Chapter 6

Continuous-time Markov Chains (CTMC)

In this chapter we turn our attention to continuous-time Markov processes that take values in a denumerable (countable) set that can be finite or infinite. Such processes are referred to as continuous-time Markov chains. As we shall see the main questions about the existence of invariant distributions, the ergodic theorem, etc. can be obtained from the corresponding techniques that we saw for discrete-time Markov chains.

6.1 Definition and Basic Properties

Let $\{X_t\}_{t\geq 0}$ be a stochastic process that takes values in a countable set E. Throughout this chapter we will use the following notation:

Definition 6.1.1 *Let $\mathcal{F}_t^X = \sigma\{X_u, u \leq t\}$ i.e. the σ-field of events generated by the process up to t. This is often referred to as the filtration generated by X_t or the* history *of the process up to t.*

$\sigma(X)$ will denote the sigma-field of events generated by the r.v. X. When the context is clear we will drop the superscript X.

A continuous-time Markov chain is a Markov process that takes values in E. More formally:

Definition 6.1.2 *The process $\{X_t\}_{t\geq 0}$ with values in E is said to a a continuous-time Markov chain (CTMC) if for any $t > s$:*

$$\mathbb{P}\left(X_t \in A|\mathcal{F}_s^X\right) = \mathbb{P}\left(X_t \in A|\sigma(X_s)\right) = \mathbb{P}\left(X_t \in A|X_s\right) \tag{6.1. 1}$$

In particular, let us denote:

$$P_{ij}(s, s+t) = \mathbb{P}(X_{t+s} = j|X_s = i) \tag{6.1. 2}$$

If $P_{ij}(s, s+t) = P_{ij}(t)$, i.e. it only depends on the difference t between $t+s$ and s then $\{X_t\}$ is referred to a a homogeneous CTMC or a homogeneous Markov chain for short. From now on we will assume that the Markov chain is homogeneous that we will abbreviate as HMC.

Let us define the matrix (could be infinite dimensional) $P(t)$ as

$$P(t) = \{P_{ij}(t)\}_{ij\in E\times E}$$

with $P(0) = I$. $P(t)$ is referred to as the transition probability matrix of X.

Then from the Markov property we have the following semigroup property satisfied by $P(t)$ that is referred to as the Chapman-Kolmogorov equations.

Proposition 6.1.1 *Let $\{X_t\}_{t\geq 0}$ be a homogeneous CTMC taking values in E. Then:*

$$P(s+t) = P(s)P(t) = P(t)P(s)$$

or equivalently:

$$P_{ij}(t+s) = \sum_{k\in E} P_{ik}(s)P_{kj}(t) = \sum_{k\in E} P_{ik}(t)P_{kj}(s)$$

Let us now define the following sequence of times, T_n referred to as the jump-times of the Markov chain when it transitions from one state to the other:

$$\begin{aligned} T_1 &= \inf\{t : X_t \neq X_0\} \\ T_n &= \inf\{t \geq T_{n-1} : X_t \neq X_{T_{n-1}}\} \end{aligned}$$

Then by definition the sequence $\{T_n\}$ forms a sequence of Markov or stopping times. This sequence allows us to characterize the sample-path trajectories of the M.C. In particular the trajectories of $\{X_t\}$ are piecewise constant between jump times. By definition $\{X_t\}$ is right-continuous.

As in the case of DTMC we can show that CTMC Markov chains satisfy the strong Markov property. We state the result below.

Proposition 6.1.2 *Let τ be a Markov or stopping time relative to $\mathcal{F}_t$ .i.e the event: $\{\tau \leq t\} \in \mathcal{F}_t$. Then the following is true:*

1. *The process $X_{\tau+t}$ is independent of $\mathcal{F}_u, u \leq \tau$ given X_τ.*
2. $\mathbb{P}(X_{\tau+t} = j | X_\tau = i) = P_{ij}(t)$.

The first important result we establish is the fact that the sequence of the inter-jump times given by $T_1, T_2 - T_1, T_3 - T_2, \cdots, T_n - T_{n-1}, \cdots$ are independent and exponentially distributed random variables and in particular can be viewed as the points of an inhomogeneous Poisson process that we will specify.

Define:

$$\tau_t = \inf\{s > t : X_s \neq X_t\} \tag{6.1. 3}$$

i.e. the first time after t when the process leaves the state X_t.

Proposition 6.1.3 *Let $\{X_t\}$ be a HMC on E. Let τ_t be as defined in (6.1. 3). Then there exists a parameter $q_i > 0$ such that:*

$$\mathbb{P}(\tau_t > u | X_t = i) = e^{-q_i t}$$

Proof:

Let us define:

$$f_i(u+v) = \mathbb{P}(\tau_t > u+v | X_t = i)$$

Now, by the definition of τ_t, if $\tau_t > u+v$ it implies that $X_{t+u} = i$ since the process has not changed state at time $t+u$.

Therefore,

$$\begin{aligned} f_i(u+v) &= \mathbb{P}(\tau_t > u, \tau_t > u+v | X_t = i) \\ &= \mathbb{P}(\tau_t > u+v | \tau_t > u, X_t = i)\mathbb{P}(\tau_t > u | X_t = i) \\ &= \mathbb{P}(\tau_t > u+v | X_{t+u} = i)\mathbb{P}(\tau_t > u | X_t = i) \\ &= f_i(v) f_i(u) \end{aligned}$$

where we have used homogeneity and the strong Markov property in the last step.

