March 2020

The Semimartingale Dynamics and Generator of a Continuous Time Semi-Markov Chain

Robert J. Elliott

*University of Calgary, Calgary, AB, Canada, relliott@ucalgary.ca*

Follow this and additional works at: [https://digitalcommons.lsu.edu/josa](https://digitalcommons.lsu.edu/josa)

Part of the [Analysis Commons](https://digitalcommons.lsu.edu/analysis-commons), and the [Other Mathematics Commons](https://digitalcommons.lsu.edu/other-mathematics-commons)

**Recommended Citation**


DOI: 10.31390/josa.1.1.01

Available at: [https://digitalcommons.lsu.edu/josa/vol1/iss1/1](https://digitalcommons.lsu.edu/josa/vol1/iss1/1)
THE SEMIMARTINGALE DYNAMICS AND GENERATOR OF A CONTINUOUS TIME SEMI-MARKOV CHAIN

ROBERT J. ELLIOTT*

Abstract. We consider a finite state, continuous time homogeneous semi-Markov chain $X = \{X_t, t \geq 0\}$. Without loss of generality the state space of the chain can be identified with the set of unit vectors $S = \{e_1, e_2, \ldots, e_N\}$ where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)' \in \mathbb{R}^N$. The probabilistic and dynamic properties of $X$ can be described by either a rate matrix $A$ or a matrix which gives the occupation times in the various states together with the probabilities of jumping to a different state. For a continuous time Markov chain the occupation times are memoryless, implying the distributions are exponential. For semi-Markov chains the occupation times can have more general distributions. The relation between these two descriptions is first investigated and the semimartingale dynamics of a semi-Markov chain obtained in contrast to the traditional description of a semi-Markov chain in terms of a renewal process.

An equation giving the dynamics of the occupation times is derived together with an equation for the density of the conditional occupation time and state. Some approximations for these dynamics are then obtained.

1. Introduction

This paper derives new results and representations for a continuous time, finite state, homogeneous semi-Markov chain $X = \{X_t, t \geq 0\}$. Without loss of generality the state space of the chain can be taken to be the set $S = \{e_1, e_2, \ldots, e_N\}$ of unit vectors in $\mathbb{R}^N$.

For a continuous time Markov chain the amount of time the chain spends in any state has an exponential distribution. This condition is relaxed for a semi-Markov chain; these occupation times may have more general distributions.

Discrete time Markov chains have occupation times which have a geometric distribution, a condition which is relaxed for discrete time semi-Markov chains as discussed in our recent paper [3]. Semi-Markov chains are discussed in [5]. This paper provides the semimartingale dynamics of a semi-Markov chain in contrast to the description in terms of a renewal process.

In many real world applications it appears that semi-Markov processes might be more appropriate models. However, their analysis is more technical and involved. We hope this paper contributes to their analysis.

Received 2019-7-17; Accepted 2019-9-23; Communicated by the editors.
2010 Mathematics Subject Classification. Primary 60K15; Secondary 90C40.
Key words and phrases. Semi-Markov chain, general occupation times, generator, semimartingale matrix dynamics.
* Corresponding author.
In Section 2 we first relate the dynamics and probabilistic properties of a finite state Markov chain and semi-Markov chain when they are described by a rate matrix or by a matrix giving the distribution of occupation times in the various states.

Dynamics for the occupation times are obtained in Section 3. In Section 4 the equation for the density of the occupation time is derived.

2. Dynamics

In this section we compare the dynamics of a continuous time Markov chain with those of a continuous time semi-Markov chain.

Our processes are defined on a probability space \((\Omega, \mathcal{F}, P)\) and time \(t \in [0, \infty)\). The state space is identified with a set of unit vectors \(S = \{e_1, e_2, \ldots, e_N\}\),

\[ e_i = (0, 0, \ldots, 1, 0, \ldots, 0)' \in \mathbb{R}^N. \]

2.1. Markov chain. We first consider a continuous time homogeneous Markov chain \(X = \{X_t, t \geq 0\}\) with \(X_t \in S\).

