
November 26, 2012, LTCC Stochastic Processes

Outline

1. Invariant distributions and limiting behaviour for continuous-time Markov Chains
   - Recap from Lecture 3
   - Examples
   - What does an invariant distribution tell us about the Markov Chain?
   - Examples

2. Embedded Markov Chain
   - Irreducible continuous-time Markov Chains

3. Time reversibility for Markov Processes
   - Discrete-time Markov Chains
     - Detailed balance equations
     - Examples
Recap from Lecture 3

**Definition**

The probability row vector $\pi$ is an **invariant distribution** if $\pi = \pi P(t)$ for all $t \geq 0$. This is equivalent to

$$\pi Q = 0.$$ 

If $p(0) = \pi$, then $p(t) = \pi P(t) = \pi$, for all $t \geq 0$.

**Definition**

An **equilibrium distribution** $\pi$ exists if $p_{ij}(t) \to \pi_j$ as $t \to \infty$ for each $j \in S$ (where $\pi$ is a probability distribution that does not depend on $i$).

If this is the case, $p'_{ij}(t) \to 0$ as $t \to \infty$.

The forward equation is

$$p'_{ij}(t) = \sum p_{ik}(t)q_{kj}, \text{ or } 0 = \sum \pi_k q_{kj} = \pi Q.$$
Example 1 (M/M/1 queue)

Assume we have a server which can service one customer at a time. The service times for customers are independent identically distributed exponential random variables with parameter $\mu > 0$. The arrival times are also assumed to be independent identically distributed exponential random variables with parameter $\lambda > 0$. The customers waiting to be serviced stay in a queue and we let $X(t)$ denote the number of customers in the queue at time $t$. Our assumption regarding the exponential arrival times implies:

$$P(X(t + h) = k + 1 | X(t) = k) = \lambda h + o(h).$$

Similarly the assumption about service times implies

$$P(X(t + h) = k - 1 | X(t) = k) = \mu h + o(h).$$
The generator is

$$Q = \begin{pmatrix}
-\lambda & \lambda & 0 & 0 & 0 & \ldots \\
\mu & -(\lambda + \mu) & \lambda & 0 & 0 & \ldots \\
0 & \mu & -(\lambda + \mu) & \lambda & 0 & \ldots \\
0 & 0 & \mu & -(\lambda + \mu) & \lambda & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots
\end{pmatrix}.$$ 

To find the invariant distribution $\pi = (\pi_0, \pi_1, \ldots)$ we need to solve $\pi Q = 0$, which is equivalent to solving

$$-\lambda \pi_0 + \mu \pi_1 = 0, \lambda \pi_{i-1} - (\lambda + \mu) \pi_i + \mu \pi_{i+1} = 0 \text{ for } i \geq 1,$$

from which we get

$$\pi_i = \left(\frac{\lambda}{\mu}\right)^i \pi_0 \text{ for all } i \geq 1.$$
Assume first $\lambda/\mu < 1$ and recall that $\sum_{i \geq 0} \pi_i = 1$. Then we have

$$\sum_{i \geq 0} \pi_i = \sum_{i \geq 0} \left( \frac{\lambda}{\mu} \right)^i \pi_0 = 1,$$

from which we obtain

$$\pi_0 = (1 - \lambda/\mu) \quad \text{and} \quad \pi_i = (1 - \lambda/\mu) \left( \frac{\lambda}{\mu} \right)^i, \quad i \geq 1.$$

If $\lambda/\mu \geq 1$, there is no invariant distribution $\pi$. 
Example 2 (Birth and death processes)

More generally, suppose that the number $X(t)$ of individuals alive at time $t$ in some population evolves continuously in the following way