Hence:

$$f_i(u+v) = f_i(u) f_i(v), \quad f_i(0) = 1$$

Hence the only solution to this equation with $f_i(t) \in (0,1)$ is

$$f_i(t) = e^{-q_i t}$$

for some $q_i > 0$.

The above result shows that the MC spends an amount of time that is exponentially distributed in a given state.

Before proceeding, let us recall some properties of exponentially distributed random variables.

Proposition 6.1.4 *Let X and Y be independent exponentially distributed r.v's. with parameters λ_x and λ_Y respectively.*

Then:

1. $\mathbb{P}(X > u+v | X > u) = \mathbb{P}(X > v) = e^{-\lambda_X v}$ *for any $u, v > 0$. This is referred to as the* memoryless property.
2. *Let S be any positive r.v. independent of X. Then :*

$$\mathbb{P}(X > S+u | X > S) = \mathbb{P}(X > u) = e^{-\lambda_X u}$$

3. *Define $Z = \min(X, Y)$ then:*

$$\mathbb{P}(Z > u) = e^{-(\lambda_X + \lambda_Y)u}$$

And moreover Z is independent of the events $\{X < Y\}$ and $\{Y < X\}$, i.e. it is independent of which of the two random variables is smaller.

Proof: 1. follows by definition. 2. follows by conditioning on S and using independence. Only 3. needs proof.

The first part of 3. is easy by the definition of Z, since:

$$\mathbb{P}(Z > u) = \mathbb{P}(X > u, Y > u) = \mathbb{P}(X > u)\mathbb{P}(Y > u) = e^{-(\lambda_X + \lambda_Y)u}$$

from independence of X and Y. For the second part :

$$\begin{aligned} \mathbb{P}(Z > u, X > Y) &= \mathbb{P}(Y > u, X > Y) \\ &= \int_u^\infty \mathbb{P}(X > x) dF_Y(x) \\ &= \int_u^\infty e^{-\lambda_X x} \lambda_Y e^{-\lambda_Y x} dx \\ &= \frac{\lambda_Y}{\lambda_X + \lambda_Y} e^{-(\lambda_X + \lambda_Y)u} \end{aligned}$$

But $\frac{\lambda_Y}{\lambda_X+\lambda_Y}$ is just the probability $\mathbb{P}(X > Y)$ and so

$$\mathbb{P}(Z > u, X > Y) = \mathbb{P}(X > Y)\mathbb{P}(Z > u)$$

establishing the independence. The independence between Z and the event $\{Y > X\}$ follows by interchanging the roles of X and Y.

Remark 6.1.1 *From Proposition 6.1.4 and the strong Markov property it readily follows that: $T_1, T_2 - T_1, \cdots$ forms a sequence of independent exponentially distributed random variables. Thus the sequence $\{T_n\}$ can be viewed as the points of an inhomogeneous Poisson process whose intensity depends on the state of the underlying MC- sometimes referred to as a doubly stochastic Poisson process.*

We now turn our attention to the study of the behavior of $P(t)$ and techniques for its computation. Let us recall the properties of $P(t)$

- $P(t)$ is continuous with $P(0) = \lim_{t\to 0} P(t) = I$
- $P(t + s) = P(s)P(t) = P(t)P(s)$

Such an operator (as a function of t) is called a C_0 semi-group. The additional requirement that $P_{ij}(t) \in [0, 1]$ actually imposes further conditions that we shall see. THe key point is that from the continuity and semi-group property the derivative of $P(t)$ exists. We state this property below as a proposition.

Proposition 6.1.5 *Let $P(t)$ be the probability transition matrix of a CTMC. Then the infinitesimal generator (or simply generator) Q exists and is defined as:*

$$Q = \lim_{t\to 0} \frac{P(t) - I}{t} \tag{6.1. 4}$$

and moreover:

$$\frac{d}{dt}P(t) = QP(t) = P(t)Q \tag{6.1. 5}$$

Proof: Rather than give the detailed proof, let us show that

$$\lim_{t\to 0} \frac{P_{ii}(t) - 1}{t} = q_{ii}$$

exists.

First note that from the semi-group property: $P(t) = [P(\frac{t}{n})]^n$ and therefore $P_{ii}(t) \geq [P_{ii}(\frac{t}{n})]^n$ for all $i \in E$. Now since $\lim_{t\to 0} P_{ii}(t) = 1$ there exists $\varepsilon > 0$ such that for all $h \in [0, \varepsilon)$ $P_{ii}(h) > 0$. Therefore since for all finite t and n sufficiently large, $\frac{t}{n} \in [0, \varepsilon)$ we have for every $t \geq 0$, $P_{ii}(t) > 0$. Define:

$$f_i(t) = -\log P_{ii}(t) < \infty$$

Since $f_i(t) \to 0$ as $t \to 0$ and from the semi-group property it readily follows that:

$$f_i(t + s) \leq f_i(t) + f_i(s)$$

or $f_i(t)$ is sub-additive.

