Convergence and Correctness of an Iterative Scheme for Calculating Passage Times in Semi-Markov Processes

Helen J. Wilson

Department of Applied Mathematics, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK

h.j.wilson@leeds.ac.uk

Tel: +44 113 343 5002; Fax: +44 113 343 5090.

Abstract

An iterative scheme was introduced [Bradley, Dingle, Knottenbelt & Wilson, LAA 2003] using matrix-vector multiplication to calculate first passage times in semi-Markov processes. We show that the spectral radius of the matrix in question is at most 1, and prove that the method converges even when eigenvalues exist with modulus exactly 1. We also demonstrate correctness of the converged result.

Keywords: Semi-Markov, passage time, iterative, converge, spectral radius.

1 Introduction

Stochastic processes are used extensively in modelling of real-world systems. A Semi-Markov process [3] is a generalisation of a Markov chain which supports arbitrary state holding times. Thus a system can be modelled as having a (typically very large) number of possible states, with a transition matrix of probabilities representing which states may succeed a given state, and a matrix of distributions of the time taken for the system to carry out each such transition.

Passage-time analysis is a tool which is increasingly used in performance modelling, service contracts and quality-of-service agreements, allowing the assessment of questions of the form "what is the probability that task A will be completed within a given time window?".

In this paper we use an iterative process for calculating passage-time distributions, first introduced in [1]. This iterative scheme is intuitively clear, but guarantees of convergence and correctness have not been proved until now.

2 Definitions and Notation

We denote the transition matrix of the underlying Markov chain to our semi-Markov process as P, with elements p_{jk} representing the probability that, after being in state j, the next state the system finds itself in will be k. From the definition it is clear that the elements of P lie on the real axis between 0 and 1, and since the system will move from state j to some next state with probability 1,

$$\sum_k p_{jk} = 1 \quad \forall j. \tag{1}$$

The probability density function for the sojourn time of the transition from state j to state k is defined as $h_{jk}(t)$, and is an arbitrary nonnegative real function defined over $t \geq 0$ such that

$$\int_{t=0}^{\infty} h_{jk}(t)\, dt = 1.$$

We define elements of a matrix $R(s)$ for any fixed complex s with nonnegative real part:

$$[R]_{jk} = r^*_{jk}(s) = p_{jk} \int_0^\infty e^{-st} h_{jk}(t)dt, \tag{2}$$

i.e. the elements of R are the elements of P multiplied by the Laplace Transforms (LTs) of the corresponding sojourn density functions. If we denote the LT of $h_{jk}(t)$ by $\hat{h}_{jk}(s)$ then we have

$$r^*_{jk}(s) = p_{jk}\hat{h}_{jk}(s) \tag{3}$$

which is the LT of the kernel of the semi-Markov process.

We also introduce a matrix U, which is equivalent to R but with a set of *target states* $\mathcal{J}$ made absorbing: that is, rows corresponding to those states are replaced by rows of zeros:

$$[U]_{jk} = \begin{cases} 0 & j \in \mathcal{J} \\ r^*_{jk}(s) & j \notin \mathcal{J}. \end{cases} \tag{4}$$

We will use the notation δ^i to denote the vector whose elements are 0 except for element i, which is 1, and correspondingly introduce the vector $\delta^{\mathcal{J}}$ with components:

$$\delta^{\mathcal{J}}_j = \begin{cases} 1 \text{ for } j \in \mathcal{J} \\ 0 \text{ otherwise.} \end{cases} \tag{5}$$

Our ultimate aim is to prove convergence of the iterative scheme

$$L^n_i = [R(I + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}}]_i, \tag{6}$$

which is easy to implement numerically and in parallel [1], to the solution of the equation

$$L_i = \sum_{j \notin \mathcal{J}} R_{ij} L_j + \sum_{j \in \mathcal{J}} R_{ij}. \tag{7}$$

Each quantity L_i represents the value, for a specific value of the parameter s, of the Laplace transform of the distribution of the passage time of the system from state i to any state in $\mathcal{J}$, for reasons to be discussed in section 4.

3 Conjectures

In this section we state and prove several facts about our system, culminating in the convergence of L^n_i as defined above.

