Chapter 1

Elements of Stochastic Processes

Definition 1.1 A family of random variables $X_t$, where $t$ is a parameter running over an index set $T$, is called a stochastic process. Write $\{X_t, t \in T\}$.

A realization or sample function of stochastic process $\{X_t, t \in T\}$ is an assignment to each $t \in T$, of a possible value of $X_t$. — function of $t$.

Classification of stochastic processes: The main elements distinguishing stochastic processes are in the nature of the state space, the index parameter $T$, and the dependence relations.

(1). State space.
This is a space in which the possible values of each $X_t$ lies. In the case that $S = \{0, 1, \ldots\}$, we refer to the process as integer valued or as a discrete state process. If $S = (-\infty, \infty)$, we call $X_t$ a real-valued stochastic process. If $S = \mathbb{R}^k$, then $X_t$ is a $k$-vector process.

(2). Index parameter $T$
If $T = \{0, 1, 2, \ldots\}$, then we say that $X_t$ is a discrete time process. In this case, we shall write $X_n$ instead of $X_t$.
If $T = [0, \infty)$, then $X_t$ is called a continuous time process.

(3). Dependent relations
Let $\{X_t, t \in T\}$ be a real-valued stochastic process with discrete or continuous parameter set.
(a). Independent increments
If r.v.’s $X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent for all values of $t_1, \ldots, t_n$ such that $t_1 < t_2 < \cdots < t_n$, then $X_t$ is a process with independent increments.

(b). Martingales
We say that $\{X_t, t \in T\}$ is a martingale if for any $t_1 < t_2 < \cdots < t_n < t_{n+1}$, we have

$$E(X_{t_{n+1}}|X_{t_1} = a_1, \ldots, X_{t_n} = a_n) = a_n$$

for all values $a_1, \ldots, a_n$. It may be considered as a model for fair games, in the sense that $X_t$ signifies the amount of money that a player has at time $t$. The martingale property states that the average amount a player will have at time $t_{n+1}$, given that he has amount $a_n$ at time $t_n$, is equal to $a_n$ regardless of what his past fortune has been. For example, if
$X_n = Z_1 + \cdots + Z_n$, $n = 1, 2, \cdots$, is a discrete time martingale if the $Z_i$ are independent and have means zero (prove it).

(c). Markov process

A Markov process is a process with the property that given the values of $X_t$, the values of $X_s$, $s > t$, don’t depend on the values of $X_u$, $u < t$. In formal terms, a process is said to be Markovian if

$$P(a < X_t \leq b|X_t = x_1, \cdots, X_{t_n} = x_n) = P(a < X_t \leq b|X_{t_n} = x_n),$$

(0.1)

where $t_1 < \cdots < t_n < t$.

Let $A$ be an interval of the real line. The function

$$P(x, s; t, A) = P(X_t \in A|X_s = x), \quad t > s$$

is called the transition probability function. We may write (0.1) as

$$P(a < X_t \leq b|X_{t_1} = x_1, \cdots, X_{t_n} = x_n) = P(x_n, t_n; t, A),$$

where $A = \{\xi : a < \xi \leq b\}$.

A Markov process with finite or denumerable state space is called a Markov chain. A Markov process for which all realizations or sample functions $\{X_t, t \in [0, \infty)\}$ are continuous functions is called a diffusion process. The Poisson process is a continuous time Markov chain and Brownian motion is a diffusion process.

(d). Stationary process

A stochastic process $\{X_t, t \in T\}$ [here $T$ could be one of the sets $(-\infty, \infty), [0, \infty)$, the set of all integers, or the set of all positive integers] is said to be strictly stationary if the joint distribution functions of the families of r.v.’s

$$(X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_n+h})$$

and

$$(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$$

are the same for all $h > 0$ and arbitrary selections $t_1, t_2, \cdots, t_n$ from $T$. It means that the process is in probabilistic equilibrium and that the particular times at which we examine the process are of no relevance.

A stochastic process $\{X_t, t \in T\}$ is said to be wide sense stationary or covariance stationary if it possesses finite second moments and if

$$\text{Cov}(X_t, X_{t+h}) = E(X_t X_{t+h}) - E(X_t)E(X_{t+h})$$

depends only on $h$ for all $t \in T$. A stationary process with finite second moments is covariance stationary. There are covariance stationary processes that are not stationary.

A Markov process is said to have stationary transition probabilities if $P(x, s; t, A)$ is a function only of $t - s$. Remember that $P(x, s; t, A)$ is a conditional probability, given the present state. So there is no reason to expect that a Markov process with stationary transition prob. is stationary.