(a) $X(t)$ is a Markov chain taking values in $S = \{0, 1, 2, \ldots\}$

(b) For small $h$ the transition probabilities are given by:

$$P(X(t + h) = k + 1 | X(t) = k) = \lambda_k h + o(h), k = 0, 1, 2 \ldots,$$

$$P(X(t + h) = k - 1 | X(t) = k) = \mu_k h + o(h), k = 0, 1, 2 \ldots$$

and

$$P(X(t + h) = k + m | X(t) = k) = o(h), |m| > 1.$$

where $\lambda_k$ is the birth rate for a population of size $k$ and $\mu_k$ is the corresponding death rate. We assume $\mu_0 = 0$, $\lambda_0 > 0$ and $\mu_k > 0$, $\lambda_k > 0$, for $k > 0$. 
The generator is

\[
Q = \begin{pmatrix}
-\lambda_0 & \lambda_0 & 0 & 0 & 0 & \ldots \\
\mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & \ldots \\
0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & \ldots \\
0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \lambda_3 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots
\end{pmatrix}.
\]

To find the invariant distribution \( \pi = (\pi_0, \pi_1, \ldots) \) we need to solve
\( \pi Q = 0 \), which is equivalent to solving

\[
\begin{align*}
-\lambda_0 \pi_0 + \mu_1 \pi_1 &= 0, \\
\lambda_{i-1} \pi_{i-1} - (\lambda_i + \mu_i) \pi_i + \mu_{i+1} \pi_{i+1} &= 0 \quad \text{for } i \geq 1,
\end{align*}
\]

from which we get

\[
\pi_i = \frac{\lambda_{i-1}}{\mu_i} \pi_{i-1} = \ldots = \frac{\lambda_{i-1} \ldots \lambda_0}{\mu_i \ldots \mu_1} \pi_0 \quad \text{for all } i \geq 1.
\]
Recall that $\sum_{i \geq 0} \pi_i = 1$. Then we have

$$\sum_{i \geq 0} \pi_i = \sum_{i \geq 0} \frac{\lambda_{i-1} \cdots \lambda_1}{\mu_i \cdots \mu_1} \pi_0 = 1,$$

from which we obtain

$$\pi_0 = \left( \sum_{i \geq 0} \frac{\lambda_{i-1} \cdots \lambda_0}{\mu_i \cdots \mu_1} \right)^{-1},$$

provided that $\sum_{i \geq 0} \frac{\lambda_{i-1} \cdots \lambda_0}{\mu_i \cdots \mu_1} < \infty$.

If no such $\pi_0$ exists (i.e., $\sum_{i \geq 0} \frac{\lambda_{i-1} \cdots \lambda_0}{\mu_i \cdots \mu_1} = \infty$ and $\pi_i = 0$, $i \geq 0$) then our assumption is false, i.e. the process cannot be in equilibrium.

*See also Example 2.5 (Poisson process) from Lecture Notes, question 1 from Exam 2011 and the in-class example.*
What does an invariant distribution tell us about the Markov Chain?

Let $X(t), t \in [0, \infty)$, be a continuous-time Markov Chain with invariant distribution $\pi$ and let

$$V_i(t) = \int_0^t 1_{\{X(s) = i\}} ds = \text{the time spent in state } i \text{ up to time } t.$$ 

Then we have

$$\lim_{t \to \infty} \frac{V_i(t)}{t} = \pi_i \text{ with probability } 1,$$

which gives the proportion of time spent in state $i$ over long periods of time.
(Note the similarity with the properties for discrete-time Markov-Chains.)
Examples

- Let \( X(t), t \in [0, \infty) \), be a continuous-time Markov chain with generator

\[
Q = \begin{pmatrix}
-3 & 3 \\
1 & -1
\end{pmatrix}.
\]

Then the long run, the chain will be in state zero with a probability of \(1/4\) as the invariant distribution is \(\pi = (1/4, 3/4)\).

- In example 1, the proportion of customers in state 1 in the long-run is \((1 - \lambda/\mu) \frac{\lambda}{\mu}\).

- See also question 1 parts (a) and (b) from Exam 2012.
Embedded Markov Chain

An interesting way of analyzing a CTMC is through the embedded discrete-time Markov chain. If we consider the Markov process only at the moments upon which the state of the system changes, and we number these instances 0, 1, 2, . . . , then we get a discrete-time Markov chain.

