OM9103: Stochastic Process

Lecture 7&8&9:
Markov Chains

A. Introduction

Consider some stochastic process with *discrete* time parameter: $\{X_n: n \in \mathbf{N}\}$. The set of values that the random variables X_n can assume is called the ***state space*** of the stochastic process, which can be finite, countable, or uncountable.

Initially, we will restrict ourselves to finite or countable state spaces. Without loss of generality, we can then assume that the state space is $\{0, 1, \ldots, N\}$ or $\mathbf{N}$. If $X_n = i$, we say *the process is in state i at time n*.

The dynamics of the stochastic process are governed by ***transition probabilities***:

$$\Pr(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0)$$

If these transition probabilities only depend on i and j, we call the process a *Markov chain*. In other words, the conditional probability distribution of X_{n+1} *given* the values of X_n, …, X_0 is only dependent on X_n. We denote

$$P_{nij} = \Pr(X_{n+1} = j \mid X_n = i)$$

The property that the future of the process is *conditionally* independent of the past *given* the present is called the *Markovian property*. We distinguish between:

- (homogeneous) Markov chains: $P_{nij} = P_{ij}$
- nonhomogeneous Markov chains

We will mainly limit ourselves to (*homogeneous*) *Markov chains*. The transition probabilities can be arranged in a matrix, say, P, with elements P_{ij}:

$$P = \begin{pmatrix} P_{00} & P_{01} & \cdots \\ P_{10} & P_{11} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

clearly, for fixed i, the vector $(P_{i0}, P_{i1}, \ldots)$ is a probability distribution. Put differently, the transition probability matrix is a *stochastic matrix*, i.e.,

- $P_{ij} \geq 0 \quad i, j = 0,1,\ldots$
- $\sum_{j=0}^{\infty} P_{ij} = 1 \quad i = 0,1,\ldots$

A Markov chain is often represented by a graph:

- the nodes of the graph corresponds to the states of the Markov chain;
- there is a *directed* arc between nodes i and j if and only if $P_{ij} > 0$.

<u>Example 1</u>: **(s, S) Inventory System with Periodic Review**

Consider a single-item inventory system that we review periodically. Demands for the item in successive periods are *independent and identically distributed random variables* D_n, $n = 0, 1, \ldots$:

$$\Pr(D_n = j) = \varphi_j \quad j = 0,1,\ldots$$

We can review the inventory position at the end of each period. Delivery of replenishment is instantaneous and available at the start of the next period. We use **an (*s, S*)-policy** ($0 \le s < S$): if the inventory position is $\le s$, we order to bring it up to S.

What will we do in case of stockout? We can assume either (a) *lost sales* – any demand in excess of inventory is lost; or (b) *backordering* – any demand in excess of inventory is satisfied immediately after replenishment. Our first Markov chain model for this system will apply to both cases.

Markov Chain Model I

For periods $n = 0, 1, \ldots$, let X_n denote the inventory level at the beginning of the n^{th} period (just after delivery of any replenishment order). The stochastic process $\{X_n: n \in \mathbf{N}\}$ has state space $\{s + 1, \ldots, S\}$. This process is a Markov chain because of two key factors:

- the demands are independent;
- the replenishment policy is given by

$$X_{n+1} = \begin{cases} X_n - D_n & \text{if } X_n - D_n > s \\ S & \text{if } X_n - D_n \le s \end{cases}$$

To determine the transition probabilities P_{ij} we need to distinguish between the cases: $j \ne S$ and $j = S$.

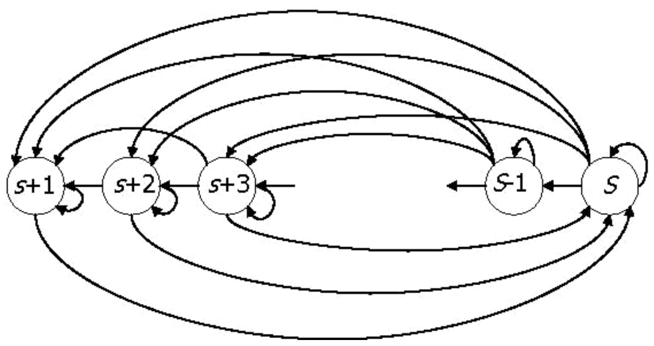
If $j \ne S$,

$$P_{ij} = P(D = i - j) = \begin{cases} \varphi_{i-j} & \text{if } j = s+1,\ldots,\min(i, S-1);\ i = s+1,\ldots,S \\ 0 & \text{if } j = i+1,\ldots,S-1;\ i = s+1,\ldots,S \end{cases}$$

If $j = S$,

$$P_{ij} = \begin{cases} P(D = 0 \text{ or } D \ge i - s) = \varphi_0 + \sum_{k=i-s}^{\infty} \varphi_k & \text{if } j = S;\ i = S \\ P(D \ge i - s) = \sum_{k=i-s}^{\infty} \varphi_k & \text{if } j = S;\ i = s+1,\ldots,S-1 \end{cases}$$

A graphical representation of the above Markov chain is as follows:



Markov Chain Model II

For periods n = 0, 1, …, let Y_n denote the inventory level at the end of the n^{th} period (just before ordering). The analysis of the stochastic process $\{Y_n, n = 0, 1, \ldots\}$ depends on stockout policy/assumptions.

<u>Case 1 – lost sales</u>: The state space of $\{Y_n, n = 0, 1, \ldots\}$ is $\{0, 1, \ldots, S\}$. This process is a Markov chain because

$$Y_{n+1} = \begin{cases} \max\{0, Y_n - D_{n+1}\} & \text{if } Y_n > s \\ \max\{0, S - D_{n+1}\} & \text{if } Y_n \le s \end{cases}$$

To determine the transition probabilities Q_{ij} we need to distinguish the cases $i \le s$ and $i > s$.

If $i \le s$, we have

$$Q_{ij} = \begin{cases} P(D \ge S) = \sum_{k=S}^{\infty} \varphi_k & \text{if } j = 0;\ i = 0, \ldots, s \\ P(D = S - j) = \varphi_{S-j} & \text{if } j = 1, \ldots, S;\ i = 0, \ldots, s \end{cases}$$

If $i > s$, we have

$$Q_{ij} = \begin{cases} \Pr(D \ge j) = \sum_{k=i}^{\infty} \varphi_k & \text{if } j = 0;\ i = s+1, \ldots, S \\ \Pr(D = i - j) = \varphi_{i-j} & \text{if } j = 1, \ldots, i;\ i = s+1, \ldots, S \\ \Pr(D = i - j) = 0 & \text{if } j = i+1, \ldots, S;\ i = s+1, \ldots, S \end{cases}$$

<u>Case 2 – backordering</u>: In this case, the state space of $\{Y_n, n = 0, 1, \ldots\}$ is $\{\ldots, -1, 0, 1, \ldots, S\}$. This process is a Markov chain because

$$Y_{n+1} = \begin{cases} Y_n - D_{n+1} & \text{if } Y_n > s \\ S - D_{n+1} & \text{if } Y_n \le s \end{cases}$$

And the transition probabilities are:

$$Q_{ij} = \begin{cases} \Pr(D = S - j) = \varphi_{S-j} & \text{if } j \le S;\ i \le s \\ \Pr(D = i - j) = \varphi_{i-j} & \text{if } j \le i;\ s < i \le S \\ \Pr(D = i - j) = 0 & \text{if } i < j \le S;\ s < i \le S \end{cases}$$

Example 2 (Example 4.1(A)): **M/G/1 Queue**

Consider a single-server system where

- customers arrive according to a Poisson process with parameter λ;
- service times are independent, and distributed according to G;
- an arriving customer waits until server is available.

Let $X(t)$ denote the number of customers in the system at time t. The stochastic process $\{X(t), t \geq 0\}$ does not possess the Markovian property because the future depends on time since start of current service.

- If we knew the number in the system at time t, then, to predict future behavior, where we would not care how much time has lapsed since the last arrival (since the arrival process is memoryless), we should care how long the person in service has already been there (since the service distribution is not memoryless).

However, we can study certain properties of this system through a so-called ***embedded Markov chain***.

Suppose that we only observe the system when a customer departs. Let X_n denote the number of customers in the system just after departure of the nth customer. It is worthwhile to highlight that

- The events (sometimes called ***epochs***) $n = 0, 1, \ldots$ are not equally spaced in time;
- In fact, the time between epochs is random.