Define

$$q_i = \sup_{t>0} \frac{f_i(t)}{t}$$

Then it follows from sub-additivity that:

$$\lim_{h \downarrow 0} \frac{f_i(h)}{h} = q_i$$

Therefore:

$$\lim_{h \downarrow 0} \frac{P_{ii}(h) - 1}{h} = \lim_{h \downarrow 0} \frac{e^{f_i(h)} - 1}{f_i(h)} \frac{f_i(h)}{h} = -q_i = q_{ii}$$

The proof of $\lim_{h \downarrow 0} \frac{P_{ij}(h)}{h} = q_{ij}$ for some $q_{ij} > 0$ is more complicated but can be shown similarly by noting that

$$P_{ij}(t) \leq 1 - P_{ii}(t)$$

and from the Markov property:

$$P_{ij}(nh) \geq nP_{ij}(h)C$$

where $C \to 1$ is a constant related to $P_{jj}(h)$ as $h \to 0$.

Remark 6.1.2 *The infinitesimal generator Q is often referred to as the rate matrix of the Markov chain and plays the same function as the transition matrix P of discrete-time chains. We will use the term rate matrix.*

From the definition of Q it readily follows that formally we can write:

$$P(t) = e^{Qt} \qquad (6.1.\ 6)$$

If $|E| < \infty$ then this is just the matrix exponential function. When $|E| = \infty$ we need to define it an appropriate way that we will discuss later. But first of all let us study some properties of Q and the relation between q_{ii} and the parameter q_i of the exponential sojourn time in state i we had shown before.

We state this below.

Proposition 6.1.6 *Let $Q = \{q_{ij}\}_{(ij) \in E \times E}$ be the rate matrix of a CTMC. Then:*

a) $\sum_{j \in E} q_{ij} = 0 \quad \forall\ i \in E$

b) $q_{ii} = q_i$ *where* q_i *is the parameter associated with the exponentially distributed sojourn time in state* i.

Proof: First note that for every $t \geq 0$

$$\sum_{j \in E} P_{ij}(t) = 1$$

Therefore differentiating we obtain:

$$\begin{aligned} \frac{d}{dt} \sum_{j \in E} P_{ij}(t) = 0 &= \sum_{j \in E} q_{ik} P_{kj}(t) \\ &= \sum_{k \in E} q_{ik} \sum_{j \in E} P_{kj}(t) \\ &= \sum_{k \in E} q_{ik} \end{aligned}$$

So in particular: $q_{ii} = -\sum_{j \neq i} q_{ij} < 0$ since $q_{ij} \geq 0$. By definition $-q_{ii}$ is the rate at which the MC exits from state i and thus from the definition of q_i earlier we have $q_i = -q_{ii}$.

The evolution equations for $P(t)$ given by $\frac{d}{dt}P(t) = QP(t) = P(t)Q$ are called the Kolmogorov equations.

Specifically the form:

$$\frac{d}{dt}P(t) = QP(t)$$

or

$$\frac{d}{dt}P_{ij}(t) = -q_i P_{ii}(t) + \sum_{k \neq i} q_{ik} P_{kj}(t)$$

is called the Kolmogorov backward equation and is written interms of the change of flux out of state i. Formally it can be viewed as taking conditioning X_{t+h} on X_h and then letting $t \to 0$.

The other equation is equivalent to stating:

$$\frac{d}{dt}P_{ij}(t) = -q_i P_{ii}(t) + \sum_{k \neq i} P_{ik}(t) q_{kj}$$

and is called the forward equation that represents the flux coming into state i. This can be viewed as conditioning X_{t+h} on X_t and letting $h \to 0$.

6.1.1 Jump or Embedded Chains

Let T_n be the sequence of jump times of a CTMC X_t. The intervals $T_{n+1} - T_n = S_n$ denote the sojourn times in a given state. Define the discrete-time process:

$$Y_n = X_{T_n}$$

Then Y_n is called the embedded or jump chain. Note that since the sequence $\{T_n\}$ forms a set of stopping or Markov times with respect to $\mathcal{F}^X$ the process $\{Y - n\}$ is a DTMC.

Let $\pi_{ij} = \mathbb{P}(Y_{n=1} = j | Y_n = i) = \mathbb{P}(X_{T_n + J_n} = j | X_{T_n} = i)$ denote the transition probabilities of the chain $\{Y_n\}$.

Lemma 6.1.1 *The embedded MC $\{Y_n\}$ has the following transition probabilities:*

$$\begin{aligned} \pi_{ij} &= 0 \quad i = j && (6.1.\ 7) \\ &= \frac{q_{ij}}{q_i} \quad j \neq i && (6.1.\ 8) \end{aligned}$$

Proof: If i=j, then by definition $\pi_{ii} = 0$ as the chain $\{Y_n\}$ only corresponds to changes in the state.

For $j \neq i$ we have by the definition of the rates q_{ij}:

$$\begin{aligned} \pi_{ij} &= \mathbf{E}_i[\int_0^{S_1} q_{ij} dt] \\ &= q_{ij} \mathbf{E}_i[S_1] \\ &= \frac{q_{ij}}{q_i} \end{aligned}$$

where we used the fact that conditioned on being in state i the mean sojourn time (given by the mean of the exponential distribution with parameter q_i) is $\frac{1}{q_i}$.

6.1.2 Regularity and Stationarity

Let us now turn our attention to the long-term behavior of CTMCs. The key issue is that of existence of a stationary distribution and convergence to stationarity. As in the case of DTMCs, it is fundamentally governed by the behavior of $P(t)$ that itself is determined by the rate matrix Q.