Neither the Poisson process nor the BM process is stationary. But for Poisson or BM process $\{X_t, t \geq 0\}$,

$$Z_t = X_{t+h} - X_t$$

is a stationary process for any fixed $h \geq 0$. 
Proposition 1.0.1 Suppose $P(X \geq 0) = 1$ and $EX < \infty$. Then $EX = \int_0^\infty P(X > x)dx$. Moreover, if $\sum_{k=0}^\infty P(X = k) = 1$, then $EX = \sum_{k=0}^\infty P(X > k)$.

Proof. First note that 
$$xP(X > x) \leq EXI(X > x) \to 0$$
as $x \to \infty$. So
\[
EX = \int_0^\infty xdP(X \leq x) \\
= -\int_0^\infty xdP(X > x) \quad \text{(integration by parts)} \\
= \int_0^\infty P(X > x)dx.
\]

If $X$ is integer valued, then $P(X > s) = P(X > k)$ for $k \leq s < k + 1$. Hence
\[
EX = \sum_{k=0}^\infty \int_k^{k+1} P(X > s)ds = \sum_{k=0}^\infty P(X > k).
\]
Chapter 2
Markov Chains

2.1 Definitions

A discrete time Markov chain \( \{X_n\} \) is a Markov stochastic process whose state space is a countable or finite set, and for which \( T = (0, 1, 2, \cdots) \).

We usually label the state space of the process by the non-negative integers \((0, 1, 2, \cdots)\) and speak of \(X_n\) being in state \(i\) if \(X_n = i\).

The probability of \(X_{n+1}\) being in state \(j\) given that \(X_n\) is in state \(i\) is denoted by \(P_{ij}^{n,n+1}\), i.e.,

\[
P_{ij}^{n,n+1} = P\{X_{n+1} = j | X_n = i\}
\]

(1.1)

\(P_{ij}^{n,n+1}\) is called a one-step transition probability.

In general the transition prob. are functions not only of the initial and final state, but also of the time of transition as well. When one-step transition probabilities are independent of the time variable, then we say that the Markov process has stationary transition probabilities and write \(P_{ij}^{n,n+1} = P_{ij}\). We can arrange the probabilities as a matrix,

\[
P = \begin{pmatrix}
P_{00} & P_{01} & P_{02} & P_{03} & \cdots \\
P_{10} & P_{11} & P_{12} & P_{13} & \cdots \\
P_{20} & P_{21} & P_{22} & P_{23} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \\
P_{i0} & P_{i1} & P_{i2} & P_{i3} & \cdots \\
\vdots & \vdots & \vdots & \vdots &
\end{pmatrix}
\]

We refer to \(P = (P_{ij})\) as the Markov matrix or transition probability matrix of the process.

The \((i+1)\)st row of \(P\) is the probability distribution of the values of \(X_{n+1}\) given \(X_n = i\). If the number of states is finite, then \(P\) is a finite square matrix whose order (the number of rows) is equal to the number of states. Clearly, \(P_{ij}\) satisfy

\[
P_{ij} \geq 0, \quad i, j = 0, 1, 2, \ldots,
\]

\[
\sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, 2, \ldots.
\]

The 2nd condition expresses the fact that some transition occurs at each trial. (One says that a transition has occurred even if the state remains unchanged.)
Proposition 2.1.1 A discrete time Markov chain \( \{X_n\} \) with stationary transition probabilities is completely determined by \((1.1)\) and the value of \(X_0\)

Proof. Let \(P\{X_0 = i\} = p_i\). It is enough to show how to compute the quantities

\[
P\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\},
\]

since any probability involving \(X_{j_1}, \ldots, X_{j_k}, j_1 < \cdots < j_k\) may be obtained by summing terms of the form \((1.2)\)

By the definition of conditional probabilities we have

\[
P\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\} = P\{X_n = i_n | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\} P\{X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\}
\]

\[
P\{X_n = i_n | X_{n-1} = i_{n-1}\} P\{X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\} = \cdots = P_{i_{n-1}i_n} P_{i_{n-2}i_{n-1}} \cdots P_{i_0i_1} p_{i_0}
\]

2.2 Examples of Markov Chains

A. SPATIALLY HOMOGENEOUS MARKOV CHAINS

Let \(\xi\) denote a discrete-valued random variable whose possible values are the non-negative integers, \(P\{\xi = i\} = a_i \geq 0\), and \(\sum_{i=0}^{\infty} a_i = 1\). Let \(\xi_1, \ldots, \xi_n, \ldots\) represent independent copies of \(\xi\).