We will study the stochastic process $\{X_n, n = 0, 1,\ldots\}$. Suppose that $X_n = 0$, i.e., the queue is empty after departure of the nth customer.

- The system is empty until arrival of $(n+1)^{st}$ customer.
- Let Y_n denote the number of customers arriving while this customer receives service.
- The number of customers in the system after its departure is Y_n.

Suppose that $X_n > 0$. The nth departure leaves behind X_n customers – of which one enters service and other $X_n - 1$ wait in line in addition to any arrivals during the service time of the $(n+1)^{st}$ customer. Hence, the number of customers in the system after n^{th} departure is $X_n + Y_n - 1$.

Therefore, we conclude that $\{X_n, n = 0, 1,\ldots\}$ is a Markov chain:

$$X_{n+1} = \begin{cases} Y_n & \text{if } X_n = 0 \\ X_n + Y_n - 1 & \text{if } X_n > 0 \end{cases}$$

To determine the transition probabilities, we need first to determine the distribution Y_n. Note that this distribution is independent of n. If the service time were fixed, say equal to t, Y_n is Poisson distributed with parameter λt. Let T be the service time. Then,

$$\varphi_j = \Pr(Y_n = j) = \int_0^\infty \Pr(Y_n = j \mid T = t)\, dG(t) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^j}{j!}\, dG(t) \quad \text{for } j = 0, 1, \ldots$$

We use this to determine the transition probabilities:

$$P_{ij}=\begin{cases}\Pr(Y_n=j)=\varphi_j & \text{if } i=0;\ j=0,1,\ldots\\ 0 & \text{if } i=1,2,\ldots;\ j=0,1,\ldots,i-2\\ \Pr(i+Y_n-1=j)=\varphi_{j-i+1} & \text{if } i=1,2,\ldots;\ j=i-1,i,\ldots\end{cases}$$

Example 3 (Example 4.1(B)): **G/M/1 Queue**

Suppose that customers arrive at a single-server service center in accordance with an arbitrary renewal process having interarrival distribution G. Suppose further that the service distribution is exponential with rate μ.

Let X_n be the number of customers in the system as seen by the nth arrival. It is clear that the process $\{X_n,\ n = 1, 2,\ldots\}$ is a Markov chain. Let us find the transition probability matrix of this Markov chain. Note that as long as there are customers to be served, the number of services in any length of time t is a Poisson random variable with mean μt. This is true since the time between successive services is exponential, implying that the number of services coincides with a Poisson process. Thus,

$$P_{i,i+1-j}=\int_0^\infty e^{-\mu t}\frac{(\mu t)^j}{j!}\,dG(t),\ j = 0, 1, 2, \ldots, i,$$

which follows from the following observations:

- If an arrival finds i customers in the system, then the next arrival will find $(i+1)$ minus the number served.
- By conditioning on the time between successive arrivals, we will get the right-hand side expression.

Alternatively,

$$P_{i,k}=\int_0^\infty e^{-\mu t}\frac{(\mu t)^{i+1-k}}{(i+1-k)!}\,dG(t),\ \text{for } k = 1, \ldots, i+1.$$

Note that P_{i0} is the probability that *at least* i+1 Poisson events occur in a random length of time having distribution G. Hence,

$$P_{i0}=\int_0^\infty \sum_{k=i+1}^{\infty} e^{-\mu t}\frac{(\mu t)^k}{k!}dG(t),\ \ i\ge 0.$$

Example 4 (Example 4.1(C)): **The General Random Walk**

Let X_i, $i \geq 1$, be a sequence of i.i.d. random variables with the

$$P(X_i = j) = a_j,\ \ j = 0, \pm1, \pm2, \ldots$$

Let $S_0 = 0$ and

$$S_n=\sum_{i=1}^{n} X_i.$$

Then $\{S_n,\ n = 0, 1, \ldots\}$ is called ***the general random walk***.

Claim: $\{S_n,\ n = 0, 1, \ldots\}$ is a Markov chain with the transition probabilities given by

$$P_{ij} = \Pr(S_{n+1} = j \mid S_n = i) = \Pr(X_{n+1} = j - i) = a_{j-i}.$$

Example 5 (Example 4.1(D)): **Absolute Value of Simple Random Walk**

A simple random walk corresponds to the following distribution for X_i: for some p, $0 < p < 1$, $P(X_i = 1) = p$ and $P(X_i = -1) = 1 - p \equiv q$. With $S_n = \sum_{i=1}^{n} X_i$, the process $\{S_n, n = 0, 1, \ldots\}$ is called ***a simple random walk***.

Claim: $\{|S_n|, n = 0, 1, \ldots\}$ is a Markov chain with the following transition probabilities:

$$P_{01} = 1;$$

$$P_{i,i+1} = \frac{p^{i+1} + q^{i+1}}{p^i + q^i} = 1 - P_{i,i-1} \quad \text{for } i > 0.$$

The key to derive these transition probabilities is to show (Proposition 4.1.1, p.166):

$$\Pr(S_n = i \,\|\, S_n | = i, | S_{n-1} | = i_{n-1}, \ldots, | S_1 | = i_1) = \frac{p^i}{p^i + q^i}.$$

B. Multi-Step Transitions

B.1 Chapman-Kolmogorov Equation

So far, we have focused on the so-called *one-step transitions*, characterized by the transition probability matrix P. We can also consider n-step transitions:

$$P_{ij}^{(n)} = \Pr(X_{m+n} = j \mid X_m = i) \quad \text{for } n = 0, 1, 2, \ldots$$

Clearly, for $n = 0$, $P_{ij}^{(0)} = 1_{\{i = j\}}$, and for $n = 1$, $P_{ij}^{(1)} = P_{ij}$. We can recursively determine the $(n+1)$-step transition probabilities from the n-step transition probabilities:

$$P_{ij}^{(n+1)} = \Pr(X_{m+n+1} = j \mid X_m = i)$$

$$= \sum_{k=0}^{\infty} \Pr(X_{m+n+1} = j \mid X_{m+n} = k) \cdot \Pr(X_{m+n} = k \mid X_m = i) = \sum_{k=0}^{\infty} P_{kj} P_{ik}^{(n)}$$

In matrix notation, this means

$$P^{(n+1)} = P^{(n)} \cdot P$$

In a similar way, we can show that

$$P^{(n+1)} = P \cdot P^{(n)}$$

In general, we have the following identity, known as *Chapman-Kolmogorov Equation*,

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)} \quad \text{for all } n \geq 0 \text{ and } m \geq 0.$$

Furthermore, by induction, we can easily show that

$$P^{(n)} = P^n, \quad \text{for } n = 0, 1, 2, \ldots$$

B.2 Communication, Classes and Irreducibility

Definition (*Accessibility*): We say that state j is ***accessible*** from state i if there exists some $n \geq 0$ such that $P_{ij}^{(n)} > 0$.

In terms of graph representation of the Markov chain, state j is accessible from state i if there exists *directed path* from i to j.

Definition (*Communication*): We say states i and j ***communicate***, denoted by $i \leftrightarrow j$, if (a) i is accessible from j and (b) j is accessible from i.

In terms of graph representation of the Markov chain, states i and j communicate if there exists a *directed path* from i to j and a *directed path* from j to i.

Proposition 4.2.1: Communication is *an equivalence relation*, that is,

- $i \leftrightarrow i$;
- $i \leftrightarrow j$ is equivalent to $j \leftrightarrow i$.
- if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$.

The third property can be proven formally as follows. Since $i \leftrightarrow j$ and $j \leftrightarrow k$, there exist values m and n such that $P_{ij}^{(m)} > 0$ and $P_{jk}^{(n)} > 0$. Thus, we have

$$P_{ik}^{(m+n)} = \sum_{l=0}^{\infty} P_{il}^{(m)} P_{lk}^{(n)} \geq P_{ij}^{(m)} P_{jk}^{(n)} > 0.$$

Definition (*Classes*): This equivalence relation decomposes the state space into a number of ***classes***:

- any two states in the same class communicate;
- any two states in different classes do not communicate.

Note that this decomposition does not correspond to *connected components* of the graph representation of the Markov chain. In particular, it is possible that j is accessible from i even if the states i and j are in different classes. However, in this case, i cannot be accessible from j.

Definition (*Irreducibility*): If *all* states of a Markov chain communicate (i.e., these is only one class), the Markov chain is called ***irreducible***.

Remarks:

- Irreducibility is a very important class of Markov chains.
- For any Markov chain that has a finite number of classes, we will eventually end up in a class from which we can never leave.