Lemma 1. *For any given value of s with nonnegative real part, the modulus of each complex number $\hat{h}_{jk}(s)$ is at most 1.*

Proof. Let $s = x + iy$. Then $x \geq 0$ and

$$\begin{aligned} \left|\hat{h}_{jk}(s)\right| &= \left|\int_{t=0}^\infty e^{-xt}(\cos yt - i \sin yt) h_{jk}(t)dt\right| \\ &\leq \int_{t=0}^\infty e^{-xt} h_{jk}(t)dt \leq \int_{t=0}^\infty h_{jk}(t)dt = 1. \end{aligned}$$

□

Corollary 1. *The spectral radius of the matrix U is at most 1.*

Proof. Since $|\hat{h}_{jk}| \leq 1$, it follows that $|r^*_{jk}| \leq p_{jk}$ for any j, k. Then the condition that $\sum_k p_{jk} = 1$ gives an upper bound on the infinity-norm of U:

$$\|U\|_\infty \leq \|R\|_\infty = \max_j \sum_k |r^*_{jk}| \leq 1$$

and thus the spectral radius of U is at most 1. ☐

Lemma 2. *Right-multiplication by U does not increase the L1-norm of any vector v:*

$$\|vU\|_1 \leq \|v\|_1 \quad \textit{and hence} \quad \|vU^n\|_1 \leq \|v\|_1 \quad \forall\, n \geq 0.$$

Proof.

$$\begin{aligned}
\|vU\|_1 &= \sum_k |(vU)_k| = \sum_k \left| \sum_j v_j U_{jk} \right| \\
&\leq \sum_k \sum_j |v_j|\,|U_{jk}| = \sum_k \sum_{j \notin \mathcal{J}} |v_j|\,|R_{jk}| \\
&\leq \sum_k \sum_{j \notin \mathcal{J}} |v_j|\, p_{jk} = \sum_{j \notin \mathcal{J}} |v_j| \sum_k p_{jk} \\
&= \sum_{j \notin \mathcal{J}} |v_j| \leq \sum_j |v_j| = \|v\|_1 . \qquad (8)
\end{aligned}$$

The extension to n multiplications by U is clear. ☐

Lemma 3. *Any eigenvalue λ of U which has modulus 1 is non-defective, that is, there is no vector w for which $w(U - \lambda I)^2 = 0$ but $w(U - \lambda I) \neq 0$.*

Proof. Assume the converse. Then let w be a principal vector of stage 1 associated with the eigenvalue in question. Then

$$w(U - \lambda I)^2 = 0 \text{ and } w(U - \lambda I) = v \neq 0.$$

It follows that v is an eigenvector corresponding to λ and

$$wU = \lambda w + v, \text{ and hence that } wU^n = \lambda^n w + n\lambda^{n-1} v.$$

Then

$$\|wU^n\| \geq n|\lambda|^{n-1}\|v\| - |\lambda|^n \|w\|,$$

and if $|\lambda| = 1$ then

$$\|wU^n\| \geq n\|v\| - \|w\|,$$

which grows without bound for large n since $\|v\|$ cannot be zero. However, lemma 2 states that $\|wU^n\|$ is bounded above by $\|w\|$ for any n, so we have a contradiction. ☐

Lemma 4. *The eigenvectors corresponding to all eigenvalues of modulus 1 have zero components corresponding to elements of $\mathcal{J}$.*

Proof. Consider the proof of lemma 2. For the norm of v to be maintained under multiplication with U (which must happen for an eigenvector whose eigenvalue has modulus 1), every inequality in the proof must be strict equality. In particular, from (8) we have

$$\sum_{j \notin \mathcal{J}} |v_j| = \sum_j |v_j|$$

which requires $v_j = 0$ if $j \in \mathcal{J}$. ☐

Lemma 5. *Let $\mathcal{V}_0$ be the vector subspace spanned by the eigenvectors of U corresponding to eigenvalues of modulus 1. Let $\mathcal{V}_1$ be the vector subspace spanned by the eigenvectors and principal vectors corresponding to eigenvalues of modulus strictly less than 1. Then $\mathbb{C}_N$ (where N is the dimension of U) is spanned by these two vector subspaces and they are invariant under right-multiplication by U.*

Proof. From lemma 1, all eigenvalues of U either have modulus 1 or modulus strictly less than 1. We know that $\mathbb{C}_N$ is spanned by the set of eigenvectors and principal vectors of U, and we can partition these vectors into those corresponding to eigenvalues of modulus 1 (which are all eigenvectors by lemma 3) and those corresponding to the other eigenvalues.