i. Consider the process \(X_n, n = 0, 1, \ldots\), defined by \(X_n = \xi_n\), \((X_0 = \xi_0\) prescribed).

\[
P_{i,j}^{n,n+1} = P\{X_{n+1} = j | X_n = i\} = P\{\xi_{n+1} = j | \xi_n = i\} = P\{\xi_{n+1} = j\} = a_j
\]

Its Markov matrix has the form

\[
P = \begin{pmatrix}
a_0 & a_1 & a_2 & \cdots \\a_0 & a_1 & a_2 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]

ii. Partial sum \(\eta_n = \xi_1 + \cdots + \xi_n, n = 1, 2, \ldots\), and by definition, \(\eta_0 = 0\).

\[
P_{i,j}^{n,n+1} = P(\eta_{n+1} = j | \eta_n = i)
\]

\[
= P(\xi_1 + \cdots + \xi_{n+1} = j | \xi_1 + \cdots + \xi_n = i)
\]

\[
= P(\xi_{n+1} = j - i | \xi_1 + \cdots + \xi_n = i)
\]

\[
= \begin{cases}
0 & j < i \\
a_{j-i} & j \geq i
\end{cases}
\]

\[
P = \begin{pmatrix}
a_0 & a_1 & a_2 & \cdots \\0 & a_0 & a_1 & \cdots \\
0 & 0 & a_0 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]
B. ONE-DIMENSIONAL RANDOM WALKS

A one-dimensional random walk is a Markov chain whose state space is a finite or infinite subset \( a, a+1, \ldots, b \) of the integers, in which the particle, if it is in state \( i \), can in a single transition either stay in \( i \) or move to one of the adjacent states \( i-1, i+1 \). If the state space is taken as the nonnegative integers, then the transition matrix of a random walk has the form

\[
P = \begin{pmatrix}
  r_0 & p_0 & 0 & 0 & \cdots \\
  q_1 & r_1 & p_1 & 0 & \cdots \\
  0 & q_2 & r_2 & p_2 & \cdots \\
  \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]

where \( p_i \geq 0, q_i \geq 0, r_i \geq 0 \) and \( p_i + q_i + r_i = 1, i = 1, 2, \ldots, p_0 \geq 0, r_0 \geq 0, p_0 + r_0 = 1 \). If \( X_n = i \), then for \( i \geq 1 \),

\[
P(X_{n+1} = i+1 | X_n = i) = p_i, \quad P(X_{n+1} = i-1 | X_n = i) = q_i, \quad P(X_{n+1} = i | X_n = i) = r_i
\]

with the obvious modifications holding for \( i = 0 \).

The designation “random walk” seems apt since a realization of the process describes the path of a person (suitable intoxicated) moving randomly one step forward or backward.

C. A DISCRETE QUEUEING MARKOV CHAIN

Customers arrive for service and take their place in a waiting line. During each period of time a single customer is served, provided that at least one customer is present. During a service period new customers may arrive. We suppose the actual number of arrivals in the \( n \)th service period is a r.v. \( \xi_n \) whose distribution function is independent of the period and is given by \( P(k \text{ customers arrive in a service period}) = P(\xi_n = k) = a_k \geq 0, k = 0, 1, \ldots \) and \( \sum_{i=0}^{\infty} a_i = 1 \). We also assume the r.v.’s \( \xi_n \) are independent. The state of the system at the start of each period is defined to be the number of customers waiting in line for service. If the present state is \( i \) then after a lapse of one period the state is

\[
j = \begin{cases} 
  i - 1 + \xi & \text{if } i \geq 1 \\
  \xi & \text{if } i = 0
\end{cases}
\]

where \( \xi \) is the number of new customers having arrived in this period while a single customer was served. We can write \( X_{n+1} = (X_n - 1)^+ + \xi_n \), where \( Y^+ = \max(Y, 0) \).

\[
P^{n,n+1}_{i,j} = P(X_{n+1} = j | X_n = i) = P((X_n - 1)^+ + \xi_n = j | X_n = i)
\]

\[
= P((i-1)^+ + \xi_n = j | X_n = i) = P(\xi_n = j - (i-1)^+)
\]

\[
P = \begin{pmatrix}
  a_0 & a_1 & a_2 & \cdots \\
  a_0 & a_1 & a_2 & \cdots \\
  0 & a_1 & a_2 & \cdots \\
  \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]

D. INVENTORY MODEL

Consider a situation in which a commodity is stocked in order to satisfy a continuing demand. We assume that the replenishing of stock takes place at successive times \( t_1, t_2, \ldots \),
and we assume that the cumulative demand for the commodity over the period \((t_{n-1}, t_n)\) is a random variable \(\xi_n\) whose distribution function is independent of the time period,

\[
P(\xi_n = k) = a_k, \quad k = 0, 1, 2, \ldots
\]

where \(a_k \geq 0\) and \(\sum_{k=0}^{\infty} a_k = 1\). The stock level is examined at the start of each period. An inventory policy is prescribed by specifying two nonnegative critical values \(s\) and \(S > s\).