Let $\pi_i(t) = \mathbb{P}(X_t = i)$. Then from the Markov property it readily follows that:

$$\pi_i(t) = \sum_{j \in E} \pi_j(0) P_{ji}(t)$$

Suppose there exists a stationary distribution (also referred to as the equilibrium distribution), that we denote by:

$$\pi = (\pi_0, \pi_1, \cdots, \pi_j, \cdots)$$

Then:

$$\pi = \pi P(t), \quad t \geq 0$$

Now differentiating both sides we obtain:

$$\begin{aligned} \frac{d}{dt}\pi = 0 &= \pi \frac{d}{dt} P(t) \\ &= \pi Q P(t), \quad \forall t \ \geq 0 \end{aligned}$$

This implies that π satisfies:

$$\pi Q = 0 \tag{6.1. 9}$$

The condition (6.1. 9) is thus necessary and the equation when expanded for each state $i \in E$ can be written as

$$\begin{aligned} \sum_{j \in E} \pi_j q_{ji} &= 0, \quad i \in E \\ -q_{ii}\pi_i &= \sum_{j \in E, j \neq i} \pi_j q_{ji} \ , \ i \in E \end{aligned} \tag{6.1. 10}$$

The equations (6.1. 10) is referred to as the (global) *balance equations* and states that in equilibrium the total probability flux out of state i given by $-q_{ii}\pi_i$ is equal to the probability flux into state i.

Let us now study the relationship between the stationary distribution of a CTMC and its embedded chain.

Proposition 6.1.7 *Let $\pi = (\pi_0, \pi_1 m \cdots,)$ denote the stationary distribution of a homogeneous CTMC $\{X_t\}_{t \geq 0}$ and $\tilde{\pi}$ denote the the stationary distribution of the embedded chain $Y_n = X_{T_n}$ where $\{T_n\}$ are the jump points of X_t.*

Let $\beta = \sum_{i \in E} \pi_i q_i < \infty$. Then:

$$\tilde{\pi}_i = \frac{\pi_i q_i}{\beta} \tag{6.1. 11}$$

Proof: To show this result we need to show that $\tilde{\pi} = \tilde{\pi} P$ where P is the transition probability matrix of Y_n. But by definition $P_{ij} = \frac{q_{ij}}{q_i}$, $i \neq j$, *and* 0 otherwise. So we need to establish that:

$$(\frac{\pi_0 q_0}{\beta}, \ldots, \frac{\pi_i q_i}{\beta}, \ldots) = (\frac{\pi_0 q_0}{\beta}, \ldots, \frac{\pi_i q_i}{\beta}, \ldots) P$$

We obtain for the i-th term:

$$\frac{\pi q_i}{\beta} = \sum_{j \neq i} \frac{\pi_j q_j}{\beta} \frac{q_{ji}}{q_j} = \sum_{j \neq i} \frac{\pi_j q_{ji}}{\beta}$$

or:

$$\pi_i q_i = \sum_{j \neq i} \pi_j q_{ji}$$

which is equivalent to

$$\pi Q = 0$$

which corresponds to π being the stationary distribution of X_t as per assumption.

Note the factor β does not play a role in the calculations. It is just required for normalization. We will now study the questions of ergodicity and convergence to steady-state for CTMCs. But before doing so we need to introduce the important concept of regularity for CTMCs that does not arise in the context of DTMCs.

The classification of states for CTMC is done via the embedded or jump chain $Y_n = X_{T_n}$.

Definition 6.1.3 *A state i is said to be recurrent (resp. transient) for X_t if and only if it is recurrent (resp. transient) for the jump chain.*

Definition 6.1.4 *The state space E is said to be irreducible for X_t if and only if it is irreducible for Y_n.*

Let $S_n = T_{n+1} - T_n$ be the sequence the sequence of sojourn times of the CTMC. By definition: $T_n = \sum_1^{n-1} S_k$ and let $\xi = \lim_{n \to \infty} T_n = \sum_{n=1}^{\infty} S_n$. Then ξ is a measure of the *lifetime of the process*, i.e. it is the total time that the process X_t spends in all the states. Now if X_t is *well behaved* we would want ξ to be infinite. This is the issue of *regularity* and we will see that it is related to the behavior of q_i. If $\xi < \infty$ it would imply that the MC has an infinite number of transitions in finite time, and thus $\xi = \infty$ means that the MC *persists* so that we can define the notion of a stationary distribution as the fraction of time it spends in a given state.

To study regularity we need the following lemma.

Lemma 6.1.2 *Let $\{S_k\}$ be independent exponentially distributed r.v's with $\mathbf{E}[S_n] = \frac{1}{\lambda_n}$. Define $T_\infty = \sum_{n=1}^{\infty} S_n$.*

$$\begin{aligned} \mathbb{P}(T_\infty < \infty) &= 0 \quad if \ \sum_n \frac{1}{\lambda_n} = \infty \\ &= 1 \quad if \ \sum_n \frac{1}{\lambda_n} < \infty \end{aligned}$$

Proof: First note that:

$$\mathbf{E}[T_\infty] = \sum_n \frac{1}{\lambda_n}$$

and therefore $\sum_n \frac{1}{\lambda_n} < \infty$ implies that $\mathbb{P}(T_\infty = \infty) = 0$.