Write

\[ p_t = E[X_t|X_0] = (p_1^t, p_2^t, \ldots, p_N^t)' \in \mathbb{R}^N. \]

The probabilistic properties of \(X\) are usually described by a ‘rate’ matrix \(A = (a_{ji}, 1 \leq i, j \leq N)\) such that

\[ \frac{dp_t}{dt} = Ap_t, \]

with \(p_0 = X_0\) given. Then \(p_t = E[X_t|X_0] = e^{At}X_0 \in \mathbb{R}^N\) and more generally for \(0 \leq s \leq t\)

\[ E[X_t|X_s] = e^{A(t-s)}X_s. \]

Define the vector process \(M := \{M_t, t \geq 0\}\) by

\[ M_t = X_t - X_0 - \int_0^t AX_u du \in \mathbb{R}^N. \]

With \(\mathcal{F}_t := \sigma\{X_u : u \leq t\}\) and \(0 \leq s \leq t\)

\[ E[M_t - M_s | \mathcal{F}_s] = E[X_t - X_s - \int_s^t AX_u du | X_s] \]

\[ = e^{A(t-s)}X_s - X_s - \int_s^t AE[X_u | X_s] du \]

\[ = e^{A(t-s)}X_s - X_s - \int_s^t Ae^{A(u-s)}X_s du \]

\[ = 0 \in \mathbb{R}^N \]

so \(M\) is an \(\{\mathcal{F}_t\}\) vector martingale. Consequently the semimartingale representation of \(X\) is, as in [4],

\[ X_t = X_0 + \int_0^t AX_u du + M_t \in \mathbb{R}^N. \] (2.1)
For \( i \neq j \), \( a_{ji} \geq 0 \) is the rate at which \( X \) jumps from \( e_i \) to \( e_j \). Also,

\[
a_{ii} < 0 \quad \text{and} \quad \sum_{j=1}^{N} a_{ji} = 0.
\]

The occupation times which describe how long the Markov chain \( X \) remains in any state are exponentially distributed. See [1]. Write

\[
0 < \tau_1 < \tau_2 < \ldots
\]

for the jump times of \( X \). If the occupation time in state \( e_i \) is exponentially distributed with parameter \( \lambda_i \) then

\[
P(\tau_n < \tau_{n+1} \leq t | X_{\tau_n} = e_i) = 1 - e^{-\lambda_i(t-\tau_n)}. \tag{2.2}
\]

If, for \( j \neq i \), the probability of jumping from \( e_i \) to \( e_j \), given there is a jump, is \( p_{ji} \) then

\[
P(X_{\tau_{n+1}} = e_j, \tau_n \leq \tau_{n+1} \leq t | X_{\tau_n} = e_i) = p_{ji} \left(1 - e^{-\lambda_i(t-\tau_n)}\right). \tag{2.3}
\]

Also,

\[
\sum_{j=1}^{N} p_{ji} = 1.
\]

The rate of jumping from \( e_i \) to \( e_j \) at time \( t \), given \( t \leq \tau_{n+1} \) and \( X_{\tau_n} = e_i \), is then \( p_{ji} \lambda_i \).

For the Markov chain \( X \) this rate is memoryless, that is, it does not depend on \( \tau_n \) or \( t \).

However, from the previous parameterization of \( X \) the rate of jumping from \( e_i \) to \( e_j \) is \( a_{ji} \), \( i \neq j \). Therefore, for \( i \neq j \)

\[
a_{ji} = p_{ji} \lambda_i.
\]

Now the \( p_{ji} \) only refer to non-zero jumps, that is \( i \neq j \), and

\[
\sum_{j=1}^{N} p_{ji} = 1.
\]

So

\[
\sum_{j=1}^{N} p_{ji} \lambda_i = \lambda_i = \sum_{j=1}^{N} a_{ji} = -a_{ii}.
\]

\( \lambda_i \) is the rate at which the chain jumps from \( e_i \), (to any other state \( e_j \), \( i \neq j \)).