B.3 Transience and Recurrence

Consider two states of a Markov chain, i and j. Given that we start in state i, what is the probability that we will ever visit j? That is, we are interested in

$$f_{ij} = \Pr(\exists n \ni X_n = j \mid X_0 = i)$$

Note that $f_{ij} > 0$ if and only j is accessible from i. We can characterize this probability as follows:

$$\begin{aligned} f_{ij} &= \Pr(\exists n \ni X_n = j \mid X_0 = i) \\ &= \Pr(\exists n \ni X_n = j, X_k \neq j \text{ for } k = 1, \ldots, n-1 \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} \Pr(X_n = j, X_k \neq j \text{ for } k = 1, \ldots, n-1 \mid X_0 = i) = \sum_{n=1}^{\infty} f_{ij}^{(n)} \end{aligned}$$

where $f_{ij}^{(n)}$ is the probability that, starting in i, the first transition into state j occurs at time n. It is evident that as long as j is accessible from i ($i \neq j$), $f_{ij} > 0$.

The following result establishes the link between $f_{ij}^{(n)}$ and $P_{ij}^{(n)}$.

Proposition: For all i, j, and $n \geq 0$,

$$P_{ij}^{(n)} = \sum_{k=0}^{n} f_{ij}^{(k)} P_{jj}^{(n-k)}$$

Proof: First, by definition, we have

$$f_{ij}^{(n)} = \Pr(X_n = j, X_v \neq j, v = 1, \ldots, n\text{-}1 \mid X_0 = i),$$

that is, it is the probability that, starting from state i, the first visit to state j occurs at nth transition.

We prove the result by decomposing the event from which $P_{ij}^{(n)}$ is computed according to the time of the *first* visit to state j. Let E_k be the event that the first visit to state j occurs at k-transition. Note that for $1 \leq k \leq n$,

$$P(E_k) = \Pr(\text{first visit to } j \text{ is at } k\text{th transition} \mid X_0 = i) P(X_n = j \mid X_k = j)$$
$$= f_{ij}^{(k)} P_{jj}^{(n-k)}$$

Because the events E_k's are mutually exclusive, it follows that

$$P_{ij}^{(n)} = P(X_n = j \mid X_0 = i) = \sum_{k=1}^{n} P(E_k) = \sum_{k=1}^{n} f_{ij}^{(k)} P_{jj}^{(n-k)} = \sum_{k=0}^{n} f_{ij}^{(k)} P_{jj}^{(n-k)}$$

since $f_{ij}^{(0)} = 0$ by definition. □

Definition (*Recurrence & Transience*): Given that we state in state j, the probability that we will ever return to state j is equal to f_{jj}.

- If $f_{jj} = 1$, we say that the state j is ***recurrent***.
- If $f_{jj} < 1$, we say that the state j is ***transient***.

Consider a transient state j, i.e., $f_{jj} < 1$. This implies that each time the process enters state i there will be a positive probability, $1 - f_{jj}$, that it will ***never*** enter that state again. Therefore, starting in state j, the probability the process will be in state j exactly n time periods is equal to $f_{jj}^{n-1}(1-f_{jj})$, $n \geq 1$. In other words, *if the state j is transient, then, starting in state j, the number of time periods that the process will be in state j has a geometric distribution with a finite mean* $1/(1-f_{jj})$.

Consider a recurrent state j, i.e., $f_{jj} = 1$. How often will we return to this state? Whether we visit this state, we will return there with probability one. Thus, starting in a recurrent state, the number of expected number of visits to that state is ∞.

The above two observations can be used to derive another characterization of whether a state is transient or recurrent.

Let $I_n = 1$ if $X_n = j$, 0 otherwise. Then the total number of visits to j is equal to $\sum_{n=1}^{\infty} I_n$ and

$$E\left(\sum_{n=0}^{\infty} I_n \mid X_0 = j\right) = \sum_{n=0}^{\infty} E\left(I_n \mid X_0 = j\right) = \sum_{n=0}^{\infty} P_{jj}^{(n)}$$

Now recall that

$$E\left(\sum_{n=0}^{\infty} I_n \mid X_0 = j\right) = \infty$$

if and only if j is recurrent. Thus, we have the following result:

Proposition 4.2.3: $f_{jj} = 1$ (that is, j is recurrent) if and only if

$$\sum_{n=1}^{\infty} P_{jj}^{(n)} = \infty$$

Any property that is shared by *all* or *none* of the states in a class is called a ***class property.***

Corollary 4.2.4: Recurrence is a class property, that is, if i is recurrent and $i \leftrightarrow j$, then j is recurrent.

Suppose that state i is recurrent and $i \leftrightarrow j$. We know that there exist values m and n such that $P_{ji}^{(m)} > 0$ and $P_{ij}^{(n)} > 0$. This implies that

$$\sum_{s=0}^{\infty} P_{jj}^{(s)} \geq \sum_{s=0}^{\infty} P_{jj}^{(m+s+n)} \geq P_{ji}^{(m)} \cdot \sum_{s=0}^{\infty} P_{ii}^{(s)} \cdot P_{ij}^{n} = \infty$$

Hence j is recurrent as well. This shows that recurrence is a class property.

Corollary 4.2.5: If $i \leftrightarrow j$ and j is recurrent, then $f_{ij} = 1$.

Let $N_j(t)$ be the number of transitions into j by time t. If j is recurrent and $X_0 = j$, then as the process probabilistically starts over upon transitions into j. This implies that $\{N_j(t), t \geq 0\}$ is a renewal process with interarrival distribution $\{f_{jj}^{(n)}, n \geq 1\}$. If $X_0 = i$, and $i \leftrightarrow j$, then $\{N_j(t), t \geq 0\}$ is a delayed renewal process with initial interarrival distribution $\{f_{ij}^{(n)}, n \geq 1\}$ and subsequent interarrival (time) distribution $\{f_{jj}^{(n)}, n \geq 1\}$.

The expected number of transitions needed to return to some state j is equal to

$$\mu_{jj} = \begin{cases} \infty & \text{if } j \text{ is transient} \\ \sum_{n=1}^{\infty} n f_{jj}^{(n)} & \text{if } j \text{ is recurrent} \end{cases}$$

Remark:

- Note that μ_{jj} is infinite if j is transient; but the converse is not true!

Definition: (*Null and Positive Recurrence*):

- If j is recurrent and $\mu_{jj} = \infty$, we call the state ***null recurrent***.
- If j is recurrent and $\mu_{jj} < \infty$, we call the state ***positive recurrent***.

Claim: Null recurrence and positive recurrence are class properties.

Exercise: Find the class(es) of the following two Markov chains:

(i) The chain has three states 0, 1 and 2 having transition probability matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

(ii) The chain has three states 0, 1, 2 and 3 having transition probability matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

B.4 Periodicity

Definition (*Period*): A state i of a Markov chain is said to be have ***period*** $d(i)$ if

- $P_{ii}^{(n)} > 0$ implies that n is divisible by $d(i)$, and
- $d(i)$ is the largest integer having this property.

A state with period 1 is called ***aperiodic***.

Claim: Periodicity is a class property, i.e., all states in a class have the same period.

Let m and n be such that $P_{ji}^{(n)}P_{ij}^{(m)} > 0$ and let $P_{ii}^{(s)} > 0$. Then, we must have

$$P_{jj}^{(n+m)} \geq P_{ji}^{(n)}P_{ij}^{(m)} > 0 \text{ and } P_{jj}^{(n+m+s)} \geq P_{ji}^{(n)} P_{ii}^{(s)} P_{ij}^{(m)} > 0$$

Thus $d(j)$ divides both $n + m$ and $n + s + m$, and therefore, also *any* s for which $P_{ii}^{(s)} > 0$. Hence $d(j)$ divides $d(i)$. The reverse argument yields that $d(i)$ divides $d(j)$. Consequently, $d(i) = d(j)$.

Definition (*Ergodic*): A state that is positive recurrent as well as aperiodic is called ***ergodic***. A Markov chain is said to be ***ergodic*** if all states are ergodic.

B.5 Markov Chains with Finite State Space

Proposition: Consider a Markov chain with a finite state space. Then this Markov chain must have at least one recurrent state!