On the other hand:

$$\begin{aligned}\mathbf{E}[e^{-T_\infty}] &= \prod_{n=1}^{\infty} \mathbf{E}[e^{-S_n}] \text{ (by independence)} \\ &= \prod_{n=1}^{\infty} \frac{\lambda_n}{1+\lambda_n} \leq [\sum_{n=1}^{\infty} \frac{1}{\lambda_n}]^{-1}\end{aligned}$$

Therefore if $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$ it implies $\mathbf{E}[e^{-T_\infty}] = 0$ or $\mathbb{P}(T_\infty = \infty) = 1$.

Hence by noting that the r.v's S_n are exponentially distributed with parameter $\lambda_n = q_{X_{T_n}}$ $\mathbb{P}(\sum_{n=1}^{\infty} \frac{1}{q_{X_{T_n}}} = \infty) = 1$ implies that $\mathbb{P}(\xi = \infty) = 1$

Now suppose the embedded MC is irreducible then $\sum_n \frac{1}{q_{X_{T_n}}} = \sum_{i \in E} N_i \frac{1}{q_i} \geq \sum_{i \in E} \frac{1}{q_i}$ where $N_i = \sum_{n=1}^{\infty} \mathbb{I}_{[X_{T_n}=i]} \geq 1$. Therefore $\sum_{i \in E} \frac{1}{q_i} = \infty$ implies that $\xi = \infty$ a.s..

This leads to the following definition:

Definition 6.1.5 *A CTMC is said to be* regular *if* $\mathbb{P}(\xi = \infty) = 1$. *If* $\mathbb{P}(\xi < \infty) > 0$ *then* ξ *is said to be an* explosion *time.*
A regular Markov process is also referred to as non-exlosive.

The discrete-time analog of a non-explosive process is $f_{ii} = \mathbb{P}(X_n = i \ \ eventually | X_0 = i) = 1$. Now if the chain is irreducible, then $f_{ii} = 1 \iff \mathbf{E}[N_i] = \infty$ and $\mathbb{P}(N_i = \infty) = 1$, i.e. regularity is equivalent to recurrence. A sufficient condition for a CTMC with $|E| = \infty$ to be regular is $\sup_i q_i = q < \infty$ since $\sum_i \frac{1}{q_i} \geq \sum_i \frac{1}{q} = \infty$.

We state and prove the result below.

Proposition 6.1.8 *Let* $\{X_t\}_{t \geq 0}$ *be a CTMC with generator* Q. *Then* X_t *is regular if any of the following hold:*

a) $\sup_i q_i < \infty$

b) The embedded process $Y_n = X_{T_n}$ *is recurrent.*

Proof: The proof of a) has been discussed above so we will restrict ourselves to the proof of b). Since Y_n is recurrent, let N_n^ibe the sequence of returns to state i. By definition $N_n^i \to \infty$ because i is recurrent.

Therefore noting that:

$$\xi \geq \sum_n S_{N_n^i}$$

and $N_n^i \to \infty$ it implies that $\xi = \infty$ or the process X_t is regular.

Remark 6.1.3 *It is useful to note that a) is true when* $|E| < \infty$. *b) is also true when* $|E| < \infty$ *and the chain* Y_n *is irreducible. Thus for finite chains we obtain positive recurrence (existence of a stationary distribution) if the state space is finite and irreducible.*

We now state the basic convergence and limit theorems for CTMCs.

Proposition 6.1.9 *Let $\{X_t\}$ be a CTMC with state space E and generator Q and suppose E is irreducible.*

Define:

$$\tau_i = \inf\{t > 0 : X_t = i | X_0 = i\}$$

Then:

i) If i is transient or null recurrent:

$$\lim_{t\to\infty} P_{ii}(t) = 0$$

ii) If j is positive recurrent then:

$$\lim_{t\to\infty} P_{ij}(t) = \pi_j = \frac{1}{q_j \mathbf{E}_j[\tau_j]}$$

iii) If$\{X_t\}$ is positive recurrent if and only if $\exists\ \pi_i > 0$ such that:

$$\begin{aligned} \pi_i &= \sum_{j\in E} \pi_j P_{ji}(t) \quad \forall t \geq 0 \\ \sum_{i\in E} \pi_i &= 1 \end{aligned}$$

We conclude with a statement of the ergodic theorem for CTMC. The proof is similar to the discrete case where we partition time into cycles where the chain regenerates and then use the SLLN on those independent components.

Proposition 6.1.10 *Let X_t be a positive recurrent CTMC then for any integrable function i.e. $\sum_{i\in E} |f(i)|\pi_i < \infty$,we have the following SLLN :*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X_s) ds = \mathbf{E}_\pi[f(X_0)] = \sum_{i\in E} f(i)\pi_i \tag{6.1. 12}$$

Canonical examples of Markov chains are Poisson processes which have stationary independent increments. Of course a Poisson process is by definition non-stationary in that $\lim_{t\to\infty} N_t = \infty$ and has no stationary distribution.