Consider an element of $\mathcal{V}_0$. It can be represented as a sum of eigenvectors x^l of U whose eigenvalues have modulus 1:

$$v_0 = \sum_l a^l x^l \text{ and } v_0 U = \sum_l a^l \lambda_l x^l \in \mathcal{V}_0$$

so $\mathcal{V}_0$ is invariant under right-multiplication by U.

Finally, consider an element of $\mathcal{V}_1$. Right-multiplication of each of its principal vector components by U generates principal vectors and eigenvectors corresponding to the same eigenvalue; and eigenvector components produce multiples of themselves so, by a similar argument to that used above, $\mathcal{V}_1$ is also invariant under right-multiplication by U. □

Lemma 6. *Split the matrix U into two components, U_0 and U_1, such that*

$$U = U_0 + U_1 \quad \textit{and} \quad v_0 U_0 = v_0 U \textit{ for all } v_0 \in \mathcal{V}_0, \quad v_1 U_0 = 0 \textit{ for all } v_1 \in \mathcal{V}_1. \tag{9}$$

Then

$$v_0 U_1 = 0 \textit{ for all } v_0 \in \mathcal{V}_0, v_1 U_1 = v_1 U \textit{ for all } v_1 \in \mathcal{V}_1, \tag{10}$$

and the matrix U_1 has spectral radius strictly less than 1.

Proof. Since $\mathcal{V}_0$ has as its basis the eigenvectors of U corresponding to eigenvalues of modulus 1, it is easy to construct U_0 to behave as above (9). Then the definition $U_1 = U - U_0$ leads trivially to (10).

Now consider any eigenvector of U_1. By the first part of (10), either it has eigenvalue 0 or it is a member of $\mathcal{V}_1$. In the latter case it is also an eigenvector of U, whose eigenvalue has modulus less than 1 from the definition of $\mathcal{V}_1$. It follows that U_1 has spectral radius strictly less than 1. □

Lemma 7. *For any element v_0 of $\mathcal{V}_0$, the scalar expression $v_0 U^n \delta^{\mathcal{J}}$ is zero for all $n \geq 0$.*

Proof. Since $\mathcal{V}_0$ is an invariant subspace under multiplication by U, we know that $v_0 U^n \in \mathcal{V}_0$. Any element of $\mathcal{V}_0$ can be decomposed into a sum of eigenvectors of U corresponding to eigenvalues of modulus 1, and by lemma 4 each eigenvector has zeros corresponding to nonzero elements of $\delta^{\mathcal{J}}$. □

Lemma 8. *The vector quantity $U_0 U^n \delta^{\mathcal{J}} = 0$ for all $n \geq 0$.*

Proof. Write the elements of the vector $U_0 U^n \delta^{\mathcal{J}}$ as $[U_0 U^n \delta^{\mathcal{J}}]_i = \delta^i U_0 U^n \delta^{\mathcal{J}}$. Then $\delta^i U_0 \in \mathcal{V}_0$ and the property follows from lemma 7. □

Theorem. *The limits as $n \to \infty$ of the sequences*

$$L_n^i = \delta^i R(I + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}}$$

are finite and satisfy

$$L_i = \sum_{j \notin \mathcal{J}} R_{ij} L_j + \sum_{j \in \mathcal{J}} R_{ij}.$$

Proof. Let us begin by defining an additional scalar quantity

$$T_n^i = [(I + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}}]_i = \delta^i (I + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}}.$$

We note that we can calculate L_n^i simply via

$$L_n^i = \sum_k R_{ik} T_n^k.$$

Now we split the vector δ^i into two components, one from each vector subspace. Thus

$$\delta^i = \delta^{i,0} + \delta^{i,1}$$

in the obvious notation, and we have

$$\begin{aligned}
T_n^i &= (\delta^{i,0} + \delta^{i,1})(I + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}} \\
&= \delta^{i,1}(I + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}} \text{ by lemma 7} \\
&= \delta^{i,1}(I + U_1 + U_1^2 + \ldots + U_1^n)\delta^{\mathcal{J}} \text{ by equations (9–10)} \\
&= \delta^{i}(I + U_1 + U_1^2 + \ldots + U_1^n)\delta^{\mathcal{J}} \text{ by lemma 7} \\
&= [(I + U_1 + U_1^2 + \ldots + U_1^n)\delta^{\mathcal{J}}]_i
\end{aligned}$$

and since U_1 has spectral radius strictly less than 1, by theorem 6.2.8 of [2] we have