The implementation of the inventory policy is as follows: If the available stock quantity is not greater than \(s\) then immediate procurement is done so as to bring the quantity of stock on hand to the level \(S\). If, however, the available stock is in excess of \(s\) then no replenishment of stock is undertaken. Let \(X_n\) denote the stock on hand just prior to restocking at \(t_n\). The states of the process \(\{X_n\}\) consist of the possible values of the stock size \(S, S - 1, \ldots, 1, 0, -1, -2, \ldots\), where a negative value is interpreted as an unfulfilled demand for stock, which will be satisfied immediately upon restocking. So the stock levels at two consecutive periods are connected by the relation

\[
X_{n+1} = \begin{cases} 
X_n - \xi_{n+1} & \text{if } s < X_n \leq S, \\
S - \xi_{n+1} & \text{if } X_n \leq s.
\end{cases}
\]

If we assume the \(\xi_n\) to be mutually independent, then the stock values \(X_0, X_1, \ldots\) constitute a Markov chain. Its transition probabilities are

\[
P(X_{n+1} = j | X_n = i) = \begin{cases} 
P(i - \xi_{n+1} = j | X_n = i) = P(\xi_{n+1} = j - i), & s < i \leq S \\
P(S - \xi_{n+1} = j | X_n = i) = P(\xi_{n+1} = S - i), & i \leq s \\
0, & i > S.
\end{cases}
\]

E. SUCCESS RUNS

Consider a Markov chain on the nonnegative integers with transition probability matrix of the form

\[
P = \begin{pmatrix} p_0 & q_0 & 0 & 0 & \ldots \\
p_1 & 0 & q_1 & 0 & \ldots \\
p_2 & 0 & 0 & q_2 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]

where \(q_i > 0, p_i > 0\) and \(q_i + p_i = 1, i = 0, 1, 2, \ldots\)

A special case of this transition matrix arises when one is dealing with success runs resulting from repeated trials each of which admits two possible outcomes, success (S) or failure (F). More explicitly, consider a sequence of trials with two possible outcomes (S) or (F). Moreover, suppose that in each trial, the probability of (S) is \(\alpha\) and the probability of (F) is \(\beta = 1 - \alpha\). We say a success run of length \(r\) happened at trial \(n\) if the outcomes in the preceding \(r + 1\) trials, including the present trial as the last, were respectively, F, S, S, ..., S. Let us now label the present state of the process by the length of the success run currently under way. In particular, if the last trial resulted in a failure then the state is zero. Similarly, when the preceding \(r + 1\) trials in order had the outcomes F, S, S, ..., S, the state variable would carry the label \(r\). The process is clearly Markovian (since the individual trials were independent of each other) and its transition matrix has the form (2.3) where \(p_n = \beta, n = 0, 1, \cdots\).

F. BRANCHING PROCESSES
Suppose an organism at the end of its lifetime produces a random number $\xi$ of offspring with probability distribution

$$P(\xi = k) = a_k \geq 0, \quad k = 0, 1, \ldots$$

(2.4)

where, $\sum_{k=1}^{\infty} a_k = 1$. We assume that all offspring act independently of each other and at the end of their lifetime individually have progeny in accordance with the probability distribution (2.4). The process $\{X_n\}$, where $X_n$ is the population size at the $n$th generation, is a Markov chain. The transition matrix is obviously given by

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

where the $\xi_i$’s are independent observations of a r.v. with probability law (2.4).

G. MARKOV CHAINS IN GENETICS

We assume that we are dealing with a fixed population size of $2N$ genes composed of type-a and type-A individuals. The make-up of the next generation is determined by $2N$ independent binomial trials as follows: If the parent population consists of $j$ a-genes and $2N - j$ A-genes then each trial results in a or A with probabilities $p_j = j/(2N)$, $q_j = 1 - p_j$ respectively. Repeated selections are done with replacement. By this procedure we generate a Markov chain $\{X_n\}$ where $X_n$ is the number of a-genes in the $n$th generation among a constant population size of $2N$ elements. The state space contains the $2N + 1$ values $\{0, 1, 2, \ldots, 2N\}$. The transition probability matrix is computed according to the binomial distribution as

$$P(X_{n+1} = k|X_n = j) = P_{jk} = \binom{2N}{k} p_j^k q_j^{2N-k} \quad (j, k = 0, 1, \cdots, 2N).$$

2.3 Transition Probability Matrices of a Markov Chain

We always assume the processes have stationary transition probability unless we specify.