Proof: Recall that the number of visits to a transient state is finite with probability one. This implies that a transient state will only be visited a finite number of times. Now suppose the result is false: all states for a finite-state Markov chain are transient. Then, after a finite amount of time, say, T_0, state 0 will never be visited; and after a finite amount of time, say T_1, state 1 will never be visited; and so on. Let $T = \max(T_0, T_1, \ldots, T_M)$. Then it is evident that after T, no states will be visited. This is of course cannot be true since the process must be at one of these states all the time.

Corollary:

(a) In an irreducible Markov chain with a finite state space, all states are either positive or null recurrent.

(b) If a Markov chain is also *aperiodic*, then

- the Markov chain is *ergodic* (all states are positive recurrent); or
- all states are null recurrent.

C. Limiting Behavior

C.1 Limiting Transition Probabilities

It seems intuitive that, if a Markov chain is irreducible, the effect of the initial state should wear off after a large number of transitions. If a Markov chain is not irreducible, the starting state may preclude some states from ever being visited, and thus clearly influences the long-run behavior of the chain.

Fix state j and let $N_j(t)$ denote the number of transitions into state j by time t. If $X_0 = j$, then $\{N_j(t), t \geq 0\}$ is *a renewal process* with interarrival distribution

$$\Pr(T = n) = f_{jj}^{(n)} \text{ for } n = 1, 2, \ldots$$

Note that this distribution is lattice with period $d(j)$. Then Blackwell's Theorem implies that

$$\lim_{n\to\infty} P_{jj}^{(nd(j))} = \frac{d(j)}{\mu_{jj}}$$

This means that, starting in state j, then long-run fraction of time that the chain is in state j is equal to $1/\mu_{jj}$.

If $X_0 = i$ and $i \leftrightarrow j$, then $\{N_j(t), t \geq 0\}$ is *a delayed renewal process*. If j is *aperiodic* and $i \leftrightarrow j$, Blackwell's Theorem for delayed renewal process says

$$\lim_{n\to\infty} P_{jj}^{(n)} = \frac{1}{\mu_{jj}} \equiv \pi_j$$

Clearly, $\pi_j > 0$ if and only if $\mu_{jj} < \infty$, i.e., if and only if j is positive recurrent.

The value π_j can be interpreted as the long-run fraction of time spent in state j. If the values π_j form a probability distribution on the state space, π can be interpreted as the *limiting distribution* of the Markov chain, provided the Markov chain is *aperiodic*. Unfortunately, the values of $f_{jj}^{(n)}$, and therefore μ_{jj}, are often very difficult to compute, we need to find another way of computing the values of π_j.

C.2 Stationary Distribution

Until now, we have assumed that the initial state is fixed. Now consider the case in which the initial state is chosen randomly:

$$\Pr(X_0 = j) = \alpha_j, \;\; j = 0, 1, \ldots$$

We then have the probability distribution of X_1 is given by

$$\Pr(X_1 = j) = \sum_{i=0}^{\infty} \Pr(X_1 = j \mid X_0 = i) \cdot \Pr(X_0 = i) = \sum_{i=0}^{\infty} P_{ij}\alpha_i$$

Now if

$$\Pr(X_1 = j) = \alpha_j, \;\; j = 0, 1, \ldots$$

we say that α is a ***stationary distribution*** of the Markov chain. That is, if X_0 is distributed according to α, X_n will have this distribution for all n.

Put it differently, a vector α, satisfying (1) $P^T\alpha = \alpha$, (2) $e^T\alpha = \mathbf{1}$, and (3) $\alpha \geq 0$, is a *stationary probability distribution* for the Markov chain.

Theorem 4.3.3: Consider an *irreducible aperiodic* Markov chain. Then we have two possible scenarios:

1. Either the states are all states are transient or null recurrent; in this case, $P_{ij}^n \to 0$ as $n \to \infty$ for all i and j and there exists no stationary distribution.
2. All states are positive recurrent (known as an ***ergodic Markov chain***), that is, $\pi_j = \lim_{n\to\infty} P_{ij}^n > 0$. In this case, $\{\pi_j, j = 0, 1, 2\ldots\}$ is a stationary distribution which specifies a *unique* limiting probabilities:
$$\pi_j = \sum_i \pi_i P_{ij}, \forall j; \quad \sum_j \pi_j = 1$$

Proof: Refer to textbook, p. 175-176.

Remark:

- π_j must be interpreted as the long-run proportion of time that the Markov chain is in state j.
- Since μ_{jj} is the expected number of transitions needed to return to the state j, it follows that $\pi_j = 1/\mu_{jj}$.

C.3 Markov Chains with Finite State Space

Recall that in an irreducible Markov chain with finite state space the states are either all positive recurrent, or all null recurrent. We clearly have that
$$\sum_{j\in S} P_{ij}^{(n)} = 1.$$
Taking limit for $n \to \infty$ on both sides yields
$$1 = \lim_{n\to\infty} \sum_{j\in S} P_{ij}^{(n)} = \sum_{j\in S} \lim_{n\to\infty} P_{ij}^{(n)}$$
if $|S| < \infty$, i.e., the number of states is finite. Therefore, an irreducible and aperiodic Markov chain with a finite state space is *always* ergodic.

C.4 Examples

Example 1: **(s, S) Inventory System with Periodic Review** (cont.)

Recall Markov model I for this example. Is this Markov process aperiodic? If we assume that $\varphi_0 > 0$, then $P_{ii} > 0$ for all i, which implies that the Markov chain is aperiodic. It can be checked that this Markov chain is irreducible as long as $\varphi_1 > 0$. Finally, the state space of the Markov chain is *finite*. Therefore, the chain is ergodic. Hence the limiting distribution is the unique solution of the system:
$$\pi_j = \sum_{i=s+1}^{S} P_{ij}\pi_i, \quad j = s+1, \ldots, S; \text{ and } \sum_{j=s+1}^{S} \pi_j = 1$$

Recall that the transition probabilities are:

$$P_{ij} = \begin{cases} \varphi_{i-j} & j = s+1,\ldots \min(i, S-1);\ i = s+1,\ldots,S \\ \sum_{k=i-s}^{\infty} \varphi_k & j = S;\ i \neq S \\ \varphi_0 + \sum_{k=S-s}^{\infty} \varphi_k & j = S;\ i = S. \end{cases}$$

Thus, the limiting distribution is the solution to

$$\begin{cases} \pi_S = \varphi_0 \pi_S + \sum_{i=s+1}^{S} \left(\sum_{k=i-s}^{\infty} \varphi_k \right) \pi_i \\ \pi_j = \sum_{i=j}^{S} \varphi_{i-j} \pi_i \quad \text{for } j = s+1,\ldots,S-1 \\ \sum_{j=s+1}^{S} \pi_j = 1 \end{cases}$$

This system can be solved by
- first setting π_S equal to some arbitrary positive value, say 1;
- then solve recursively for the other values π_j, $j = s+1, \ldots, S-1$;
- finally, renormalize the values so that the sum is equal to one.

Note that one of the equations is *redundant*. In this case, it is most convenient to use all but the first equation.

<u>Example 2 (Example 4.3(A))</u>: **M/G/1 Queue** – Limiting Probabilities of the Embedded Markov Chain

The embedded Markov chain for this system is irreducible and aperiodic. Recall that the transition probabilities are:

$$P_{ij} = \begin{cases} \varphi_j & \text{if } i = 0;\ j = 0,1,\ldots \\ 0 & \text{if } i = 1,2,\ldots;\ j = 0,1,\ldots,i-2 \\ \varphi_{j-i+1} & \text{if } i = 1,2,\ldots;\ j = i-1, i,\ldots \end{cases}$$

where φ_j is the probability of j arrivals during a service period, given by

$$\varphi_j = \int_0^{\infty} e^{-\lambda x} \frac{(\lambda x)^j}{j!} dG(x)$$

Define $\rho = \sum_j j\varphi_j$, which is the expected number of arrivals during a service period. It is easy to check that $\rho = \lambda\mu_G$, where μ_G is the expected value of the service with the distribution G.

A stationary distribution if the Markov chain satisfies

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij} = \pi_0 \varphi_j + \sum_{i=1}^{j+1} \pi_i \varphi_{j-i+1} \quad j = 0,1,\ldots \Rightarrow \pi_{j+1} = \frac{\pi_j - \pi_0 \varphi_j - \sum_{i=1}^{j} \pi_i \varphi_{j-i+1}}{\varphi_0} \quad j = 0,1,\ldots$$

Note that $\varphi_0 > 0$.

This means that we can try to solve this system by

- fixing an arbitrary positive value for π_0;
- solve recursively for the remaining values of π_j;
- renormalize so that the values add to one.

