SPECTRAL REPRESENTATIONS OF THE
TRANSITION PROBABILITY MATRICES FOR
CONTINUOUS TIME FINITE MARKOV CHAINS

NAN FU PENG,* National Chiao Tung University

Abstract

Using an easy linear-algebraic method, we obtain spectral representations, without the
need for eigenvector determination, of the transition probability matrices for completely
general continuous time Markov chains with finite state space. Comparing the proof
presented here with that of Brown (1991), who provided a similar result for a special
class of finite Markov chains, we observe that ours is more concise.

MARKOV CHAINS; TRANSITION PROBABILITY MATRICES; SPECTRAL REPRESENTATIONS

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60J35
SECONDARY 60J27

1. Introduction

It is undoubtedly important to calculate numerically the time-dependent transition
probabilities of continuous time Markov chains. We focus our attention on those with
a finite state space. Keilson developed in his book [5] the methods of spectral decompo-
sition and the uniformization technique. Ross [10] found the external uniformization; this
was followed by related work such as [7] and [12]. Some results on finite queues can be
of the transition probability matrices of finite continuous time Markov chains with
diagonalizable infinitesimal matrices (see also theorem 5 of [3]). Here we present an easy
linear-algebraic technique which enables us to extend the result of [2] to completely
general continuous time Markov chains with finite state space. The method used in this
paper is also more concise and efficient than that of [2].

2. A simple linear-algebraic method

Consider a Markov chain \((X(t))\) defined on a finite state space \(\{0, 1, 2, \ldots, N\}\). Denote
by \(\lambda_0 = 0, \lambda_1, \ldots, \lambda_N\) (maybe complex) the eigenvalues of its infinitesimal matrix \(Q\). It is
well known [5] that the transition probability matrix \(P(t)\) of \(X(t)\) is

Received 15 July 1994; revision received 19 December 1994.
* Postal address: Institute of Statistics, National Chiao Tung University, 1001 TA Hsueh Road, Hsinchu,
Taiwan.
Spectral representations of the transition probability matrices

\[ P(t) = e^{Qt} = \sum_{n=0}^{\infty} \frac{Q^n t^n}{n!}. \]

Obviously, (1) implies the following:

\[ P(0) = 1 \quad \text{and} \quad \left. \frac{d^n P(t)}{dt^n} \right|_{t=0} = \left( \frac{d^n P(t)}{dt^n} \right) = Q^n, \quad \forall n \geq 1. \]

If \( P(t) \) is a transition function or, more generally, sufficiently smooth, then (2) implies (1); hence we obtain the equivalence of (1) and (2). The linear algebra used below can be found in many textbooks, e.g. [4].

**Lemma 1.** Let \( A \) and \( B \) be two complex \( n \times n \) matrices and \( \{ \bar{x}_1, \ldots, \bar{x}_n \} \) be any basis of \( \mathbb{C}^n \). Then \( A\bar{x}_i = B\bar{x}_i \) for all \( i \) implies \( A = B \).

Although Theorem 1 is a special case of Theorem 3 below, it is worth listing the proof here for comparison with that of Theorem 3 and that of [2].

**Theorem 1.** If the \( \lambda_i \) are all distinct, then

\[ P(t) = \prod_{i=1}^{N} (I - Q/\lambda_i) \]

\[ + \sum_{m=1}^{N} (Q/\lambda_m) \prod_{i \neq m, 0} [(I - Q/\lambda_i)/(1 - \lambda_m/\lambda_i)] \exp(\lambda_m t). \]

**Proof.** Call the right-hand side of (3) \( \tilde{P}(t) \). It is easy to see that, for \( m = 0, 1, \ldots, N \),

\[ \left. \frac{d^n \tilde{P}(t)}{dt^n} \right|_{t=0} \bar{x}_m = \lambda_m^n \bar{x}_m = Q^n \bar{x}_m, \quad n = 0, 1, 2, \ldots \]

where \( \bar{x}_m \) is an eigenvector associated with the eigenvalue \( \lambda_m \). Since the \( \lambda_m \) are all distinct, the \( \bar{x}_m \) form a basis of \( \mathbb{C}^{N+1} \). The \( \tilde{P}(t) \) is obviously smooth, hence we obtain (3) from the fact that (2) implies (1) and Lemma 1.

The above proof gives us a natural extension of Theorem 1 to Theorem 2 below. We allow repeated eigenvalues here, and relabel them \( \lambda_0 = 0, \lambda_1, \ldots, \lambda_M \) as the distinct values.

**Theorem 2.** If the minimal polynomial of \( Q \) is of the form

\[ g(x) = x \prod_{i=1}^{M} (x - \lambda_i), \quad M \leq N, \]

with distinct \( \lambda_0 = 0, \lambda_1, \ldots, \lambda_M \), then \( P(t) \) is of the form (3) with \( N \) replaced by \( M \).

The next corollary also appeared in [2].

**Corollary 1.** If \( (X(t)) \) is a finite birth and death process, then \( P(t) \) is of the form (3).
Proof. The infinitesimal matrix $Q$ of $(X(t))$ is tridiagonal and it is shown in [2] that its eigenvalues are real and distinct.

