing the survival function.

For general \(j < n\) ((1.2) had \(j = 0\)), let \(w_1, \ldots, w_j\) be the eigenvalues of \(A(j)\). Consider the case where \(w_1, \ldots, w_j\) are distinct and \(\lambda_1, \ldots, \lambda_n\) are distinct. Define \(D = \{i: \lambda_i \cap \{w_1, \ldots, w_r\} = \phi\}\) and \(E = \{i: w_i \cap \{\lambda_1, \ldots, \lambda_n\} = \phi\}\). Then:

\[
\Pr(T_{j,n} > t) = \sum_{i \in D} \left( \prod_{j \neq i}^{j} \left( \frac{\lambda_j}{\lambda_i} \right)^{-\lambda_i t} \prod_{j \in E} \frac{w_j}{\lambda_i} \right) e^{-\lambda_i t}.
\]
The class of processes under consideration can be described as birth and death processes with multiple deaths (but not multiple births) allowed. Birth and death and related processes have significant application in queueing, reliability, inventory and biological models. Good references for these applications are Cinlar (1975), Gertsbakh (1984), Karlin (1966), Keilson (1975), and Ross (1983).

In section 3 we attack the problem of identifying $\gamma_1, \ldots, \gamma_n$ in (1.1) for a general Markov process.

2. Derivation of Results. Consider a Markov process $\{X(t), t \geq 0\}$ with state space $N = \{0, 1, 2, \ldots\}$. The infinitesimal generator matrix $A$ is defined by $A_{ij} = \lim_{h \to 0} \frac{\Pr(X(h)=j|X(0)=i)}{h}$ for $i \neq j$, and $A_{ii} = \lim_{h \to 0} \frac{\Pr(X(h)=i|X(0)=i)-1}{h}$. We consider processes which satisfy:

(i) $A_{i,i+1} = b_i > 0$

(ii) $A_{i,j} = 0$ for $j \geq i+2$

(iii) $-A_{i,i} < -\infty$ for all $i$

The $n \times n$ matrix $A(n)$ is the submatrix of $A$ consisting of the first $n$ rows and first $n$ columns, i.e. those corresponding to states $0, 1, \ldots, n-1$. It is not hard to show that under (i), (ii) and (iii) above, $A(n)$ is similar to a diagonal matrix if and only if its eigenvalues $\lambda_1, \ldots, \lambda_n$ are distinct. Thus the spectral representation (1.1) holds if and only if $\lambda_1, \ldots, \lambda_n$ are distinct.
Proof of (1.2). In order for state \( n \) to be reached there must be at least one transition from \( i \) to \( i+1 \) for \( i = 0, 1, \ldots, n-1 \).

Thus \( T_{0,n} \) is stochastically larger than the convolution of \( n \) exponentially distributed random variables with parameters \( b_0, \ldots, b_{n-1} \), which in turn is stochastically larger than \( M \), the maximum of \( n \) independent exponentially distributed random variables with parameters \( b_0, \ldots, b_{n-1} \). Define \( F_{0,n}^{(t)} = \Pr(T_{0,n} \leq t) \). Then:

\[
F_{0,n}^{(t)} \leq \Pr(M \leq t) = t^n \left( \prod_{i=0}^{n-1} b_i \right) + O(t^n) \quad \text{as} \quad t \to 0.
\]

Assume that \( \lambda_1, \ldots, \lambda_n \) are distinct, and thus that (1.1) holds. Then \( F_{0,n}^{(t)} \) is an analytic function which by (2.1) is \( O(t^{n-1}) \) as \( t \to 0 \). It is clear from the Taylor series expansion of \( F_{0,n}^{(t)} \) around zero that \( F_{0,n}^{(t)} = O(t^{n-1}) \) implies:

\[
\left. \frac{d^k}{dt^k} F_{0,n}^{(t)} \right|_{t=0} = 0 \quad \text{for} \quad k = 1, \ldots, n-1.
\]

From (1.1) and (2.2):

\[
\sum_{i=1}^{n} \gamma_i = \begin{cases} 1 \text{ for } k = 0 \\ 0 \text{ for } k = 1, \ldots, n-1 \end{cases}
\]