However,

- does the solution exist?
- is it unique?
- is it the limiting distribution of the chain?

It can be shown that a solution exists if and only if $\rho < 1$, equivalently, $\mu_G < 1/\lambda$, i.e., the expected service time is smaller than the expected interarrival time. This means that this Markov chain is positive recurrent if and only if $\rho < 1$.

The textbook (p.178 – 179) proves the above result by showing that $\pi_0 = 1 - \rho$, by establishing the relationship between two moment generating functions:

$$\pi(s) = \sum_{j=0}^{\infty} \pi_j s^j, \quad \varphi(s) = \sum_{j=0}^{\infty} \varphi_j s^j.$$

That is, we have

$$\pi(s) = \frac{(s-1)\pi_0 \varphi(s)}{s - \varphi(s)}$$

Since $\varphi(1) = 1$, then by using L'hospital's rule, we obtain:

$$\lim_{s \to 1} \pi(s) = \frac{\pi_0}{1 - \varphi'(1)} = \frac{\pi_0}{1-\rho}$$

which should be equal to one. Therefore, the stationary distribution exists if and only if $\pi_0 = 1 - \rho > 0$, i.e., $\rho < 1$.

Since the Markov chain is aperiodic and irreducible, it follows that the stationary distribution (if it exists) is unique, and it also the limiting distribution of the chain. In that case, all states are positive recurrent.

<u>Example 3 (Example 4.3(B))</u>: **G/M/1 Queue** – Limiting Probabilities of the Embedded Markov Chain

Note that the embedded Markov chain has the following transition probabilities:

$$P_{i,k} = \int_0^{\infty} e^{-\mu t} \frac{(\mu t)^{i+1-k}}{(i+1-k)!} dG(t) \text{, for } k = 1, \ldots, i+1.$$

$$P_{i0} = \int_0^{\infty} \sum_{k=i+1}^{\infty} e^{-\mu t} \frac{(\mu t)^k}{k!} dG(t), \ i \geq 0.$$

Then we have

$$\pi_k = \sum_{i=1}^{\infty} \pi_i P_{ik} = \sum_{i=k-1}^{\infty} \pi_i \left(\int_0^{\infty} e^{-\mu t} \frac{(\mu t)^{i+1-k}}{(i+1-k)!} dG(t) \right) \text{ for all } k \geq 1.$$

$$\sum_{k=0}^{\infty} \pi_k = 1$$

(We can ignore the case k = 0 since one of the equation is redundant.)

Because of special structure of a Poisson distribution, let us try $\pi_k = c\beta^k$ for $k \geq 0$. Under this specification, it follows,

$$c\beta^k = c\sum_{i=k-1}^{\infty} \beta^i \int_0^{\infty} e^{-\mu t} \frac{(\mu t)^{i+1-k}}{(i+1-k)!} dG(t) \; = c\int_0^{\infty} e^{-\mu t} \beta^{k-1} \sum_{i=k-1}^{\infty} \frac{(\beta\mu t)^{i+1-k}}{(i+1-k)!} dG(t)$$

$$= c\beta^{k-1}\int_0^{\infty} e^{-\mu(1-\beta)t} dG(t)$$

which implies that β must satisfies the following equation:

$$\beta = \int_0^{\infty} e^{-\mu(1-\beta)t} dG(t) \qquad (*)$$

There is no guarantee that such an β indeed exists. A sufficient condition is given below:

Exercise: Show that if $\mu_G > 1/\mu$, then (*) has a unique solution that is between 0 and 1.

Once the existence of β is established, it is easy to check that

$$\pi_k = (1 - \beta)\beta^k, \text{ for } k = 0, 1, \ldots$$

where β is the solution to (*).

Example 4: Consider a Markov chain with state space S and transition probability matrix P. In addition, suppose that the matrix P is *doubly stochastic*:

- $\sum_{i \in S} P_{ij} = 1$ for all $j \in S$; (columns)
- $\sum_{j \in S} P_{ij} = 1$ for all $i \in S$. (rows)

Any stationary distribution of this Markov chain should satisfy

$$\pi_j = \sum_{i \in S} \pi_i P_{ij} \quad \text{for all } j \in S$$

Note that $\pi_j = c$ for $j \in S$ is a solution to the above system!

Now note that, if the Markov chain is irreducible and aperiodic, the stationary distribution should be unique *if it exists*.

If $|S| < \infty$, the chain is ergodic, and we have

$$\pi_j = \frac{1}{|S|} \quad \text{for al } j \in S,$$

which is the limiting distribution of the Markov chain.

If $|S| = \infty$, the limiting distribution does not exist, and all states are either transient or null recurrent.

In conclusion, an irreducible and aperiodic Markov chain with finite state space and doubly-stochastic transition probability matrix has the uniform distribution over the state space as its limiting distribution.

Stationary probabilities can sometimes be determined not by algebraically solving the stationary equations, but by reasoning directly that when the initial state is chosen according to a certain set of probabilities, then the resulting chain is stationary.

Example 5 (Example 4.3(C)): **Age of a Renewal Process** (p.180)

At the beginning, an item is put in use; and when it fails it will be replaced at the beginning of the next time period by a new item. Assume that the lives of the items are independent and each will fail in its j-th period of use with probability P_j, where the distribution $\{P_j\}$ is aperiodic and $\sum_j jP_j < \infty$. Let X_n be the age of the item in use at time n – that is, the number of periods it has been in use. Then $\{X_n, n \geq 0\}$ is a Markov chain with transition probabilities:

$$P_{i,1} = 1 - P_{i,i+1} = \frac{P_i}{\sum_{j=1}^{\infty} P_j} \equiv \lambda(i), \quad i \geq 1.$$

Then the limiting probabilities $\{\pi_i, i \geq 1\}$are given by

$$\pi_1 = \frac{1}{E(X)};$$

$$\pi_i = \frac{P(X \geq i)}{E(X)} \quad i \geq 1.$$

The following two examples show how the stationary distributions can be derived by reasoning directly when the initial state is chosen according to a certain set of probabilities such that the resulting chain is stationary, instead of solving the stationary equations.

Example 6 (Example 4.3(D)): **Population Growth Model** (p.182)

During each time period, every member of a population independently dies with a probability p, and also that the number of new members that join the population in each period is a Poisson random variable with mean λ. Let X_n be the number of members in the population at the beginning of period n. Then $\{X_n, n \geq 1\}$ is a Markov chain. The stationary distribution is Poisson with mean λ/p:

$$\pi_j = e^{-\lambda/p} \frac{(\lambda/p)^j}{j!}, \; j = 0,1,...$$

The key is to assume that we start with X_0 to be a Poisson distributed with mean α. Then X_1 follows a Poison with mean $\alpha(1-p) + \lambda$. Then stationarity of the chain requires that

$$\alpha(1-p) + \lambda = \alpha \Leftrightarrow \alpha = \lambda/p.$$

Example 7 (Example 4.3(E)): **The Gibbs Sample** (p.182 – 183)

D. Transitions among Classes

D.1 Decomposition of General Markov Chains & Transition among Classes

It is easy to show that a recurrent class is ***a closed class*** in the sense that once entered it is never left. (Proposition 4.4.1)

Any Markov chain can be decomposed into (a) a number of *closed classes*; (b) the union of all other classes. The transition probability matrix can then (possibly by reordering the states) be written as:

$$P = \begin{pmatrix} P_0 & A_1 & A_2 & \cdots \\ 0 & P_1 & 0 & \cdots \\ 0 & 0 & P_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Then it is clear that:

- The limiting probability of being in one of the *transient states* (corresponding to P_0 and A_1, A_2, …) is 0.
- Each of the remaining classes of states can itself be viewed as an irreducible Markov chain (i.e., a *subchain*). So we can study the limiting behavior of these chains.

If a subchain is ergodic, it has a limiting distribution. Note that each of these limiting distributions is a stationary distribution of the original Markov chain!

Now let us study the probabilities of entering each of the closed sets of recurrent states from any of the transient states. Denote the set of transient states by T and the r^{th} set of recurrent states by R. Consider $i \in T$ and $j \in R$. We are interested in characterizing f_{ij}, the probability of ever entering a recurrent state j given that the process starts in a transient state i.

Note that we can go from i to j by

- Going from i to some other state $k \in T$ in one step, and then from k eventually to j; or
- Going from i to some state $k \in R$ in one step, and then from k eventually to j.