**Theorem 2.3.1** If the one-step transition probability matrix of a Markov chain is $P = (P_{ij})$, then

$$P^n_{ij} = \sum_{k=0}^{\infty} P^r_{ik} P^s_{kj}$$

(3.5)

for any fixed pair of nonnegative integers $r$ and $s$ satisfying $r + s = n$, where we define

$$P^0_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

$$P^n_{ij} = P(X_{m+n} = j|X_m = i)$$

independent of $m$

**Proof.** We prove the case $n = 2$. The event of going from state $i$ to state $j$ in two transitions can be realized in the mutually exclusive ways of going to some intermediate state $k$ $(k = 0, 1, 2, \ldots)$ in the first transition and then going from state $k$ to state $j$ in the second transition.
Because of the Markovian assumption the probability of the second transition is \( P_{kj} \), and that of the first transition is \( P_{ik} \). If we use the law of total probabilities (3.5) follows.

Remark: It is easy to check that the \( n \)th step transition probability matrices

\[
P^n = \overbrace{P \times P \times \cdots \times P}^{n}
\]

### 2.4 Classification of States of a Markov Chain

State \( j \) is said to be accessible from state \( i \) if for some integer \( n \geq 0 \), \( P^n_{ij} > 0 \). Two states \( i \) and \( j \), each accessible to the other, are said to communicate and we write \( i \leftrightarrow j \). The concept of communication is an equivalence relation. i.e.,

i. Reflexivity. \( i \leftrightarrow i \)

ii. Symmetry. If \( i \leftrightarrow j \), then \( j \leftrightarrow i \).

iii. Transitivity. If \( i \leftrightarrow j \), \( j \leftrightarrow k \), then \( i \leftrightarrow k \).

The proof of transitivity proceeds as follows: \( i \leftrightarrow j \) and \( j \leftrightarrow k \) imply that there exist integers \( n \) and \( m \) such that \( P^n_{ij} > 0 \) and \( P^m_{jk} > 0 \). Consequently by (3.5) and the nonnegativity of each \( P^n_{rs} \) we conclude that

\[
P^n_{ik} = \sum_{r=0}^{\infty} P^n_{ir} P^m_{rk} \geq P^n_{ij} P^m_{jk} > 0.
\]

A similar argument shows the existence of an integer \( v \) such that \( P^v_{ki} > 0 \), as desired.

We can partition the states into equivalence classes.

We say that the Markov chain is irreducible if the equivalence relation induces only one class, that is all states communicate with each other.

Example:

\[
P = \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 0 & 1 & 0
\end{pmatrix} = \begin{pmatrix}
P_1 & 0 \\
0 & P_2
\end{pmatrix}
\]

This is the Markov chain divided into two classes \( \{0, 1\} \) and \( \{2, 3, 4\} \).

The period of state \( i \), written \( d(i) \), to be the greatest common divisor (g.c.d.) of all integers \( n \geq 1 \) for which \( P^n_{ii} > 0 \). If \( P^n_{ii} = 0 \) for all \( n \geq 1 \) define \( d(i) = 0 \).

Example:

\[
P = \begin{pmatrix}
0 & 1 & 0 & 0 & \ldots & 0 \\
0 & 0 & 1 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & \ldots & 0 \\
1 & 0 & 0 & 0 & \ldots & 0
\end{pmatrix}
\]

Each state has period \( n \).
\[ n = 3 \]
\[
P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \Rightarrow P_{ii} = 0, P^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow P^2_{ii} = 0, P^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow P^3_{ii} = 1.
\]

**Theorem 2.4.1**

i. If \( i \leftrightarrow j \) then \( d(i) = d(j) \).

ii. If state \( i \) has period \( d(i) \), then there exists an integer \( N \) depending on \( i \) such that for all integers \( n \geq N \)
\[
P^{nd(i)}_{ii} > 0
\]

iii. If \( P^m_{ji} > 0 \), then \( P^{m+nd(i)}_{ji} > 0 \) for all \( n \) sufficiently large.

**Lemma 2.4.1**

Let \( n_1, \ldots, n_k \) be positive integers with greatest common divisor \( d \). Then \( \exists \) a positive integer \( M \) such that \( m \geq M \) implies existing nonnegative integers \( \{c_{ji}\}_{j=1}^k \) s.t. \( md = \sum_{j=1}^k c_{ji}n_j \).