The following example makes Theorem 1 more plausible.

Example 1. Consider a continuous time Markov chain having state space \{0, 1, 2, 3\} and starting from state 0 with infinitesimal matrix

$$
Q = \begin{bmatrix}
0 & 1 & 2 & 3 \\
0 & -\lambda & \lambda & 0 \\
2 & 0 & -\lambda & \lambda \\
3 & \lambda & 0 & -\lambda
\end{bmatrix}.
$$

A simple argument shows that

$$
P_{03}(t) = \sum_{n=1}^{\infty} P(T = 4n - 1) = e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(4n-1)!},
$$

where $T$ is a random variable distributed as Poisson $(\lambda t)$. In a similar fashion, we have

$$
P_{12}(t) = e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-3}}{(4n-3)!}.
$$

Alternatively, observing that $0, -2\lambda, -\lambda + i\lambda$ and $-\lambda - i\lambda$ are the eigenvalues of $Q$, we obtain from (3) that

$$
P_{03}(t) = e^{-\lambda t} \left[ \frac{1}{4} e^{\lambda t} - \frac{1}{4} e^{-\lambda t} - \frac{1}{2} \sin(\lambda t) \right]
$$

and

$$
P_{12}(t) = e^{-\lambda t} \left[ \frac{1}{4} e^{\lambda t} - \frac{1}{4} e^{-\lambda t} + \frac{1}{2} \sin(\lambda t) \right].
$$

By introducing the Taylor expansions of the terms in the brackets of the right-hand sides of (6) and (7), we obtain the respective equivalence of (6) and (7) to (4) and (5).

3. The general result

A matrix $Q$ is defined to be lower semitriangular if $Q_{ij} = 0$ for $j > i + 1$. It was claimed in Theorem 1.2 of [6] that, if $Q$ is lower semitriangular with $Q_{i,i+1} \neq 0$ for all $i$, then its eigenvalues are distinct but may be complex. This statement is incorrect as the next simple counterexample shows.

Example 2. Let the matrix $Q$ be

$$
Q = \begin{bmatrix}
0 & 1 & 2 \\
0 & -1 & 1 & 0 \\
2 & 1 & 0 & -1
\end{bmatrix}.
$$

The eigenvalues of $Q$ are 0, $-2$ and $-2$. Neither Theorem 1 nor Theorem 2 can be applied to this case because the null space of $Q + 2I$ is of dimension 1. Theorem 3 below...
Spectral representations of the transition probability matrices

deals with general $Q$ and provides us with a way to settle the problem. Several lemmas are needed in order to prove that theorem.

**Lemma 2.**

\[
\frac{d^n(t^k e^{\lambda t})}{dt^n}\bigg|_{t=0} = \begin{cases} 
0 & n < k \\
k! & n = k \\
\binom{n}{k}k!\lambda^{n-k} & n > k.
\end{cases}
\]

**Proof.** By the product rule of derivatives, it is easy to see that if $f(t)$ and $g(t)$ are continuously differentiable functions,

\[
(fg)^{(n)} = \sum_{i=0}^{n} \binom{n}{i} f^{(i)}g^{(n-i)}.
\]

We immediately obtain the lemma by letting $f(t) = t^k$ and $g(t) = e^{\lambda t}$.

**Lemma 3.** For given $M \geq 1$, let

\[
f(t) = \left[\prod_{m=1}^{M-1} \left(\frac{t}{a_m} + 1\right)^{d_m}\right] \left(1 + \sum_{i=1}^{k} c_i t^i\right)
\]

where the $d_m$ are non-negative integers and $a_m \neq 0$ for $m = 1, \cdots, M - 1$. Then $f^{(n)}(0) = 0$, $n = 1, 2, \cdots, K$, if and only if the $c_n$ satisfy

\[
-c_n = \sum_{\sum_{m=0}^{M} d_m \leq \pi} \left[\prod_{m=1}^{M-1} \frac{i_m}{a_m^{d_m}}\right] c_{n-i_1-\cdots-i_{M-1}}, \quad n = 1, 2, \cdots, K,
\]

with the conventions that $c_0 = 1$ and the right-hand side of (9) is zero when $M = 1$.

**Proof.** A quick application of (8) shows, for $M = 2, 3, \cdots$ and any $f_1, \cdots, f_M$,

\[
\frac{d^n}{dt^n} \left[\prod_{i=1}^{M} f_i(t)\right] = \sum_{0 \leq i_1 + \cdots + i_{M-1} \leq n} \binom{n}{i_1i_2\cdots i_M} \left(\prod_{m=1}^{M-1} f_m^{(i_m)}(t)\right) f_M^{n-i_1-\cdots-i_{M-1}}(t),
\]

with $i_M = n - i_1 - \cdots - i_{M-1}$ here. Hence for $n = 1, 2, \cdots, K$,...
\[
\frac{d^n f(t)}{dt^n} \bigg|_{t=0} = \sum_{0 \leq i_1 + \cdots + i_{M-1} \leq n} \frac{n!}{i_1! \cdots i_{M-1}!(n-i_1-\cdots-i_{M-1})!} \\
\times \left( \prod_{m=1}^{M-1} \left( \frac{d_{m!}}{(d_m-i_m)!a_{m^n}} \right) (n-i_1-\cdots-i_{M-1})! c_{n-i_1-\cdots-i_{M-1}} \right)
\]

\[
= n! \sum_{0 \leq i_1 + \cdots + i_{M-1} \leq n} \left( \prod_{m=1}^{M-1} \frac{\binom{d_m}{i_m}}{a_{m^n}} \right) c_{n-i_1-\cdots-i_{M-1}} = 0
\]

if and only if (9) holds.