The distribution describing how long the chain stays in state \( e_i \) and does not jump is

\[
P(\tau_n \leq t < \tau_{n+1} | X_{\tau_n} = e_i) = e^{-\lambda_i(t-\tau_n)}.
\]

Suppose \( N = 3 \). We have shown that

\[
A = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & 33
\end{pmatrix} = \begin{pmatrix}
-\lambda_1 & p_{12} \lambda_2 & p_{13} \lambda_3 \\
p_{21} \lambda_1 & -\lambda_2 & p_{23} \lambda_3 \\
p_{31} \lambda_1 & p_{32} \lambda_2 & -\lambda_3
\end{pmatrix}.
\]
4 ROBERT J. ELLIOTT

With

\[ \Pi := \begin{pmatrix} -1 & p_{12} & p_{13} \\ p_{21} & -1 & p_{23} \\ p_{31} & p_{32} & -1 \end{pmatrix} \]

\[ D := \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \]

we see

\[ A = \Pi D. \quad (2.4) \]

2.2. Semi-Markov chain. For a semi-Markov chain \( X = \{X_t, t \geq 0\} \) the occupation times in any state \( e_i \) have more general distributions. Suppose for \( 1 \leq i \leq N \) there is a probability distribution \( G_i \) such that

\[ P(\tau_n \leq t < \tau_{n+1} | X_{\tau_n} = e_i) = G_i(t - \tau_n). \]

Note that \( G_i(t - \tau_n) \) is decreasing in \( t \), or at least non-increasing, so assuming \( G_i \) is differentiable with derivative \( g_i \), \( g_i(t - \tau_n) \leq 0 \). Then

\[ P(\tau_n < \tau_{n+1} \leq t | X_{\tau_n} = e_i) = 1 - G_i(t - \tau_n). \]

Suppose

\[ p_{ji} = P(X_{\tau_{n+1}} = e_j | X_{\tau_n} = e_i) \quad \text{for} \quad i \neq j. \]

We consider only non-zero jumps so

\[ \sum_{j=1}^{N} p_{ji} = 1. \]

Then for \( i \neq j \):

\[ P(X_{\tau_{n+1}} = e_j, \tau_n < \tau_{n+1} \leq t | X_{\tau_n} = e_i) = p_{ji} (1 - G_i(t - \tau_n)). \quad (2.5) \]

We consider the analog of the rate matrix \( A = (a_{ij}, 1 \leq i, j \leq N) \) for this semi-Markov chain. From (2.5) the rate of jumping from \( e_i = X_{\tau_n} \) to \( e_j = X_{\tau_{n+1}}, i \neq j \), given \( X_t = X_{\tau_n} = e_i \) and \( \tau_{n+1} > t \) is

\[ \frac{\partial}{\partial t} \left( \frac{p_{ji} G_i(t - \tau_n)}{G_i(t - \tau_n)} \right) = -\frac{p_{ji} g_i(t - \tau_n)}{G_i(t - \tau_n)}. \]

That is, the off-diagonal elements of a rate matrix for a semi-Markov chain have the form

\[ a_{ji}(t - \tau_n) = \frac{-p_{ji} g_i(t - \tau_n)}{G_i(t - \tau_n)}. \]

Now

\[ \sum_{j=1}^{N} p_{ji} = 1 \]
so, as in the Markov case,

\[ a_{ii}(t - \tau_n) = \sum_{j=1}^{N} \sum_{i \neq j} p_{ji} \frac{g_i(t - \tau_n)}{G_i(t - \tau_n)} \]

This is the infinitesimal rate of there not being a jump at \( t \), given \( X_t = e_i \) and \( \tau_{n+1} > t \).

The matrix \( A(t - \tau_n) = (a_{ji}(t - \tau_n), 1 \leq i, j \leq N) \) is the analog for the semi-Markov chain of the rate matrix for the Markov chain. Unlike the Markov case the entries now depend on \( (t - \tau_n) \), the time since the last jump.

The semimartingale dynamics for the semi-Markov chain are

\[ X_t = X_{\tau_n} + \int_{\tau_n}^{\tau_{n+1} \land t} A(u - \tau_n)X_u du + M_t \in R^N. \]  

In the book [2] we show that if \( \tau \) is a jump time with \( F_t = P(\tau > t) \) then

\[ I_{t \geq \tau} + \int_0^{t \land \tau} \frac{dF_u}{F_{u-}} \]

is a martingale. The martingale property of \( M_t \in R^N \) is a consequence of this result for each component.