Proposition 4.4.2: If j is recurrent, then the set of probabilities $\{f_{ij},\, i \in T\}$ satisfies

$$f_{ij} = \sum_{k\in T} P_{ik} f_{kj} + \sum_{k \in R} P_{ik} \text{ , } i \in T$$

where R is the set of states communicating with j.

Proof: Note that

$$\begin{aligned} f_{ij} &= P(N_j(\infty) > 0 \mid X_0 = i) \\ &= \sum_{\text{all } k} P(N_j(\infty) > 0 \mid X_1 = k, X_0 = i) P(X_1 = k \mid X_0 = i) \\ &= \sum_{k \in T} f_{kj} P_{ik} + \sum_{k\in R} f_{kj} P_{ik} + \sum_{\substack{k \notin T \\ k \notin R}} f_{kj} P_{ik} = \sum_{k\in T} f_{kj} P_{ik} + \sum_{k \in R} P_{ik} \end{aligned}$$

because of facts: $f_{kj} = 1$ for $k \in R$ and $f_{kj} = 0$ for $k \notin T$, $k \notin R$.

D.2 Gambler's Ruin Problem (Example 4.4(A))

Consider the following problem:

- A gambler has a probability p of winning 1 unit and a probability $1 - p$ of losing 1 unit at each play of the game.
- Assume successive games are independent.

Find the probability that starting with i units, the gambler's fortune will reach N before reaching 0.

Let X_n be the player's fortune at time n. The $\{X_n, n = 0, 1, \dots\}$ is a Markov chain with transition probabilities:

$$P_{00} = P_{NN} = 1;$$
$$P_{i,i+1} = p = 1 - P_{i,\,i-1},\ \ i = 1, 2, \dots, N\text{-}1.$$

This Markov chain has three classes:

- Two recurrent classes: {0}, {N};
- One transient class: $\{1, 2, \dots, N\text{-}1\}$.

Since each transient state is only visited finitely often, the gambler will, after a finite amount of time, will either attain her goal of N or go broke.

Let $f_i \equiv f_{iN}$ be the probability that starting with i, the gambler's fortune will eventually reach N. Then it is easy to see that,

$$f_i = p\, f_{i+1} + (1\text{-}p)\, f_{i\text{-}1},\quad i = 1, 2, \dots, N\text{-}1.$$

Let $q = 1 - p$. We then have the following recursive relations:

$$f_i - f_{i-1} = \frac{q}{p}(f_{i-1} - f_{i-2}) = \dots = \left(\frac{q}{p}\right)^{i-1} f_1,\ \text{ for all } i \ge 1.$$

Therefore, we have

$$f_i - f_1 = f_1\left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \dots + \left(\frac{q}{p}\right)^{i-1}\right]$$

or alternatively,

$$f_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)}\, f_1 & \text{if } \dfrac{q}{p} \ne 1 \\[2ex] i \cdot f_1 & \text{if } \dfrac{q}{p} = 1 \end{cases}$$

Since $f_N = 1$, it follows that

$$f_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)^N} & \text{if } p \ne \tfrac{1}{2} \\[2ex] \dfrac{i}{N} & \text{if } p = \tfrac{1}{2} \end{cases}$$

Note that

$$\lim_{N\to\infty} f_i = \begin{cases} 1-(q/p)^i & \text{if } p > \tfrac{1}{2} \\ 0 & \text{if } p \le \tfrac{1}{2} \end{cases}$$

which implies that if $p > ½$, then there is a positive probability that the gambler's fortune will converge to infinity; if $p \le ½$, then, with a probability 1, the gambler will eventually go broke when playing an infinitely rich opponent.

Suppose now that we want to determine the expected number of bets that the gambler, starting at i, makes before reaching either 0 or n.

Let X_j be the winnings on the jth bet, $j \ge 1$ and B be the number of bets until the gambler's fortune reaches 0 or N. Then,

$$B = \min\left\{ m : \sum_{j=1}^{m} X_j = -i \ \text{ or } \ \sum_{j=1}^{m} X_j = N - i \right\}$$

Since X_j's are independent with $E(X_j) = 2p - 1$, and B is a stopping time for X_j, it follows

$$E\left[\sum_{j=1}^{B} X_j \right] = (2p-1)E(B)$$

Now note that, for $p \ne ½$,

$$\sum_{j=1}^{B} X_j = \begin{cases} N - i & \text{with probability } f_i \\ -i & \text{with probability } 1 - f_i \end{cases}$$

which implies that

$$(2p - 1)E(B) = Nf_i - i.$$

Hence

$$E(B) = \frac{N}{2p-1}\left(\frac{1-(q/p)^i}{1-(q/p)^N} - \frac{i}{N} \right)$$

D.3 Mean Times in Transient States

For a finite state Markov chain, let $T = \{1, 2, \dots, t\}$ be the set of transient states with the following transition probability sub-matrix:

$$\begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1t} \\ P_{21} & P_{22} & \cdots & P_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ P_{t1} & P_{t2} & \cdots & P_{tt} \end{bmatrix}$$

(Note: this is not a transition probability matrix since the sum of its row elements may be less than 1.)

For $i, j \in T$, let m_{ij} be the expected total number of time periods spent in state j given that the chain starts in state i. Then it follows that (conditioning on the initial transition):

$$m_{ij} = \delta_{ij} + \sum_{k=1}^{t} P_{ik} m_{kj} \Rightarrow \boldsymbol{M} = \boldsymbol{I} + \boldsymbol{Q}\,\boldsymbol{M}$$

where $\boldsymbol{M} = [m_{ij}]$ and $\boldsymbol{I}$ is the identity matrix. Therefore,

$$\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}.$$

E. Branching Process

E.1 Definition and Examples

Suppose an organism at the end of its lifetime produces a random number ξ of offspring with the probability distribution:

$$P(\xi = k) = a_k, \quad k = 0,1,2,...;$$

$$a_k \geq 0, \quad \sum_{k=0}^{\infty} a_k = 1.$$

Assume that all offspring act independently of each other and at the end of their lifetime (for simplicity, the lifespan of all organisms are assumed to be the same) individually have progeny in accordance with the above distribution. Let X_n be the population size at nth generation. Then, for given initial state X_0, the random sequence $\{X_n: n \geq 0\}$ is a Markov chain, which also called a ***branching process***.

The Markovian property follows immediately from the following relationship between X_n and X_{n-1}:

$$P_{ij} = P(X_{n+1} = j \mid X_n = i) = P(\xi_1 + \cdots + \xi_i = j)$$

where ξ_i represents the number of offspring of i-individual of the n-th generation.

<u>Example 1 (Electron Multipliers)</u>: An electron multiplier is a device that amplifies a weak current of electrons. A series of plates are set up in the path of electrons emitted by a source. Each electron, as it strikes the first plate, generates a random number of new electrons, which in turn strike the next plate and produce more electrons, etc. Let X_0 be the number of electrons initially emitted. Define X_n as the number of electrons from the nth plate due to electrons emanating from (n-1)st plate.

<u>Example 2 (Neutron Chain Reaction)</u>: A nucleus is split by a chance collision with a neutron. The resulting fission yields a random number of new neutrons. Each of these secondary neutrons may hit some other nucleus producing a random number of additional neutrons, etc. In this case, the initial number of neutrons is $X_0 = 1$. The first generation of neutrons comprises all those produced from fission caused by the initial neutron. The size of the first generation is a random variable X_1. In general the population X_n at the $n$the generation is produced by the chance hits of the X_{n-1} individual neutrons of the (n – 1)st generation.

<u>Example 3 (Survival of Family Names)</u>: The family name is inherited by sons only. Suppose that each individual has probability p_k of having k male offspring. Then from one individual there result in 1st, 2nd, …., n-th, … generations of descendants. We may investigate the distribution of such random variables as the number of descendants in the nth generation, or the probability that the family name will eventually become extinct. Such questions are related to the analysis of branching processes.

<u>Example 4 (Survival of Mutant Genes)</u>: Each individual gene has a chance to give birth to k offspring, k = 1, 2, …, which are genes of the same kind. However, any individual has a chance to transform into a different type or mutant gene. This gene may become the first in a sequence of generations of a particular mutant gene. We may inquire about the chances of the survival of the mutant gene within the population of the original process.