In applications we most often encounter a very important class of Markov chains referred to as birth-death processes. These CTMCs are defined through the Q matrix where:

$$\begin{aligned} q_{ij} &= p_i \quad j = i+1 \\ &= q_i \quad j = i-1, i \geq 1 \\ &= 0 \quad \textit{otherwise} \end{aligned}$$

It is easy to see that the CTMC possesses a stationary distribution iff $\sum_{n=1}^\infty \rho_n < \infty$ where $\rho_n = \prod_{k=0}^{n-1} \frac{p_k}{q_{k+1}}$ (use iii) of Proposition 6.1.9) . In particular when $p_n = \lambda$ and $q_n = \mu$ then a necessary and sufficient condition for the CTMC to be positive recurrent is $\rho = \frac{\lambda}{\mu} < 1$ and then the stationary distribution is given by:

$$\pi_k = \rho^k(1-\rho), \quad k = 0, 1, 2, \cdots ,$$

6.1.3 Uniformization of CTMC

Uniformization is a technique by which we construct a discrete-time MC whose distribution is the same as the CTMC. Note the embedded chain is also a DTMC but the stationary distribution differs from the stationary distribution of the CTMC.

Let us begin by considering the simple case first. If $q_i = q$ for all i in a CTMC then it is clear that the stationary distribution of the embedded chain given by:

$$\tilde{\pi}_i = \frac{\pi_i q_i}{\sum_j \pi_j q_j} = \frac{\pi q}{q \sum_j \pi_j} = \pi_i$$

or the two stationary distributions coincide. However if the q_i's are all the same then by construction the jump times form the points of a Poisson process , say N_t, with rate q and thus:

$$X_t = Y_{N_t}$$

where Y_n is the embedded chain.

We will now show that when $\sup_i q_i = q < \infty$ then we can perform a similar construction. Now we know that:

$$\frac{d}{dt}\pi(t) = \pi(t)Q$$

Define the matrix:

$$R = I + \frac{1}{q}Q$$

.

By construction the row sums of R are all 1, all elements are take values in $(0,1)$ and therefore R is a stochastic matrix that can be associated with a DTMC, say Y_n.

Therefore substituting we have:

$$\frac{d}{dt}\pi(t) = \pi(t)q(-I + R)$$

Since $\sup_i q_i < \infty$ we can write the solution as

$$\begin{aligned}
\pi(t) &= \pi(0)e^{Qt} \\
&= \pi(0)e^{q[-I+R]t} \\
&= \pi(0)e^{-qt}e^{qRt} \\
&= \pi(0)\sum_{n=0}^{\infty} e^{-qt}\frac{(qt)^n R^n}{n!} \\
&= \sum_{n=0}^{\infty} \hat{\pi}(n)e^{-qt}\frac{(qt)^n}{n!} \qquad (6.1.\ 13)
\end{aligned}$$

where $\hat{\pi}(n) = \pi(0)R^n$ denotes the probability distribution of Y_n starting with initial distribution $\pi(0)$.

Now define:

$$\hat{X}_t = Y_{N_t}$$

where N_t is a Poisson process with rate q. Then by definition: $\hat{X}_t = Y_n \quad T_n \leq t < T_{n+1}$ where T_n are the points of the Poisson process.

Now:

$$\begin{aligned}\mathbb{P}(\hat{X}_t = i) &= \sum_{k=0}^{\infty} \mathbb{P}(Y_k = i)\mathbb{P}(N_t = k) \\ &= \sum_{k=0}^{\infty} (\pi(0)R^k)_i e^{-qt}\frac{(qt)^k}{k!} \\ &= \mathbb{P}(X_t = i) \quad from\ (6.1.13)\end{aligned}$$

In other words we have constructed the process $\hat{X}_t$ from the DTMC process Y_n that has the same distributions as X_t for all t not only the stationary distribution.

Uniformization is thus a very useful technique to convert a CTMC and view it in terms of a DTMC with the same transient and stationary distribution. One important point to note, in the jump or embedded chain the transition from i to i is forbidden while in the uniformized chain there can be transitions from i to i as can be seen from the (i,i) element of R i.e. $R_{ii} = 1 - \frac{q_i}{q} > 0$.

6.1.4 Hitting times

Let $A \subset E$ be any subset of E.

Define $T^A = \inf\{t \geq 0 : X_t \in A\}$ denote the hitting or entrance time of X_t to A and $\tau_A = \inf\{n \geq 0 : X_{T_n} \in a\}$ denotes the hitting time to A for the jump chain.

Clearly, since the entrance time to A of X_t must coincide with a jump in the embedded or jump chain:

$$\mathbb{P}(\tau_A < \infty) = \mathbb{P}(T^A < \infty)$$

Define: $h_i^A = \mathbb{P}_i(T^A < \infty)$.

Proposition 6.1.11 *Let Q denote the generator of $\{X_t\}$. Then:*

$$\begin{aligned} h_i^A &= 1 \quad i \in A \\ \sum_{j\in E} q_{ij}h_j^A &= 0 \quad i \notin A \end{aligned} \qquad (6.1.\ 14)$$

bf Proof: We exploit the results for discrete-time chains to show this.

If $i \in A$ then by definition $\tau_A = 0$ and so there is nothing to prove.