### 3. Occupation Times

Consider a finite state, homogeneous semi-Markov chain \( X = \{X_t, t \geq 0\} \) with the semimartingale representation as above:

\[ X_t = X_{\tau_n} + \int_{\tau_n}^{\tau_{n+1} \land t} A(u - \tau_n)X_u du + M_t \in R^N. \]

For \( 1 \leq i \leq N \) define the process \( h^i_t, t \geq 0, \) by

\[ h^i_t = \int_0^t \langle X_u, e_i \rangle du + \int_0^t h^i_{u-} \langle X_u-, e_i \rangle \langle e_i, dX_u \rangle. \]

Suppose initially the chain \( X \) is in state \( e_i \). The second integral only contributes when \( X \) jumps from \( e_i \) to another state \( e_j \) with \( i \neq j \). At the first jump time it contributes the amount

\[ -h^i_{\tau_1-} \langle X_{\tau_1-}, e_i \rangle. \]

As \( X_{\tau_1-} = e_i \) this is \( -h^i_{\tau_1-} \).

However, the value of the first integral at \( \tau_1- \) is

\[ h^i_{\tau_1-} = \int_0^{\tau_1} \langle X_u, e_i \rangle du = \tau_1. \]
Consequently,
\[ h_{\tau_1}^i = h_{\tau_1^-}^i - h_{\tau_1^-}^i \langle X_{\tau_1^-}, e_i \rangle = 0. \]

A similar argument shows that at all subsequent jump times \( \tau_k \), \( h_{\tau_k}^i = 0 \).

Therefore, \( h_{\tau_k}^i \) is re-set to 0 at each jump time \( \tau_k \) so \( h_{\tau_k}^i \) measures the amount of time the chain has spent in state \( e_i \) since it last jumped to \( e_i \). That is, \( h_{\tau_k}^i = (t - \tau_k) \), given \( X_{\tau_k} = e_i \) and \( \tau_k \leq t < \tau_{k+1} \). Consider
\[ h_t := \sum_{i=1}^{N} h_t^i. \] (3.3)

Then
\[ h_t = t + \int_0^t h_{u^-} \langle X_{u^-}, dX_u \rangle. \] (3.4)

A similar argument shows that \( h_t = t - \tau_k \) where \( \tau_k \leq t < \tau_{k+1} \).

Consequently, we can write
\[ a_{ji}(t - \tau_k) = a_{ji}(h_t^i) = -p_{ji} \frac{g_i(h_t^i)}{G_i(h_t^i)}, \quad i \neq j, \]
\[ a_{ii}(t - \tau_k) = a_{ii}(h_t^i) = \frac{g_i(h_t^i)}{G_i(h_t^i)}, \]
and
\[ A(h_t) = \left( a_{ji}(h_t^i) \right). \]

The semimartingale dynamics of \( X \) are then
\[ X_t = X_0 + \int_0^t A(h_u)X_u du + M_t \in R^N. \] (3.5)

### 4. Semi-Markov Chain Generator

As above, \( X = \{X_t, t \geq 0\} \) is a finite state continuous time homogeneous semi-Markov chain defined on the probability space \((\Omega, \mathcal{F}, P)\). The state space is the set \( S = \{e_1, e_2, \ldots, e_N\} \) of unit vectors in \( R^N \). \( h_t \) is the occupation time in the current state since the last jump. We have established the following dynamics for \( X \) and \( h \):
\[ X_t = X_0 + \int_0^t A(h_u)X_u du + M_t \in R^N \] (4.1)
\[ h_t = t + \int_0^t h_{u^-} \langle X_{u^-}, dX_u \rangle \in R. \] (4.2)

From (4.1) and (4.2) we see the process \((X_t, h_t)\) is Markov.