Let $X(t), t \in [0, \infty)$ be a CTMC with generator $Q$ and initial state $i$. Then the chain leaves state $i$ at a rate

$$\sum_{j \neq i} q_{ij} = -q_{ii},$$

hence stays in $i$ for a holding time $T_1$ that is exponentially distributed with mean $-1/q_{ii}$. When the chain leaves state $i$, it jumps to $j$ ($j \neq i$) with probability $p_{ij} = -q_{ij}/q_{ii}$. The chain then stays in $j$ for an exponential holding time $T_2$ (mean $-1/q_{jj}$) before jumping to $k$ ($k \neq j$) with probability $-q_{jk}/q_{jj}$. And so on. Successive holding times are independent.
More precisely, let
\[ T_n = \inf\{t \geq T_{n-1} : X(t) \neq X(T_{n-1})\} \], with \( T_0 = 0 \).

Define
\[ Y(n) := X(T_n) \text{ if } T_n < \infty, \text{ and } Y(n) := \Delta \text{ if } T_n = \infty, \]
where \( \Delta \) is an arbitrary element not in \( S \). Then \( Y_n, n = 0, 1, 2, \ldots \) is a discrete-time Markov Chain with
\[ p_{ij} = -q_{ij}/q_{ii}, i \neq j \text{ and } p_{ii} = 0. \]

Every continuous-time Markov chain with transition probabilities has an associated embedded discrete-time Markov chain. While the transition matrix \( P \) completely determines the probabilistic behaviour of the embedded discrete-time Markov chain, it does not fully capture the behaviour of the continuous-time process because it does not specify the rates at which transitions occur.

See also the in-class example.

Irreducible continuous-time Markov Chains

Definition

State $j$ is said to be reachable from state $i$ for a CTMC if

$$P(X(s) = j|X(0) = i) = p_{ij}(s) > 0 \text{ for some } s \geq 0.$$  

Definition

As with discrete-time chains, states $i$ and $j$ are said to communicate if state $j$ is reachable from state $i$, and state $i$ is reachable from state $j$.

It is immediate that $i$ and $j$ communicate in continuous time if and only if they do so for the embedded discrete-time chain $Y_n$. They communicate in continuous-time if and only if they do so at transition epochs. Thus once again, we can partition the state space up into disjoint communication classes $S = C_1 \cup C_2 \cup \ldots$. 
An **irreducible chain** is a chain for which all states communicate (i.e., $S = C_1$ and we have a single communication class).

**Definition**

A continuous-time Markov chain is defined to be **irreducible** if, for every $i$ and $j$, $p_{ij}(t) > 0$ for some $t$.

*See Example 2 for an irreducible CTMC and example 2.5 from lecture notes for a CTMC with is not irreducible.*

Notions of recurrence and transience are similar as for discrete-time chains. Let $T_{ii}$ denote the amount of (continuous) time until the chain re-visits state $i$ (at a later transition) given that $X(0) = i$ (defined to be $\infty$ if it never does return).

**Definition**

State $i$ is called **recurrent** if, with probability 1, the chain will re-visit state $i$ with certainty, that is, if $P(T_{ii} < \infty) = 1$. The state is called **transient** otherwise.

- Embedded Markov Chain
- Irreducible continuous-time Markov Chains

**Theorem**

- A CTMC is irreducible if and only if its embedded chain is irreducible
- A state $i$ is recurrent/transient for a CTMC if and only if it is recurrent/transient for the embedded discrete-time chain.

**Theorem**

*For an irreducible continuous-time Markov chain, we have*

- (a) If there exists an invariant distribution $\pi$ then it is unique and $p_{ij}(t) \to \pi$ as $t \to \infty$.
- (b) If there is no invariant distribution then $p_{ij}(t) \to 0$ as $t \to \infty$. 
Time reversibility for Markov Processes

For Markov chains, the past and future are independent given the present. This property is symmetrical in time and suggests looking at Markov chains with time running backwards. On the other hand, convergence to equilibrium shows behaviour which is asymmetrical in time: a highly organised state such as a point mass decays to a disorganised one, the invariant distribution. So if we want complete time-symmetry we must begin in equilibrium. The next result shows that a Markov chain in equilibrium, run backwards, is again a Markov chain. The transition matrix may however be different.