**Theorem 3.** Let the minimal polynomial of \(Q\) be of the form \(f(x) = \prod_{i=0}^M (x-\lambda_i)^{d_i}\) where the \(\lambda_i\) are distinct and \(d_i \geq 1\). Then

\[
P(t) = \sum_{i=0}^{M} \left( \sum_{j=0}^{d_i-1} \frac{R_{(i,j)}}{j!} (Q-\lambda_i I)^j \right) e^{\lambda_i t}
\]

where

\[
R_{(i,j)} = \left( \prod_{m \neq i} \frac{(Q-\lambda_m I)^{d_m}}{(\lambda_i-\lambda_m)^{d_m}} \right) I + \sum_{n=1}^{d_i-1} c_{i,n} (Q-\lambda_i I)^n
\]

and

\[
-c_{i,n} = \sum_{k_m \leq m \neq i, i_m \neq 0} \frac{\prod_{m \neq i} \binom{d_m}{k_m}}{(\lambda_i-\lambda_m)^{k_m}} c_{i,n-k_m,i_m}, \quad 1 \leq n \leq d_i - 1,
\]

with \(c_{i,0} = 1\).

**Remark.** It is easy to check that Theorem 3 reduces to Theorem 2 when \(d_i = 1\) for \(i = 0, \cdots, M\).

**Proof.** Call the right-hand side of (14) \(\hat{P}(t)\). Due to the fact \((Q-\lambda_m I) = (Q-\lambda_i I) + (\lambda_i-\lambda_m)I\) and Lemma 3, \(R_{(i,j)}(Q-\lambda_i I)^j\) for \(0 \leq j < d_i\) can be written as

\[
R_{(i,j)}(Q-\lambda_i I)^j = w_\beta (Q-\lambda_i I)^\beta + \cdots + w_d (Q-\lambda_i I)^d + (Q-\lambda_i I)^j
\]

where the \(w\) are complex scalars depending on \(i\) and \(\beta = \Sigma d_m - 1\).

With some algebra, Lemma 2 together with (10), (11) and (12) yield the following:

\[
\hat{P}(0) = I \xi_i, \quad \text{and}
\]

\[
\frac{d^n \hat{P}(t)}{dt^n} \bigg|_{t=0} = \sum_{m=0}^{n} \binom{n}{m} (Q-\lambda_i I)^m (\lambda_i I)^{n-m} \xi_i
\]

\[
= (Q-\lambda_i I + \lambda_i I)^n \xi_i = Q^n \xi_i
\]
Spectral representations of the transition probability matrices

where \( x_i \) belongs to the null space of \((Q - \lambda I)^d\). Note that \((Q - \lambda I)^m x_i = 0\) if \( m \geq d \). Since these \( x_i \) form a basis for \( C^{N+1} \) and \( P(t) \) is sufficiently smooth, Lemma 1 and the implication of (2) to (1) yield the desired result.

**Remark.** Supposing the minimal polynomial is difficult to obtain, Theorem 3 still holds if we replace it with the characteristic polynomial.

**Corollary 2.** If \((X(t))\) is ergodic, then \( \pi' = (1/(N+1)) \bar{I} \cdot \Pi_{i=1}^{M} (I - Q/\lambda_i)^d \) is the unique stationary vector of \( P(t) \), where \( \bar{I} \) is the vector with all entries equal to 1.

**Proof.** Since \( 0 \leq P_j(t) \leq 1 \), the real part of each \( \lambda_k (k \neq 0) \) is strictly negative and \( d_0 = 1 \). Hence
\[
P(t) \rightarrow \Pi_{i=1}^{M} (I - Q/\lambda_i)^d \text{ as } t \rightarrow \infty.
\]
Since \((X(t))\) is ergodic, each row of \( \Pi_{i=1}^{M} (I - Q/\lambda_i)^d \) is the unique stationary vector \( \pi' \).

Note that irreducibility of \((X(t))\) implies ergodicity of \((X(t))\) [5].

**Example 2 (Continued).** The probability transition matrix \( P(t) \) corresponding to the infinitesimal matrix \( Q \) is
\[
P(t) = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/4 \end{pmatrix} + \begin{pmatrix} 1/2 & -1/4 & -1/4 \\ -1/2 & 3/4 & -1/4 \\ -1/2 & -1/4 & 3/4 \end{pmatrix} e^{-2t} + \begin{pmatrix} 0 & 1/2 & -1/2 \\ 0 & -1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix} te^{-2t}.
\]

**Acknowledgment**

I am very thankful to the referee for many helpful comments.

**References**