E.2 Moment Generating Functions

Let $X_0 = 1$. Then we have the following relation:

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_i$$

Let us introduce the moment generating function

$$\varphi(t) = \sum_{k=0}^{\infty} a_k t^k$$

and

$$\varphi_n(t) = \sum_{k=0}^{\infty} P(X_n = k) t^k, \quad n = 0, 1, 2, \ldots$$

Clearly, we have $\varphi_0(t) = t$ and $\varphi_1(t) = \varphi(t)$. In general,

$$\begin{aligned}
\varphi_{n+1}(t) &= \sum_{k=0}^{\infty} P(X_{n+1} = k) t^k \\
&= \sum_{k=0}^{\infty} \left[\sum_{j=0}^{\infty} P(X_{n+1} = k \mid X_n = j) P(X_n = j) \right] t^k \\
&= \sum_{k=0}^{\infty} \left[\sum_{j=0}^{\infty} P(X_n = j) P(\xi_1 + \cdots + \xi_j = k) \right] t^k \\
&= \sum_{j=0}^{\infty} P(X_n = j) \left(\sum_{k=0}^{\infty} P_{jk} t^k \right) = \sum_{j=0}^{\infty} P(X_n = j) \left(\varphi(t) \right)^j
\end{aligned}$$

That is,

$$\varphi_{n+1}(t) = \varphi_n(\varphi(t)).$$

By induction, we can show that for $k = 0, 1, \ldots, n$,

$$\varphi_{n+1}(t) = \varphi_{n-k}(\varphi_{k+1}(t))$$

In particular,

$$\varphi_{n+1}(t) = \varphi(\varphi_n(t)).$$

From the above discussion, the following fact becomes evident:

Fact: Assume $X_0 = 1$ and let $\varphi(t)$ be the moment generating function of ξ. Then the transition probability P_{ij} is the j-the coefficient in the power series expansion of $[\varphi(t)]^i$.

Proposition: Assume $X_0 = 1$. Let $\mu = E(\xi) = E(X_1) < \infty$ and $\sigma^2 = Var(\xi) = Var(X_1) < \infty$. Then, for all $n \geq 1$,

(i) $E(X_n) = \mu^n$;

(ii) $$Var(X_n) = \begin{cases} \sigma^2 \mu^{n-1} \dfrac{\mu^n - 1}{\mu - 1} & \text{if } m \neq 1 \\ n\sigma^2 & \text{if } m = 1 \end{cases}$$

Proof: Using induction and the recursive relationship among the m.g.f.

E.3 Extinction Probabilities

We want to determine the probability that the population will eventually die out, i.e.,

$$P(X_n = 0 \text{ for some } n).$$

Without loss of generality, we assume $0 < a_0 < 1$. Let

$$q_n = P(X_n = 0) = \varphi_n(0).$$

From the above discussion, we have the following recursive relation:

$$q_{n+1} = \varphi(q_n).$$

By induction, we show that the sequence $\{q_n\}$ is a monotone increasing sequence bounded by 1. Therefore

$$\pi_0 = \lim_{n\to\infty} q_n$$

exists and $0 < \pi_0 \le 1$. Consequently,

$$\pi_0 = \varphi(\pi_0).$$

Note that π_0 is in fact the probability of eventual extinction.

Theorem 4.5.1:

(i) π_0 is the *smallest* positive root of $\pi_0 = \varphi(\pi_0)$.

(ii) $\pi_0 = 1$ if and only if $\mu \le 1$.

Proof: Refer to textbook (p. 192-193)

F. Time-Reversible Markov Chains

F.1 Definition and Examples

Consider an ergodic Markov chain $\{X_n, n \in \mathbf{N}\}$. Let the initial state be chosen according to the (unique) stationary distribution, in other words, we are in steady-state. Now imagine we run the Markov chain *backwards* in time. Is the reversed stochastic process a Markov chain?

We need to show that

$$\Pr(X_m = j \mid X_{m+1} = i, X_n = i_n, \text{ for } n \ge m+2\}$$
$$= \Pr(X_m = j \mid X_{m+1} = i)$$

which is true because of the fact that, given X_{m+1}, X_n is independent of X_{m+2}, X_{m+3}, …. Hence, the reversed chain is a Markov chain with transition probabilities:

$$P_{ij}^* = \Pr(X_m = j \mid X_{m+1} = i) = \frac{\Pr(X_m = j, X_{m+1} = i)}{\Pr(X_{m+1} = i)}$$
$$= \frac{P(X_{m+1} = i \mid X_m = j)P(X_m = j)}{\Pr(X_{m+1} = i)} = \frac{\pi_j P_{ji}}{\pi_i}$$

Definition (***Time Reversibility***) We say Markov chain $\{X_n, n \in \mathbf{N}\}$ is *time reversible* (*with respect to* π) if $P_{ij}^* = P_{ij}$ for all $i, j \in S$, which is equivalent to requiring that

$$\pi_i P_{ij} = \pi_j P_{ji}, \text{ for all } i, j \in S,$$

i.e., $P_{ij}\ \pi_i$ is *symmetric* function of i and j.

When in steady-state, the Markov chain looks the same looking forward and backward. Sometimes, we say that the *rate* at which the process goes from i to j (namely, $\pi_i P_{ij}$) is equal to the rate at which the process goes from j to i (namely, $\pi_j P_{ji}$).

In reverse, if we have a positive sequence, summing to 1, such that

$$x_i P_{ij} = x_j P_{ji} \text{ for all } i, j \in S; \sum x_i = 1.$$

Then it is clear that:

$$x_i = \sum_{j \in S} x_j P_{ji}\ , \text{ for all } i \in S.$$

Since the stationary distribution of the corresponding Markov chain π is unique, it follows that $\pi_i = x_i$.

Time reversibility is a very useful concept that can be used to *design* Markov chains that have a desired stationary or limiting distribution, sometimes called a *target distribution*.

Example: Suppose we have a finite set S. An ergodic Markov chain with state space S has the uniform distribution on S as its limiting distribution if the transition probability matrix is *symmetric*.

Example 4.7(A): (***An Ergodic Random Walk***) Consider a random walk with states 0, 1, …, M and transition probabilities:

$$P_{i,i+1} = \alpha_i = 1 - P_{i,i-1},\ i = 1, \ldots, M\text{-}1;$$
$$P_{0,1} = \alpha_0 = 1 - P_{0,0},$$
$$P_{M,M} = \alpha_M = 1 - P_{M,M-1}.$$

We can argue, without computation, that the above ergodic chain is time reversible.

- The number of transitions from i to i+1 must at all times be within 1 of the number from i+1 to i.
- Between any two transitions from i to i+1 there must be one from i+1 to i (and conversely) since the only way to re-enter state i from a higher state s by way of state i+1.
- Hence, the rate of transitions from i to i+1 equals the rate from i+1 to i. So the process is reversible.

Let us now find the limiting probabilities by equating for each state i = 0, 1, …, M-1 the rate at which the process goes from i to i+1 with the rate at which it goes from i+1 to i.

$$\pi_0\ \alpha_0 = \pi_1\ (1 - \alpha_1),$$
$$\pi_1\ \alpha_1 = \pi_2\ (1 - \alpha_2),$$
$$\vdots$$
$$\pi_i\ \alpha_i = \pi_{i+1}\ (1 - \alpha_{i+1}),\ i = 0, 1, \ldots, M\text{-}1$$

Solving in terms of π_0 yields

$$\pi_1 = \frac{\alpha_0}{1-\alpha_1}\pi_0,$$

$$\pi_2 = \frac{\alpha_1}{1-\alpha_2}\pi_1 = \frac{\alpha_1\alpha_0}{(1-\alpha_2)(1-\alpha_1)}\pi_0$$
$$\vdots$$

$$\pi_i = \frac{\alpha_{i-1}\cdots\alpha_0}{(1-\alpha_i)\cdots(1-\alpha_1)}\pi_0, \quad i = 1, 2, ..., M$$

Since $\sum_0^M \pi_i = 1$, it follows that

$$\pi_0 = \left[1 + \sum_{j=1}^{M} \frac{\alpha_{j-1}\cdots\alpha_0}{(1-\alpha_j)\cdots(1-\alpha_1)}\right]^{-1}$$

If $\alpha_i = p$ for all i, we get

$$\pi_0 = \frac{1-\beta}{1-\beta^{M+1}},$$

$$\pi_i = \frac{(1-\beta)\beta^i}{1-\beta^{M+1}}, \quad i = 0, 1, ..., M$$

where $\beta = p/(1-p)$.