If $i \notin A$ then:

$$\begin{aligned}\mathbb{P}_i(\tau_A < \infty) &= \sum_{j\in E} \mathbb{P}_i(\tau_A < \infty, j) \\ &= \sum_{j\in E} \mathbb{P}(\tau_A < \infty | X_0 = i, X_1 = j)P_{ij} \\ &= \sum_{j\in E, j\neq i} h_j^A P_{ij}\end{aligned}$$

where $P_{ij} = \frac{q_{ij}}{q_i}$ Therefore substituting we obtain: $\sum_{j\neq i} h_j^A \frac{q_{ij}}{q_i} = h_i^A$ or:

$$-q_i h_i^A + \sum_{j\neq i} h_j^A q_{ij} = 0$$

which gives the result noting that $q_i = -q_{ii}$.

Remark 6.1.4 *Note that $h_i^A = 1 \ \ \forall i$ is a solution to (6.1. 14). And thus if the chain is recurrent it is the only solution. On the other hand one can show that h_i^A is the minimal solution in that if there is any other solution to (6.1. 14) then it will be larger.*

In a similar way we can compute the mean hitting times.

Proposition 6.1.12 *Let $k_i^A = \mathbf{E}_i[T^A]$ where T^A is the hitting time. Then:*

$$\begin{aligned} k_i^A &= 0 \quad i \in A \\ q_i k_i^A &= 1 + \sum_{j \neq i} q_{ij} k_j^A \end{aligned} \qquad (6.1.\ 15)$$

Proof: The only twist in the proof is that we need to take into account the sojourn times.

If $i \in A$ the result is trivial. So for $i \notin A$, let $J)i$ denote the sojourn time in state i, we have:

$$\begin{aligned} \mathbf{E}_i[T^A] = k_i^A &= \mathbf{E}_i[J_i + \sum_{j \in E} \mathbf{E}_i[T^A - J_i | X_{T_1} = j, X_0 = i] \\ &= \mathbf{E}_i[J_i] + \sum_{j \neq i, j \in E} \mathbf{E}_j[T^A] \frac{q_{ij}}{q_i} \\ &= \frac{1}{q_i} + \frac{1}{q_i} \sum_{j \neq i} q_{ij} k_j^A \end{aligned}$$

whence the stated result follows.

6.1.5 Reversibility

We conclude our discussion of CTMCs with the notion of reversibility of Markov chains. This property plays a very important role in the structure and behavior of many useful queueing models.

Recall a MC remains a MC in reverse time too. However, even if a MC is homogeneous in forward time the reverse time MC need not be homogeneous. Let us show this result in the CTMC case.

Let T be fixed and define $Y_t = X_{T-t}, \quad t \geq 0$. Then Y_t is a process that evolves in reverse time. Let us show that $\{Y_t\}$ is Markov. This follows from the fact that given $Y_t = X_{T-t}$ the events $A \in \sigma\{Y_s, s < t\} = \sigma\{X_{T-s}, s < t\}$ and $B \in \sigma\{Y_u, u > t\} = \sigma\{X_{T-u}, u > t\}$ are conditionally independent belonging to the past and future of $T - t$ from the assumption that X_t is Markov. It remains to calculate the transition probability matrix of Y_t.

$$\begin{aligned} \tilde{P}_{ij}(t,s) = \mathbb{P}(Y_{t+s} = j | Y_s = i) &= \frac{\mathbb{P}(Y_{t+s} = j, Y_s = i)}{\mathbb{P}(Y_s = i)} \\ &= \frac{\mathbb{P}(Y_s = i | Y_{t+s} = j) \mathbb{P}(Y_{t+s} = j)}{\mathbb{P}(Y_s = i)} \\ &= \frac{\mathbb{P}(X_{T-s} = i | X_{T-(t+s)} = j) \mathbb{P}(X_{T-(t+s)} = j)}{\mathbb{P}(X_{T-s} = i)} \\ &= P_{ji}(t) \frac{\pi_j(t+s)}{\pi_i(s)} \end{aligned}$$

It is important to note that the rhs is a function of t and s and not on their difference and thus the reverse chain in inhomogeneous.

Now suppose $\{X_t\}$ is stationary. Then $\pi_i(t) = \pi_i$ where $\pi = (\pi_0, \pi_1, \cdots, \pi_i, \ldots)$ is the stationary distribution . In that case:

$$\tilde{P}_{ij}(t,s) = P_{ji}(t)\frac{\pi_j}{\pi_i}$$

or the reverse chain is also homogeneous.

Let us also compute its generator.

Let $\tilde{Q}$ denote the generator then:

$$\frac{d}{dt}\tilde{P}_{ij}(t)|_{t=0} = \tilde{q}_{ij} = \frac{\pi_j}{\pi_i}\sum_k q_{jk}P_{ki}(t)|_{t=0} = \frac{\pi_j}{\pi_i}q_{ji}$$

Clearly:

$$\sum_j \tilde{q}_{ij} = \sum_j \frac{\pi_j}{\pi_i}q_{ji} = 0$$

since by definition $\sum_j \pi_j q_{ji} = 0$ by definition of the stationary distribution. So $\tilde{Q}$ is a bona fide rate matrix.

Let us show that π remains the stationary distribution of Y_t.

We know by definition that:

$$\begin{aligned}(\pi Q)_j = \sum_i \pi_i q_{ij} = 0 &= \sum_i \pi_i \frac{\pi_j}{\pi_i}\tilde{q}_{ji} \\ &= \pi_j \sum_i \tilde{q}_{ji} = 0 \\ &= (\pi\tilde{Q})_j\end{aligned}$$

establishing that π is the stationary distribution for Y_t.