We shall use the following differentiation rules for semimartingales with no continuous martingale part. The proofs are in [2] Chapter 14. If \( Z = \{Z_t, t \geq 0\} \) is a scalar semimartingale with no continuous martingale component and \( F \) a differentiable function
\[ F(Z_t) = F(Z_0) + \int_0^t F'(Z_{u^-})dZ_u + \sum_{0 < u \leq t} [F(Z_u) - F(Z_{u^-}) - F'(Z_{u^-}) \Delta Z_u]. \] (4.3)
For two such semimartingales:

\[ Y_t Z_t = Y_0 Z_0 + \int_0^t Y_u - dZ_u + \int_0^t Z_u - dY_u + [Y, Z]_t \]  

(4.4)

where

\[ [Y, Z]_t = \sum_{0 < u \leq t} \Delta Y_u \Delta Z_u. \]

Recall from Section (2.2) that

\[ A(h_t) = (a_{ji}, 1 \leq i, j \leq N) \]

where for \( i \neq j \)

\[ a_{ji}(t - \tau_k) = a_{ji}(h_t) = -\frac{p_{ji} g_i(h_t)}{G_i(h_t)} \]

and for \( i = j \)

\[ a_{ii}(t - \tau_k) = a_{ii}(h_t) = \frac{g_i(h_t)}{G_i(h_t)}. \]

Recall the state space of \( X \) is the set

\[ S = \{e_1, e_2, \ldots, e_N\} \]

of unit vectors in \( \mathbb{R}^N \).

Suppose \( \phi : S \times \mathbb{R} \to \mathbb{R} \) is an arbitrary \( C^1 \) function. Then for \( \Delta \geq 0 \):

\[
\phi(X_{t+\Delta}, h_{t+\Delta}) - \phi(X_t, h_t) = \int_t^{t+\Delta} \phi'(X_u, h_u) du + \sum_{t < u \leq t+\Delta} (\phi(X_u, h_u) - \phi(X_{u-}, h_{u-})).
\]

(4.5)

Now

\[
\sum_{i,j=1 \atop i \neq j}^N \left( \phi(X_u, h_u) - \phi(X_{u-}, h_{u-}) \right) = \sum_{i,j=1 \atop i \neq j}^N \left( \phi(X_u, h_u) - \phi(X_{u-}, h_{u-}) \right) (X_u, e_j)(X_{u-}, e_i).
\]

For \( i \neq j \)

\[
(X_u, e_j)(X_{u-}, e_i) = (X_u - X_{u-}, e_j)(X_{u-}, e_i) \]

so

\[
(\phi(X_u, h_u) - \phi(X_{u-}, h_{u-})) (X_u, e_j)(X_{u-}, e_i)
\]

\[
= (\phi(e_j, 0) - \phi(e_i, h_{u-}) \{\Delta X_u, e_j\}(X_{u-}, e_i)
\]

\[
= (\phi(e_j, 0) - \phi(e_i, h_{u-}) \{dX_u, e_j\}(X_{u-}, e_i).
\]

From (3.5)

\[ dX_u = A(h_u)X_u du + dM_u \]

where \( M \) is a martingale so this is

\[ = (\phi(e_j, 0) - \phi(e_i, h_{u-}) \{A(h_u)X_u du + dM_u, e_j\}(X_{u-}, e_i). \]
Therefore
\[
\sum_{i,j=1 \atop i \neq j}^{N} \left( \phi(X_u, h_{u-}) - \phi(X_{u-}, h_{u-}) \right) \langle X_u, e_j \rangle \langle X_{u-}, e_i \rangle
\]
\[
= \sum_{i,j=1 \atop i \neq j}^{N} \left( \phi(e_j, 0) - \phi(e_i, h_{u-}) \right) \langle X_{u-}, e_i \rangle \langle A(h_u)X_u, e_j \rangle du
\]
\[
\quad \quad + \text{a martingale increment}
\]
\[
= \sum_{i,j=1 \atop i \neq j}^{N} \left( \phi(e_j, 0) - \phi(e_i, h_{u-}) \right) \langle X_u, e_i \rangle a_{ji}(h_u) du
\]
\[
\quad \quad + \text{a martingale increment.}
\]

Consequently, assuming \(X_t = e_i\),
\[
\phi(X_{t+\Delta}, h_{t+\Delta}) - \phi(X_t, h_t)
\]
\[
= \int_t^{t+\Delta} \phi'(X_u, h_u) du + \sum_{i,j=1 \atop i \neq j}^{N} \int_t^{t+\Delta} \left( \phi(e_j, 0) - \phi(e_i, h_{u-}) \right) \langle X_{u-}, e_i \rangle a_{ji}(h_u) du
\]
\[
\quad \quad + \text{a martingale term.}
\]