Define \( \gamma' = (\gamma_1, \ldots, \gamma_n) \), \( \bar{W} \) to be an \( n \times n \) matrix with components \( W_{ik} = \lambda_i^{k-1} \), \( i, k = 1, 2, \ldots, n \), and \( \bar{\delta} \) to be an \( n \times 1 \) vector with \( \delta_1 = 1, \delta_i = 0, i = 2, \ldots, n \). Rewrite (2.3) as:
Now \( \mathbf{W} \) is the well-known Vandermonde matrix (Cullen (1967) p. 72). For \( \lambda_1, \ldots, \lambda_n \) distinct \( \mathbf{W} \) is invertible. From (2.4) we see that \( y' \) concides with the first row of \( \mathbf{W}^{-1} \), thus:

\[
\gamma_1 = (\mathbf{W}^{-1})_{11} = \prod_{j \neq 1} \left( \frac{\lambda_j}{\lambda_j - \lambda_1} \right)
\]

Substituting \( \gamma \) into (1.1) yields (1.2). Next note that if \( \lambda_1, \ldots, \lambda_n \) are distinct then from (1.2):

\[
\psi(s) = \text{Ee}^{-sT_{0,n}} = \sum_{i=1}^{n} \left( \frac{\lambda_i}{\lambda_i - \lambda_1} \right) \frac{\lambda_i}{\lambda_i + s} = p(s) \prod_{i=1}^{n} \frac{\lambda_i}{\lambda_i + s}
\]

where

\[
p(s) = e^{\prod_{i=1}^{n} \left( \frac{\lambda_i + s}{\lambda_j - \lambda_i} \right)}.
\]

Now \( p(s) - 1 \) is a polynomial of degree \( n - 1 \) with \( n \) distinct roots \( \lambda_1, \ldots, \lambda_n \). Thus \( p(s) - 1 \equiv 0 \) and it follows from (2.6) that:

\[
\psi(s) = \prod_{i=1}^{n} \frac{\lambda_i}{\lambda_i + s}.
\]

When \( \lambda_1, \ldots, \lambda_n \) are real they must be strictly positive (since \( \Pr(T_{0,n} > t) \to 0 \)), thus if \( \lambda_1, \ldots, \lambda_n \) are real and distinct it follows from (2.8) that \( T_{0,n} \) is distributed as a convolution of exponentials with parameters \( \lambda_1, \ldots, \lambda_n \).
Now consider the case of distinct real eigenvalues $\lambda_1, \ldots, \lambda_r$ with multiplicity $m_1, \ldots, m_r$. We can obviously construct a sequence of vectors $\lambda^{(k)}$ converging to this vector $\lambda$ with each $\lambda^{(k)}$ having distinct components. This can be done through a sequence of matrices $A^{(k)}(n)$ converging to $A(n)$. For the process corresponding to $A^{(k)}(n)$ the distribution of $T^{(k)}(0,n)$ is the convolution of $n$ exponentials and thus has Laplace transform:

\begin{equation}
\psi_k(s) = \prod_{i=1}^{n} \left[ \frac{\lambda_i^{(k)}}{\lambda_i^{(k)}+s} \right].
\end{equation}

Letting $k \to \infty$, $T^{(k)}(0,n)$ converges in distribution to $T_{0,n}$ and (2.7) converges to:

\begin{equation}
\psi(s) = \lim_{k \to \infty} \psi_k(s) = \prod_{i=1}^{r} \left[ \frac{\lambda_i}{\lambda_i+s} \right]^{m_i}.
\end{equation}

Since for $\lambda_1, \ldots, \lambda_r$ real (2.8) is the Laplace transform of a convolution of $\Gamma(m_i, \lambda_i)$ random variables, we see that $T_{0,n}$ has this distribution.

We next look at general $j < n$ and prove (1.3).

**Proof of (1.3).** Consider the Markov process with the restrictions previously imposed and assume that $\lambda_1, \ldots, \lambda_n$ are distinct as are $\omega_1, \ldots, \omega_j$ the eigenvalues of $A(j)$. Define $D$ and $E$ as in section 1, and $D^* = \{1, \ldots, n\}-D$. Since $A_{i,j} = 0$ for $j \geq i+2$, it follows that

\begin{equation}
T_{0,n} = T_{0,j} \star T_{j,n}
\end{equation}
where * denotes convolution. Applying (2.8) to both $T_{0,j}$ and $T_{0,n}$ and (1.1) to $T_{j,n}$ we obtain:

\[ (2.10) \quad \prod_{r=1}^{n} \frac{\lambda_r}{\lambda_r^{+s}} = \left( \prod_{l=1}^{j} \frac{w_l}{w_l^{+s}} \right) \sum_{k=1}^{n} \frac{Y_{k,k}}{\lambda_k^{+s}}. \]

As some $\lambda_r$ and $w_l$ may coincide we divide both sides of (2.10) by these common terms obtaining:

\[ (2.11) \quad \prod_{r \in D} \frac{\lambda_r}{\lambda_r^{+s}} = \frac{\prod_{l \in E} w_l}{\lambda_j^{+s}} \sum_{k=1}^{n} \frac{Y_{k,k}}{\lambda_k^{+s}}. \]