Proof of Theorem 2.4.1: (ii) If \( d(i) = 0 \), then obviously \( P^0_{ii} = 1 \). So we assume \( d(i) > 0 \). \( \exists n_0 \) s.t. \( P^{n_0}_{ii} > 0 \Rightarrow P^{nd(i)}_{ii} > 0 \) for any positive integer \( l \). Because \( P^{nd(i)}_{ii} = \sum_{k=0}^{n_0} \sum_{l=1}^{\infty} i_{kl} P^{n_0(l-1)}_{ii} \geq \sum_{k=0}^{n_0} P^{n_0} \sum_{l=1}^{\infty} P^{n_0(l-1)}_{ii} \geq (P^{n_0})^2 \). The definition of \( d(i) \) shows that there exist \( n_1, \ldots, n_k \) s.t.
\[
d(i) = \text{g.c.d.}(n_1, n_2, \ldots, n_k), \ P^{n_l}_{ii} > 0, \ l = 1, 2, \ldots, k.
\]

Then by Lemma 2.4.1, there exists an integer \( N \) depending on \( i \) such that for all integers \( n \geq N \),
\[
P^{nd(i)}_{ii} = P^\sum_{j=1}^k c_{ji}n_j \geq P^{n_1}_{ii} P^{n_2}_{ii} \cdots P^{n_k}_{ii} > 0.
\]

(iii) \( P^{m+nd(i)}_{ji} \geq P^{m}_{ji} P^{nd(i)}_{ii} \)

(i) From (ii), we know that \( \exists N \) s.t. \( \forall n \geq N, P^{nd(i)}_{ii} > 0 \)

\( i \leftrightarrow j \) : \( \exists n_0 \) s.t. \( P^{n_0}_{ij} > 0 \), \( \exists n_1 \) s.t. \( P^{n_1}_{ji} > 0 \). Then
\[
P^{m+n+nd(i)}_{ii} = \sum_{k=0}^{\infty} P^{n_0}_{ik} P^{m+nd(i)}_{kj} \geq P^{n_0}_{ij} P^{n_1}_{ji} P^{nd(i)}_{ii} > 0.
\]

This gives that \( d(j)|m_0 + m_1 + nd(i) \) and \( d(j)|m_0 + m_1 + (n+1)d(i) \). So \( d(j)|d(i) \). Similarly, \( d(i)|d(j) \). So \( d(j) = d(i) \).

(ii) Method 2. \( \exists l_1, l_2 \) s.t. \( P^{l_1d(i)}_{ij} > 0, P^{l_2d(i)}_{ij} > 0 \) and G.C.D. \( \{l_1, l_2\} = 1 \). Since \( i \leftrightarrow j \), \( \exists P^{l_i}_{ij} > 0 \), \( P^{m}_{ji} > 0 \). So \( P^{m+n+ld(i)}_{ji} > 0 \). Similarly, \( P^{m+n+ld(i)}_{jj} > 0 \), \( P^{m+n+ld(i)}_{jj} > 0 \). \( P^{m+n+2ld(i)}_{jj} > 0 \Rightarrow l_1d(i) = k_3d(j), l_2d(i) = k_4d(j) \Rightarrow d(i) = k_5d(j) \). (Note that if G.C.D. \( (a, b) = 1 \), then there exist \( \alpha \) and \( \beta \) s.t. \( \alpha a + \beta b = 1 \).) In a similar way, \( d(j) = k_6d(i) \).

(iii) Method 3. Since \( i \leftrightarrow j \), \( \exists m, n \) s.t. \( P^{m}_{ii} > 0, P^{n}_{ij} > 0 \). If \( P^s_{ii} > 0 \), then \( P^{s+m+n}_{ii} \geq P^{m}_{ii} P^{s}_{ij} P^{n}_{ji} > 0 \Rightarrow s + m + n = ld(j) \). It follows from \( P^{s}_{ii} > 0 \), then \( P^{2s}_{ii} \geq P^{s}_{ii} P^{s}_{ii} > 0 \). So \( P^{2s+m+n}_{ii} > 0 \Rightarrow 2s + m + n = ld(j) \). Thus \( s = (l_1 - l)d(j) \Rightarrow d(j)|d(i) \). Similarly, \( d(i)|d(j) \).