Consider a stochastic process $X(t)$ ($-\infty < t < \infty$ or $t = \ldots, -2, -1, 0, 1, 2, \ldots$). The process $X(t)$ is reversible if its stochastic behaviour remains the same when the direction of time is reversed.
Definition

The process $X(t)$ is **reversible** if $(X(t_1), X(t_2), \ldots, X(t_n))$ has the same joint distribution as $(X(\tau - t_1), X(\tau - t_2), \ldots, X(\tau - t_n))$ for all $t_1, t_2, \ldots, t_n, \tau$ and $n$. Setting $\tau = 0$, it follows that $(X(t_1), X(t_2), \ldots, X(t_n))$ and $(X(-t_1), X(-t_2), \ldots, X(-t_n))$ have the same joint distribution.

Definition

A process $X(t)$ is **stationary** if $(X(t_1), X(t_2), \ldots, X(t_n))$ has the same (joint) distribution as $(X(t_1 + \tau), \ldots, X(t_n + \tau))$ for all $\tau$, all times $t_1, \ldots, t_n$ and all $n$.

Theorem

*A reversible process is stationary.*

(For proof, see page 21 in the lecture notes). *See also the in-class examples of time-reversibility.*
In order to prove this, we need to introduce the concept of time reversibility for Markov Processes.

Let $X_n, n = \ldots, -2, -1, 0, 1, 2 \ldots$, be a stationary, irreducible discrete-time Markov chain with transition matrix $P = (p_{ij})_{i,j \in S}$ and equilibrium distribution $\pi$. Then we have the following: The reversed-time process $X_n^* := X_{-n}$ is also a stationary, irreducible discrete-time Markov chain with equilibrium distribution $\pi$ and transition matrix $P^* = (p_{ij}^*)_{i,j \in S}$ where

$$ p_{ij}^* = \pi_j p_{ji} / \pi_i. $$

(See the proof on page 22 in the lecture notes).

If the chain is reversible, then

$$ P(X_n = i, X_{n+1} = j) = P(X_{n+1} = i, X_n = j). $$
Thus

\[ \pi_i p_{ij} = \pi_j p_{ji} \text{ for all } i, j \in S. \]

These are called the **detailed balance equations**. In words, they say that, for all pairs of states, the rate at which transitions occur from state \( i \) to state \( j \) (\( \pi_i p_{ij} \)) balances the rate at which transitions occur from \( j \) to \( i \) (\( \pi_i p_{ij} \)).

**Theorem**

An irreducible stationary Markov chain is reversible if and only if there exists a probability row vector \( \pi (0 \leq \pi_i \leq 1, \sum_{i \in S} \pi_i = 1) \) such that

\[ \pi_i p_{ij} = \pi_j p_{ji} \text{ for all } i, j \in S. \]

If such a \( \pi \) exists, it is the equilibrium distribution of the chain.

**Note** that if a chain is stationary i.e. has an equilibrium distribution, but there is no solution to the detailed balance equations, then the chain is not reversible.
Examples

- Assume that $X_n, n = 0, 1, 2\ldots$, is a stationary discrete-time MC with

  $$P = \begin{pmatrix} 1/2 & 1/2 \\ 2/3 & 1/3 \end{pmatrix}.$$ 

  Then the MC is reversible and has an equilibrium distribution.

- Find the solution to the detailed balance equations for the discrete MC with

  $$P = \begin{pmatrix} 0 & a & 1 - a \\ 1 - a & 0 & a \\ a & 1 - a & 0 \end{pmatrix}.$$
The detailed equations are:

\[ \pi_1 a = \pi_2 (1 - a), \pi_2 a = \pi_3 (1 - a), \pi_3 a = \pi_1 (1 - a). \]

Solving them, we get a solution if \( a = 1/2 \) and no solution if \( a \neq 1/2 \).

- See also the in-class Examples and Example 2.6 and Example 2.7 from the lecture notes.

- Time reversibility for Markov Processes
- Discrete-time Markov Chains

ANY QUESTIONS?