From this we immediately deduce:

**Theorem 4.1.** The generator of the Markov process \((X_t, h_t)\) is
\[
\lim_{\Delta \to 0} \frac{1}{\Delta} E \left[ \phi(X_{t+\Delta}, h_{t+\Delta}) - \phi(X_t, h_t) | X_t, h_t \right]
\]
\[
= \phi'(X_t, h_t) + \sum_{j=1 \atop j \neq i}^{N} \left( \phi(e_j, 0) - \phi(X_t, h_t) a_{ji}(h_t) \right). \tag{4.6}
\]

5. Density Dynamics

**Notation 5.1.** Write
\[
q_s^i(y) dy = P \{ h_s \in [y, y + dy] \text{ and } X_s = e_i | X_0 \} \tag{5.1}
\]
and
\[
q_s(y) = (q_s^1(y), q_s^2(y), \ldots, q_s^N(y))' \in \mathbb{R}^N. \tag{5.2}
\]

**Remark 5.2.** Suppose \(0 < \tau_1 < \tau_2 < \cdots < \tau_n < \cdots\) are the jump times of \(X\). Write \(\{ \tau_k^i \}\) for the jump times when \(X_{\tau_k^i} = e_i\). Then
\[
\int_0^t f'(h_{s-}) \langle X_{s-}, e_i \rangle ds = \sum_{\tau_k^i \leq t} \int_{\tau_k^i}^{\tau_k^i + \Delta} f'(s - \tau_k^i) ds
\]

\[
\int_0^t f'(h_{s-}) \langle X_{s-}, e_i \rangle ds = \sum_{\tau_k^i \leq t} \int_{\tau_k^i}^{\tau_k^i + \Delta} f'(s - \tau_k^i) ds
\]
where \( \tau_{k'} \) is the next jump time after \( \tau_k \)
\[
= \sum_{\tau_k^* \leq t} (f(h_{\tau_k^*}) - f(h_{\tau_k})).
\]

If \( \tau_k^* \leq t < \tau_{k'} \) this term is
\[
(f(t - \tau_k^*) - f(0)).
\]

If \( \tau_k^* < \tau_{k'} \leq t \) this term is
\[
(f(\tau_{k'} - \tau_k^*) - f(0)).
\]

Also, if \( s \) is a jump time, \( \tau_n \), then \( h_s = 0 \) so
\[
f(h_s) - f(h_{s-}) = f(0) - f(h_{s-}).
\]

**Theorem 5.3.** The density \( q_i \) satisfies the dynamics
\[
q_i(y) = \delta_0(y) - \int_0^\infty q_i(y)dy - \int_0^t \frac{\partial q_i(y)}{\partial y} ds
\]
\[
- \delta_0(y) \int_0^t (\int_0^\infty a_{ii}(y)q_i(y)dy)ds + \int_0^t a_{ii}(y)q_i(y)ds.
\]

**Proof.** Suppose \( f \) is an arbitrary differentiable function with compact support. Then using (4.3) and (4.4):
\[
f(h_t) \langle X_t, e_i \rangle = f(0) \langle X_0, e_i \rangle + \int_0^t f(h_s) \langle A(h_s)X_s, e_i \rangle ds
\]
\[
+ \int_0^t f(h_{s-}) \langle dM_s, e_i \rangle + \int_0^t f'(h_{s-}) \langle X_{s-}, e_i \rangle ds
\]
\[
+ \sum_{0<s \leq t} [f(h_s) - f(h_{s-})] \langle -X_{s-}, \Delta X_s \rangle \langle X_{s-}, e_i \rangle
\]
\[
+ \sum_{0<s \leq t} [f(h_s) - f(h_{s-})] \langle -X_{s-}, \Delta X_s \rangle \langle \Delta X_s, e_i \rangle.
\]

Now
\[
\langle -X_{s-}, \Delta X_s \rangle \langle \Delta X_s, e_i \rangle = \langle \Delta X_s, e_i \rangle
\]
so the last sum in (5.4) is:
\[
\sum_{0<s \leq t} [f(h_s) - f(h_{s-})] \langle \Delta X_s, e_i \rangle.
\]