A Markov chain in which each state has period one is called aperiodic. The vast majority of Markov chain processes we deal with are aperiodic. Random walks usually typify the periodic cases arising in practice. Results will be developed for the aperiodic case and the modified conclusions for the general case will be stated usually without proof.
2.5 Recurrence

For any fixed \( i \), we define for each integer \( n \geq 1 \),
\[
f_n^{ij} = P\{X_n = j, X_v \neq 1, v = 1, 2, \ldots, n - 1 | X_0 = i\}
\]
That is the probability that the first passage from state \( i \), to state \( j \) at the \( n \)th transition.
We define \( f_0^{ij} = 0 \) for all \( i \) and \( j \). Note that \( P_0^{ij} = 1 \) if \( i = j \), \( = 0 \) if \( i \neq j \).

Clearly, \( f_1^{ij} = P^{ij} \) and \( f_n^{ij} \) may be calculated recursively by
\[
P_n^{ij} = \sum_{k=0}^{n} f_k^{ij} P_n^{jj}, \quad i \neq j, \quad n \geq 0,
\]
\[
P_n^{ii} = \sum_{k=0}^{n} f_k^{ii} P_n^{jj}. \quad n \geq 1.
\]

**Definition** The generating function \( P_{ij}(s) \) of the sequence \( \{P_n^{ij}\} \) is
\[
P_{ij}(s) = \sum_{n=0}^{\infty} P_n^{ij} s^n, \quad \text{for } |s| < 1.
\]
The generating function \( F_{ij}(s) \) of the sequence \( \{f_n^{ij}\} \) is
\[
F_{ij}(s) = \sum_{n=0}^{\infty} f_n^{ij} s^n, \quad \text{for } |s| < 1.
\]

Note that for \( |s| < 1 \),
\[
F_{ij}(s) P_{jj}(s) = (\sum_{n=0}^{\infty} f_n^{ij} s^n) (\sum_{n=0}^{\infty} P_n^{jj} s^n)
\]
\[
= \sum_{n=0}^{\infty} \left( \sum_{r=0}^{n} f_r^{ij} P_r^{jj} \right) s^n
\]
\[
= \begin{cases} 
\sum_{n=0}^{\infty} P_n^{jj} s^n = P_{jj}(s), & i \neq j, \\
\sum_{n=1}^{\infty} P_n^{jj} s^n = P_{ii}(s) - 1. & i = j.
\end{cases}
\]

We say a state \( i \) is recurrent if and only if \( \sum_{n=1}^{\infty} f_n^{ii} = 1 \). This says that a state \( i \) is recurrent if and only if, starting from state \( i \), the probability of returning to state \( i \) after some finite length of time is one. A non-recurrent state is said to be transient, i.e. \( \sum_{n=1}^{\infty} f_n^{ii} < 1 \). (Why \( \sum_{n=1}^{\infty} f_n^{ii} \leq 1 \)?)

We will prove a theorem that relates the recurrence or nonrecurrence of a state to the behavior of the n-step transition probabilities \( P_n^{ii} \). Before proving the theorem, we need the following:

**Lemma 2.5.1 (Abel)** (a) If \( \sum_{k=0}^{\infty} a_k \) converges, then
\[
\lim_{s \to 1^-} \sum_{k=0}^{\infty} a_k s^k = \sum_{k=0}^{\infty} a_k =: a
\]
(\( \lim_{s \to 1^-} \) means \( s \) tend to 1 from values less than 1).

(b) If \( a_k \geq 0 \) and \( \lim_{s \to 1^-} \sum_{k=0}^{\infty} a_k s^k = a \leq \infty \), then
\[
\sum_{k=0}^{\infty} a_k = \lim_{N \to \infty} \sum_{k=0}^{N} a_k = a.
\]
Theorem 2.5.1 (i) A state $i$ is recurrent iff $\sum_{n=0}^{\infty} P_{ii}^n = \infty$.

(ii) If $i \leftrightarrow j$ and $i$ is recurrent, then $j$ is recurrent.

Proof. (i)$\Rightarrow$ Assume $i$ is recurrent, i.e. $\sum_{n=1}^{\infty} f_{ii}^n = 1$. Then by Lemma 2.5.1(a) $\lim_{s \to 1^-} F_{ii}(s) = 1$. Thus

$$\lim_{s \to 1^-} P_{ii}(s) = \lim_{s \to 1^-} (1 - F_{ii}(s))^{-1} = \infty.$$ 

Applying Lemma 2.5.1(b), we have $\sum_{n=0}^{\infty} P_{ii}^n = \infty$.

$\Leftarrow$ Assume $i$ is transient, i.e. $\sum_{n=1}^{\infty} f_{ii}^n < 1$.

Using Lemma 2.5.1(a) we have $\lim_{s \to 1^-} F_{ii}(s) < 1$, so $\lim_{s \to 1^-} P_{ii}(s) < \infty$. Now appealing to Lemma 2.5.1(b), we have $\sum_{n=0}^{\infty} P_{ii}^n < \infty$. Contradiction.