Exercise: Suppose that M molecules are distributed among two urns; and at each time point one of the molecules is chosen randomly, removed from its urn, and placed in the other one. A well-known physicist Ehrenfest proposes the following random walk model to describe the movements of molecules:

$$\alpha_i = \frac{M-i}{M}, \quad i = 0, 1, ..., M$$

Find the stationary probabilities.

Example (**Random Walk on an Edge Weighted Graph**): Consider a graph having a positive number w_{ij} associated with each edge (i, j), and suppose that a particle moves from vertex to vertex in the following manner: If the particle is presently at vertex i then it will next move to vertex j with probability

$$P_{ij} = \frac{w_{ij}}{\sum_k w_{ik}}$$

$w_{ij} = 0$ if (i, j) is not on the graph. This Markov chain is called as a random walk on an edge weighted graph.

Proposition 4.7.1: If the Markov chain is irreducible, then it is time reversible with stationary probabilities given by

$$\pi_i = \frac{\sum_j w_{ij}}{\sum_i \sum_j w_{ij}}$$

Proof. The time reversibility equations

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

reduce to

$$\pi_i \frac{w_{ij}}{\sum_j w_{ij}} = \pi_j \frac{w_{ji}}{\sum_i w_{ji}}$$

or equivalently, since $w_{ij} = w_{ji}$,

$$\frac{\pi_i}{\sum_j w_{ij}} = \frac{\pi_j}{\sum_i w_{ji}}$$

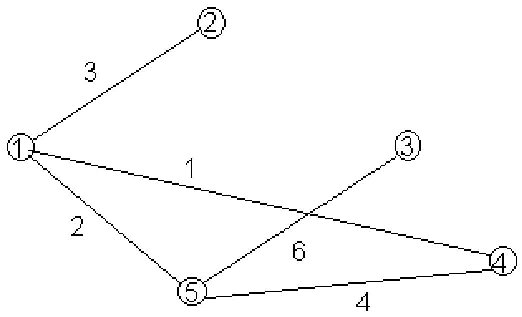
which is equivalent to

$$\frac{\pi_i}{\sum_j w_{ij}} = c \Rightarrow \pi_i = c\sum_j w_{ij}$$

Since $\sum_i \pi_i = 1$, it follows that

$$\pi_i = \frac{\sum_j w_{ij}}{\sum_i \sum_j w_{ij}}.$$

Consider the following graph:



Then it is easy to see that $P_{12} = 3/(3+1+2) = 0.5$. Furthermore, the limiting probabilities for this graph are given by:

$$\pi_1 = \tfrac{6}{32}, \pi_2 = \tfrac{3}{32}, \pi_3 = \tfrac{6}{32}, \pi_4 = \tfrac{5}{32}, \pi_5 = \tfrac{12}{32}$$

F.2 Characterization of Time Reversibility

If we try to solve the following system of equations:

$$\pi_i P_{ij} = \pi_j P_{ji}, \text{ for all } i, j \in S,$$

for an arbitrary Markov chain with states 0, 1, …, M, it usually has no solution. For example, under time reversibility, we should have

$$\pi_i P_{ij} = \pi_j P_{ji},$$
$$\pi_k P_{kj} = \pi_j P_{jk}$$

implying (if $P_{ij} P_{jk} > 0$) that

$$\frac{\pi_i}{\pi_k} = \frac{P_{ji}P_{kj}}{P_{ij}P_{jk}}$$

which in general need not equal to P_{ki}/P_{ik}. Thus, we see that a necessary condition for time reversibility is that

$$P_{ik} P_{kj} P_{ji} = P_{ij} P_{jk} P_{ki} \quad \text{for all } i, j, k$$

which equivalent to the statement that, starting in state i, the path $i \to k \to j \to i$ has the same probability as the reversed path $i \to j \to k \to i$. Note that time reversibility implies that the rate at which a sequence of transactions from i to k to j to i occurs must equal to rate of ones from i to j to k to i. This implies

$$\pi_i P_{ik} P_{kj} P_{ji} = \pi_i P_{ij} P_{jk} P_{ki}$$

hence the result follows when $\pi_i > 0$.

Theorem 4.7.2: An ergodic Markov chain for which $P_{ij} = 0$ whenever $P_{ji} = 0$ is time reversible if and only if starting in state i, any path back to i has the same probability as the reversed path, that is,

$$P_{i,i_1} P_{i_1,i_2} \cdots P_{i_k,i} = P_{i,i_k} P_{i_k,i_{k-1}} \cdots P_{i_1,i}$$

for all i, i_1, …, i_k.

Proof. Only need to prove the sufficiency. Fix i and j. Since

$$P_{i,i_1} P_{i_1,i_2} \cdots P_{i_k,j} P_{ji} = P_{ij} P_{j,i_k} P_{i_k,i_{k-1}} \cdots P_{i_1,i}$$

Summing the above expression for all possible states i_1, …, i_k, we have the following

$$P_{ij}^{(k+1)} P_{ji} = P_{ij} P_{ji}^{(k+1)}$$

Taking $k \to \infty$, we have $\pi_j P_{ji} = \pi_i P_{ij}$, which proves the theorem. □

Example 4.7(D) (**A List Problem**), p. 210

The **concept of the reversed chain** is useful even when the process is not time reversible.

Theorem 4.7.3: Consider an irreducible Markov chain with transition probabilities P_{ij}. If we can find positive numbers π_i, $i \geq 0$, summing to one, and a transition probability matrix $\mathbf{Q} = [Q_{ij}]$ such that

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

then the Q_{ij} are the transition probabilities of the reversed chain and π_i are the stationary probabilities both for the original and the reversed chain.

Proof. Refer to textbook (p.211).

Example 4.7(C): (Continuation of Example 4.3(C).) The transition probabilities for the reversed chain are:

$$Q_{i,i-1} = 1,\ i > 1$$
$$Q_{1,i} = P_i,\ \ i \geq 1$$

G. Semi-Markov Processes

A semi-Markov process is one that changes the states according to a Markov chain but takes a random amount of time between change.

Definition: A stochastic process with states 0, 1, 2…, which is such that, whenever it enters state i, $i \geq 0$:

(i) The next state it will enter is state j with probability P_{ij}, $i, j \geq 0$.

(ii) Given that the next state to be entered is state j, the time until the transition from i to j occurs has distribution F_{ij}.

If we let $Z(t)$ denote the state at time t, then $\{Z(t), t \geq 0\}$ is called ***a semi-Markov process***.

Remarks:

- A semi-Markov process does not possess the Markovian property.
- A normal Markov chain is a semi-Markov process with $F_{ij}(t) = \delta_{[t \geq 1]}$, namely, all transition times are identically 1.

Let H_i be the distribution of time that the semi-Markov process spends in state i before making a transition. Then it is easy to check

$$H_i(t) = \sum_j P_{ij} F_{ij}(t)$$

Let $\mu_i = \mathrm{E}[H_i]$.

If we let X_n denote the nth state visited, then $\{X_n, n \geq 0\}$ is a Markov chain with transition probabilities P_{ij}. We call this the *embedded* Markov chain of the semi-Markov process. We say the semi-Markov process is *irreducible* if the embedded Markov chain is irreducible.

Let T_{ii} be the time between successive transitions into state i and let $\mu_{ii} = \mathrm{E}[T_{ii}]$. We can now using the results on alternating renewal processes to derive limiting probabilities of a semi-Markov process.

Proposition 4.8.1: If the semi-Markov process is irreducible and if T_{ii} has a nonlattice distribution with finite mean, then,

$$P_i = \lim_{t \to \infty} \Pr(Z(t) = i \mid Z(0) = j)$$

exists and is independent of the initial state. Furthermore,

$$P_i = \mu_i \,/\, \mu_{ii}.$$

Proof. In line with a delayed alternating renewal process, we have the following:

- A cycle begins when the process enters state i;
  - The process is "on" when in state i and "off" when not in state i.
  - On time has the distribution H_i.
  - Cycle time is T_{ii}.
- There will be a delay when $Z(0) \neq i$.

Theorem 4.8.3: Suppose the semi-Markov process is irreducible and T_{ii} has a nonlattice distribution with finite mean. Further assume that the embedded Markov chain $\{X_n: n \geq 0\}$ is positive recurrent. Then

$$P_i = \frac{\pi_i \mu_i}{\sum_j \pi_j \mu_j}$$

where $\{\pi_j, j \geq 0\}$ is the stationary probability for the embedded Markov chain:

$$\pi_j = \sum_i \pi_i P_{ij}, \ \sum_j \pi_j = 1.$$

Proof. Refer to textbook, p. 215 – 216.