If there is a jump at \( s \) \( h_s = 0 \) so this is
\[
\int_0^t [f(0) - f(h_{s-})] \langle dX_s, e_i \rangle
\]
as \( X \) is piecewise constant. From (3.5) this is
\[
f(0) \langle X_t - X_0, e_i \rangle - \int_0^t f(h_{s-}) \langle A(h_s)X_s ds + dM_s, e_i \rangle.
\]
The integral here cancels the first two integrals in (5.4) so:

\[
f(h_t)\langle X_t, e_i \rangle = f(0)\langle X_t, e_i \rangle + \int_0^t f'(h_s) \langle X_s, e_i \rangle ds
\]

\[
+ \int_0^t [f(0) - f(h_s-)] \langle X_s, e_i \rangle \langle -e_i, dX_s \rangle.
\]

The final integral here is

\[
- \int_0^t [f(0) - f(h_s-)] \langle X_{s-}, e_i \rangle \langle A(h_s)X_s ds + dM_s, e_i \rangle
\]

\[
= -f(0) \int_0^t a_{ii}(h_s) \langle X_s, e_i \rangle ds + \int_0^t f(h_s) a_{ii}(h_s) \langle X_s, e_i \rangle ds
\]

\[
+ \text{Martingale.}
\]

Consequently,

\[
f(h_t) \langle X_t, e_i \rangle = f(0) \langle X_t, e_i \rangle + \int_0^t f'(h_s) \langle X_s, e_i \rangle ds
\]

\[
- f(0) \int_0^t a_{ii}(h_s) \langle X_s, e_i \rangle ds + \int_0^t f(h_s) a_{ii}(h_s) \langle X_s, e_i \rangle ds
\]

\[
+ \text{Martingale.} \quad (5.5)
\]

Recall \( q^i_t(y) \) is the probability density associated with \( h_s \langle X_s, e_i \rangle \), given \( X_0 \), so for example

\[
E[f(h_t) \langle X_t, e_i \rangle | X_0] = \int_0^\infty f(y) q^i_t(y) dy.
\]

Note the domain of \( q^i_t(y) \) is the range of \( h_t \) which is at most \([0, t]\); however, we write the \( y \) integrals over \([0, \infty)\). Therefore, taking expectations in (5.5), given \( X_0 \), we have

\[
\int_0^\infty f(y) q^i_t(y) dy = f(0) \int_0^\infty q^i_t(y) dy + \int_0^t \int_0^\infty f'(y) q^i_t(y) dy ds
\]

\[
- f(0) \int_0^t \left( \int_0^\infty a_{ii}(y) q^i_t(y) dy \right) ds
\]

\[
+ \int_0^t \int_0^\infty (f(y) a_{ii}(y) q^i_t(y) dy) ds.
\]
Integrating by parts, with $q_i(0) = q_i(1) = 0$ we have

$$\int_0^1 f(y)q_i(y)\,dy = f(0) \int_0^\infty q_i(y)\,dy - \int_0^t \left( \int_0^\infty f(y) \frac{\partial q_i(y)}{\partial y} \,dy \right) ds$$

$$- f(0) \int_0^t \left( \int_0^\infty a_{ii}(y)q_i(y)\,dy \right) ds$$

$$+ \int_0^t \left( \int_0^\infty f(y)a_{ii}(y)q_i(y)\,dy \right) ds.$$

As $f$ is arbitrary we see $q_i$ satisfies

$$q_i(y) = \delta_0(y) \int_0^\infty q_i(y)\,dy - \int_0^t \frac{\partial q_i(y)}{\partial y} \,ds$$

$$- \delta_0(y) \int_0^t \left( \int_0^\infty a_{ii}(y)q_i(y)\,dy \right) ds + \int_0^t a_{ii}(y)q_i(y)\,ds.$$  \hfill \square

**Corollary 5.4.** Taking marginals

$$\int_0^\infty q_i(y)\,dy = \left( \int_0^\infty q_1(y)\,dy, \int_0^\infty q_2(y)\,dy, \ldots, \int_0^\infty q_N(y)\,dy \right)' = E[X_t|X_0]$$

(5.6)

and

$$\sum_{i=1}^{\infty} q_i(y)\,dy = p(h_i \in [y, y+dy] | X_0).$$

(5.7)

**6. Approximations**

Suppose $X = \{X_t, t \geq 0\}$ is a finite state, continuous time homogeneous semi-Markov chain.