$$\lim_{n\to\infty}(I + U_1 + U_1^2 + \ldots + U_1^n) = (I - U_1)^{-1}$$

so

$$T_i = \lim_{n\to\infty} T_n^i = [(I - U_1)^{-1}\delta^{\mathcal{J}}]_i.$$

We left-multiply the vector T by $(I - U_1)$ to obtain

$$\begin{aligned}
T &= \delta^{\mathcal{J}} + U_1 T \\
&= \delta^{\mathcal{J}} + UT - U_0 T
\end{aligned}$$

and we note that $U_0 T = \lim_{n\to\infty} U_0(1 + U + U^2 + \ldots + U^n)\delta^{\mathcal{J}} = 0$ by lemma 8 to obtain

$$\begin{aligned}
T &= \delta^{\mathcal{J}} + UT \\
T_j &= [\delta^{\mathcal{J}}]_j + \sum_k U_{jk} T_k
\end{aligned}$$

Finally we left-multiply by R to have

$$\begin{aligned}
L_i &= \sum_j R_{ij}[\delta^{\mathcal{J}}]_j + \sum_j \sum_k R_{ij} U_{jk} T_k \\
L_i &= \sum_{j\in\mathcal{J}} R_{ij} + \sum_{j\notin\mathcal{J}} \sum_k R_{ij} R_{jk} T_k \\
L_i &= \sum_{j\in\mathcal{J}} R_{ij} + \sum_{j\notin\mathcal{J}} R_{ij} L_j.
\end{aligned}$$

□

4 Physical Significance of the Results

We have shown that the iterative scheme of equation (6) converges to a finite quantity L_i satisfying (7):

$$L_i = \sum_{j \notin \mathcal{J}} R_{ij} L_j + \sum_{j \in \mathcal{J}} R_{ij}.$$

To understand the meaning of the quantity L_i, we first note the following rules for composition of Laplace transforms of distribution functions of random variables. In each case $\hat{f}$ is the LT of the distribution of a random variable (RV) $\mathcal{F}$, and $\hat{g}$ is the LT of the distribution of the RV $\mathcal{G}$.

- The LT of the distribution of the RV $\mathcal{F} + \mathcal{G}$ is $\hat{f}\hat{g}$.
- The LT of the distribution of the RV $\mathcal{H}$ which samples from $\mathcal{F}$ with probability p and from $\mathcal{G}$ with probability $1 - p$ is $p\hat{f} + (1 - p)\hat{g}$.
- The LT of the instantaneous distribution (that is, the RV whose value is always zero) is 1.

Now let us consider the random variable whose value is the time taken for the system to reach any state in $\mathcal{J}$, starting from state i at time $t = 0$. We will impose the additional constraint that if $i \in \mathcal{J}$ the system must make one transition before it is deemed to have reached the target: this convention is usual in passage-time definitions, in order to enable questions of the type "how long will it take to return to this state?". We will show that L_i is the LT of the distribution of this RV.

Now let us consider the sequence of states of a single instance of this system starting from state i. Because we have imposed the condition that at least one transition occurs, we know that the sequence of states has at least two elements, so we look at the initial transition out of state i. This will be to state j with probability p_{ij} for any j, and the LT of the distribution of time taken to make the transition is $\hat{h}_{ij}(s)$.

For the second transition, there are two possibilities: either j is in $\mathcal{J}$, in which case the sequence terminates and no more time is spent, or j is not in $\mathcal{J}$, in which case the time remaining is simply the first passage time from j to $\mathcal{J}$. This sequence also has the constraint that at least one transition must be made, since j is outside $\mathcal{J}$ so at least one transition is required to reach $\mathcal{J}$.

Expressing in equation form the choice described above gives

$$L_i = \sum_{j \notin \mathcal{J}} p_{ij} \hat{h}_{ij} L_j + \sum_{j \in \mathcal{J}} p_{ij} \hat{h}_{ij} 1$$

which, substituting the definition of the matrix R, is exactly the recurrence relation (7). It follows that L_i is the LT of the first passage time from state i to the target states $\mathcal{J}$, and the iterative scheme (6) always converges to this quantity.

Acknowledgement

The author would like to thank Dr. J. T. Bradley for helpful discussions.

References

[1] J T Bradley, N J Dingle, W J Knottenbelt, and H J Wilson. Hypergraph-based parallel computation of passage time densities in large semi-Markov models. *Linear Algebra and its Applications*, in press, 2003.

[2] R A Horn and C R Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1986.

[3] R Pyke. Markov renewal processes: Definitions and preliminary properties. *Annals of Mathematical Statistics*, 32:1231–1242, 1961.