We can collect all these results and state stem as a proposition:

Proposition 6.1.13 *Let $\{X_t\}$ be a stationary CTMC on E that is irreducible with stationary distribution π. If the generator Q satisfies the detailed balance equation given by:*

$$\pi_i q_{ij} = \pi_j q_{jI} \tag{6.1. 16}$$

Then the reversed process $\{X_{-t}\}$ properly modified to be right-continuous is distributionally equivalent to $\{X_t\}$.

Proof: By the definition of the generator of the reverse process $\tilde{Q}$ we see that $\tilde{Q} = Q$ and therefore both Markov processes are stationary with the same generator and are thus distributionally equivalent.

Reversibility of the underlying CTMC has many nice properties that follow as a result of the detailed balance equation (6.1. 16).

Lemma 6.1.3 *Let Q be the generator of a reversible CTMC on E. Then, $\exists\ g_i > 0$ such that:*

$$\frac{q_{ij}}{q_{ji}} = \frac{g_j}{g_i} \tag{6.1. 17}$$

proof: Since the CTMC is reversible it must satisfy the detailed balance: we can take $g_i = C\pi_i$ for any constant C and hence we get the required result. On the other hand, if $\frac{q_{ij}}{q_{ji}}$ satisfies the above relation then g_i will satisfy the detailed balance equation and we can then normalize the $g_.$'s to obtain the stationary distribution.

We conclude our study of CTMCs by pointing two important properties associated with reversible processes. The first is the notion of partial balance.

Proposition 6.1.14 *Let$\{X_t\}$ be a reversible Markov process. Let $A \subset E$ be any closed*[1] *subset of E and denote $A^c = E - A$. Then the process satisfies the following partial balance equations :*

$$\sum_{j \in A^c} \sum_{i \in A} \pi_i q_{ij} = \sum_{j \in A^c} \sum_{i \in A} q_{ji} \pi_j \tag{6.1. 18}$$

In order to check for reversibility it appears that one needs to compute the stationary distribution π_i. This it turns out is not necessary. The result is called the Kolmogorov loop criterion.

Definition 6.1.6 *Given a collection of states $i, i_1, \cdots, i_n, i$ that defines a closed path from i to i we say that Kolmogorov's loop criterion is satisfied if:*

$$q_{ii_1} q_{i_1 i_2} \cdots q_{i_n i} = q_{ii_n} q_{i_n i_{n-1}} \cdots q_{i_1 i} \tag{6.1. 19}$$

Proposition 6.1.15 *A stationary CTMC is reversible if and only if $\forall\ i, j \in E$, $q_{ij} = 0 \implies q_{ji} = 0$ and Kolmogorov's loop criterion is satisfied for every closed path.*

A simple application of these results relates to truncated Markov Chains.

Example 1:

Let $\{X_t\}$ be a reversible MC and let $A \subset E$. Define the truncated markov process X_t^A as one with the following generator:

$$\begin{aligned} \bar{q}_{ij} &= q_{ij} \quad if \quad i, j \in A \\ \bar{q}_{ii} &= -\sum_{i,j \in A, j \neq i} q_{ij} \end{aligned}$$

Then X_t^A is a Markov is a CTMC on A with stationary distribution denoted by π^A given by:

$$\pi_i^A = \frac{\pi_i}{\sum_{j \in A} \pi_j} \quad i \in A \tag{6.1. 20}$$

Reversibility is an idea that translates to discrete-time Markov chains too. Indeed we can state the following result.

Proposition 6.1.16 *Let $\{X_n\}$ be a discrete time MC on (E, P) with stationary distribution π. If:*

$$\pi_i P_{ij} = \pi_j P_{ji}$$

Then $\{X_n\}$ is reversible and the process $\{X_{-n}\}_{n \geq 0}$ has the same probabilistic behavior as $\{X_n\}$.

[1] By a closed set it means the states within that set can communicate with each other without the need to exit that set.

A simple example of a reversible MC is the random graph model below.

A random graph model

Consider a finite connected random graph with n nodes. A pair of nodes (i, j) are said to be connected if there is a link or arc between them. Let w_{ij} be a positive weight on arc (i, j) with $w_{ij} = w_{ji}$. Define $w - ij = 0$ if there is no arc between i and j.

Define a random walk on the graph as the probability of going from $i \to j$ as $\frac{w_{ij}}{\sum_k w_{ik}}$. Then since the P matrix is completely symmetric one would expect the MC to be reversible. Let π denote the stationary distribution.

Indeed for all i and j we have:

$$\frac{\pi_i}{\sum_k w_{ik}} = \frac{\pi_j}{\sum_k w_{jk}} = C$$

and therefore

$$\pi_i = C \sum_k w_{ik}$$

.

Noting $\sum_{i=1}^n \pi_i = 1$ we obtain $C = [\sum_k \sum_i w_{ik}]^{-1}$ and hence:

$$\pi_i = \frac{\sum_k w_{ik}}{\sum_k \sum_i w_{ik}}$$

If $w_{ij} = 1$ then it corresponds to a random walk on a random graph where the probability of going to a neighboring node is uniform.

This concludes the key results about CTMCs that will be of use when we study queueing systems.