Equation (5.7) gives an expression for

$$E[X_t|X_0].$$

However, its calculation using the density $q_i(y)$ is not straightforward.

We now suggest an approximate procedure. With

$$p_i^t = P(X_t = e_i|X_0) = E[(X_t, e_i)|X_0]$$

we have

$$p_t := (p_1^t, p_2^t, \ldots, p_N^t)' = E[X_t|X_0] \in \mathbb{R}^N.$$  

Write $\hat{p}_t$ and $\hat{h}_t$ for $\mathcal{F}_t$-measurable approximations of $p_t$ and $h_t$, respectively. Consider a discretization $0 < \delta < 2\delta < \cdots < n\delta < (n + 1)\delta < \cdots$ of the time variable.

Suppose $t = n\delta$. Then from (3.5)

$$X_{t+\delta} = X_t + \int_t^{t+\delta} A(h_u)X_u\,du + (M_{t+\delta} - M_t).$$
Taking expectations, as 

\[ E[M_{t+\delta} - M_t | \mathcal{F}_t] = 0 \in \mathbb{R}^N \]

\[ p_{t+\delta} = p_t + \int_t^{t+\delta} E[A(h_u)X_u|X_0]du. \]  

(6.1)

Suppose we have \( \mathcal{F}_t \)-measurable approximations \( \hat{h}_t \) and \( \hat{p}_t \) of \( h_t \) and \( p_t \) have been found. Then from (5.1) an approximation for \( p_{t+\delta} \) is

\[ \hat{p}_{t+\delta} := \hat{p}_t + A(\hat{h}_t)\hat{p}_t \delta \in \mathbb{R}^N. \]  

(6.2)

Write \( e_{i^*} \) for a related MAP estimate of \( X_t \), that is a state \( e_{i^*} \) such that \( \hat{p}_t^{i^*} \geq p_t^i \) for all \( j, 1 \leq j \leq N \).

The recurrence (6.2) gives an expression for

\[ \hat{p}_{t+\delta} = (\hat{p}_{t+\delta}^1, \hat{p}_{t+\delta}^2, \ldots, \hat{p}_{t+\delta}^N)'. \]

Again write \( e_{j^*} \) for the related MAP estimate of \( X_{t+\delta} \), that is the state \( e_{j^*} \) such that \( \hat{p}_{t+\delta}^{j^*} \geq \hat{p}_{t+\delta}^\ell \) for all \( \ell, 1 \leq \ell \leq N \).

We then set \( \hat{h}_{t+\delta} = \hat{h}_t + \delta \) if \( i^* = j^* \) or

\[ \hat{h}_{t+\delta} = 0 \text{ if } i^* \neq j^*. \]

We proceed in this way so that at each time \( t + kh \), if estimates \( \hat{p}_{t+kh} \) and \( \hat{h}_{t+kh} \) are known we can obtain estimates \( \hat{p}_{t+(k+1)h} \) and \( \hat{h}_{t+(k+1)h} \). This can be used in calculations involving \( p_t \) and \( h_t \).

7. Conclusion

This paper discusses a continuous time, finite state, homogeneous semi-Markov chain \( X \). Its probabilistic and dynamic properties can be described by either a rate matrix, (or \( Q \)-matrix), or a matrix which describes occupation times in the various states. The rate matrix gives the rates of jumping from one state to another. The relation between these two descriptions is derived and the semimartingale decomposition of the semi-Markov chain derived.

An equation giving the dynamics of the occupation times is obtained. The dynamics for the conditional expectation of the occupation time and state is then derived. Finally some approximations for the dynamics are developed. Semi-Markov processes often fit observed data better than Markov processes. The results of this paper might find applications in fields such as signal processing and financial modelling.

Acknowledgment. The support of NSERC and the Australian Research Council is gratefully appreciated.

References


Robert J. Elliott: University of South Australia, Adelaide, Australia and University of Calgary, Calgary, AB, Canada

*E-mail address*: relliott@ucalgary.ca