Multiply both sides of (2.11) by \( \prod_{j=1}^{n} (\lambda_j^{+s}) \) to obtain:

\[ (2.12) \quad \left( \prod_{r \in D} \lambda_r \right) \prod_{j \in D^*} (\lambda_j^{+s}) = \left( \prod_{l \in E} w_l \right) \sum_{k=1}^{n} \frac{Y_{k,k}}{\lambda_k^{+s}} \prod_{j \neq k} (\lambda_j^{+s}). \]

Set $s = -\lambda_i$ in (2.12). For $i \in D$ we obtain:

\[ (2.13) \quad y_i = \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} / \prod_{l \in E} \frac{w_l}{w_l - \lambda_i} \quad \text{for } i \in D. \]

For $i \in D^*$, when we set $s = -\lambda_i$ in (2.12), the left side vanishes while the right side equals $y_i$ multiplied by a non-zero quantity. Thus:

\[ (2.14) \quad y_i = 0 \quad \text{for } i \in D^*. \]

Combining (2.13) and (2.14) we obtain (1.3).
In the case where either the $\lambda$'s or $w$'s are not distinct we can obtain the Laplace transform of $T_{j,n}$ (the Laplace transform of $T_{0,n}$ divided by that of $T_{0,j}$) but have no convenient representation for the distribution function.

3. Further Comments.

(3.1) Our method worked well because the first $n-1$ derivatives of $F_{0,n}(t)$ were equal to zero for the class of processes considered. Applying the method to a general Markov process we obtain a representation for $F_{0,n}(t)$ but it is not usually as computationally simple.

Assume that $\lambda_1, ..., \lambda_n$ are distinct which implies that (1.1) holds. Note that:

\[ v_k = \frac{d^k}{dt^k} \Pr(T_{j,n} \leq t) \bigg|_{t=0} = \sum_{i=1}^{n} \gamma_i \lambda_i^k. \]

(3.1a)

Note that $v_0=1$ and $v_k=0$ if it requires at least $k+1$ changes of state to go from $j$ to $\{k \geq n\}$. If $v' = (v_0, ..., v_{n-1})$ is known then:

\[ v' = v'w^{-1}. \]

(3.1b)

The problem reduces to computing $v_k$. Toward this end consider the lossy process with state space $0, ..., n-1$. The states $\{k \geq n\}$ are made absorbing and the Markov process restricted to $0, ..., n-1$ steadily loses probability mass \( \left( \sum_{j=0}^{n-1} \Pr(X(t)=j) < 1 \right) \) and thus called
lossy. Since $\Pr(T_0,n > t) = \Pr(X(t) \leq n-1)$ (recall that states \{k:k \geq n\} are absorbing):

\[(3.1c) \quad \Pr(T_j,n > t) = (e(n)^t)_{1}^{A(n)};\]

where $1$ is a column vector of $n$ 1's.

It follows that:

\[\frac{d^k}{dt^k} \Pr(T_j,n > t) \bigg|_{t=0} = (A_{(n)}^k)_{11};\]

and thus:

\[(3.1d) \quad v_k = (-1)^k \frac{d^k}{dt^k} \Pr(T_j,n > t) \bigg|_{t=0} = (-1)^k (A_{(n)}^k 1);\]

Thus $v_k$ equals $(-1)^k$ multiplied by the sum of the entries in the $j$th row of the matrix $A_{(n)}^k$.

(3.2) Keilson (1975) proves, using different methodology, that for a birth and death process with birth and death rates strictly positive, the eigenvalues of $A_{(n)}$ are indeed distinct and real and that $T_{0,n}$ is distributed as the convolution of $n$ exponentials with parameters $\lambda_1,\ldots,\lambda_n$.

(3.3) For the class of processes considered we proved that for $\lambda_1,\ldots,\lambda_n$ real $T_{0,n}$ is a convolution of exponentials with not necessarily distinct parameters. Such a distribution belongs to
the class $PF_n$ (Polya frequency function of order $n$). This class enjoys lots of nice properties (see Barlow and Proschan (1975)). The pdf. is log concave which leads to many interesting inequalities. It is also a subclass of IFR distributions thus all IFR inequalities are applicable.

(3.4) For the class of processes considered we do not know of any convenient to check conditions which insure that the eigenvalues are distinct or real.
Acknowledgement.

The authors are grateful for the helpful suggestions of Professors Stanley Kaplan and Julien Keilson.

Bibliography.


END
10-86
DTIC