(ii) Since $i \leftrightarrow j$, there exist $m, n \geq 1$ s.t. $P_{ij}^m > 0$, $P_{ji}^m > 0$. Let $v > 0$. We obtain

$$P_{jj}^{m+n+v} \geq P_{ji}^m P_{ii}^n P_{ij}^v,$$

and on summing,

$$\sum_{v=0}^{\infty} P_{jj}^{m+n+v} \geq \sum_{v=0}^{\infty} P_{ji}^m P_{ii}^n P_{ij}^v = P_{ji}^m P_{ij}^v \sum_{v=0}^{\infty} P_{ii}^v.$$

Hence if $\sum_{v=0}^{\infty} P_{ii}^v = \infty$, then $\sum_{v=0}^{\infty} P_{jj}^v = \infty$.

2.6 Examples of Recurrent Markov Chains

Example 1 Consider the one-dimensional random walk on the positive and negative integers, where at each transition the particle moves with probability $p$ one unit to the right and with probability $q$ one unit to the left $(p + q = 1)$.

$$P_{00}^{2n+1} = 0, n = 1, 2, \ldots \text{ and } P_{00}^{2n} = \binom{2n}{n} p^n q^n$$

We appeal now to Stirling’s formula, $n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}$

$$P_{00}^{2n} \sim \frac{(pq)^n}{\sqrt{\pi n}} = \frac{(4pq)^n}{\sqrt{\pi n}}. \quad pq \leq \frac{1}{4} \text{ with equality holding iff } p = q = \frac{1}{2}$$

Hence $\sum_{n=0}^{\infty} P_{00}^n = \infty$ if and only if $p = q = \frac{1}{2}$.

2.7 More on Recurrence

Define $Q_{ii} = P\{ \text{ a particle starting in state } i \text{ returns infinitely often to state } i \}$.

Theorem 2.7.1 State $i$ is recurrent or transient according to whether $Q_{ii} = 1$ or 0, respectively.

Proof. Let $Q_{ii}^N$ be defined as

$$Q_{ii}^N = \{ \text{ a particle starting in state } i \text{ returns to state } i \text{ at least } N \text{ times} \}$$

Then $Q_{ii}^N = \sum_{k=1}^{\infty} f_{ii}^k Q_{ii}^{N-1} = Q_{ii}^{N-1} f_{ii}^*$, where $f_{ii}^* = \sum_{k=1}^{\infty} f_{ii}^k$

Proceeding recursively, we obtain

$$Q_{ii}^N = f_{ii}^* Q_{ii}^{N-1} = (f_{ii}^*)^2 Q_{ii}^{N-2} = \cdots = (f_{ii}^*)^{N-1} Q_{ii}^1 = (f_{ii}^*)^N$$

Since $\lim_{N \to \infty} Q_{ii}^N = Q_{ii}$, we have $Q_{ii} = 1$ or 0 according to $f_{ii}^* = 1$ or < 1, respectively, or equivalently, according to whether state $i$ is recurrent or transient.
Theorem 2.7.2 If $i \leftrightarrow j$ and the class is recurrent, then

$$f_{ij}^* = \sum_{n=1}^{\infty} f_{ij}^n = 1.$$  

Proof. Since $i \leftrightarrow j$, $\exists N$, s.t. $f_{ji}^N > 0$. So

$$1 - f_{jj}^* = P\{\text{a particle starting in state } j \text{ never returns to state } j\} \geq f_{ji}^N (1 - f_{ij}^*)$$

noting that $1 - f_{ij}^*$ is the prob of never going to state $j$ from state $i$. If $f_{ij}^* < 1$, then $1 - f_{jj}^* > 0 \Rightarrow f_{ji}^* = 1 \Rightarrow$ state $j$ is transient. So we must have $f_{ij}^* = 1$.

Define $Q_{ij} = P\{ \text{a particle starting in state } i \text{ visits state } j \text{ infinitely often}\}$.

Corollary 2.7.1 If $i \leftrightarrow j$ and the class is recurrent, then $Q_{ij} = 1$.

Proof. It is easy to see that

$$Q_{ij} = f_{ij}^* Q_{jj}$$

Since $j$ is a recurrent state, by Theorem 2.7.1, $Q_{jj} = 1$. By Theorem 2.7.2 $f_{ij}^* = 1$. Hence $Q_{ij} = 1$. 


Chapter 3

The Basic Limit Theorem of Markov Chains and Applications

3.1 Discrete Renewal Equation

Theorem 3.1.1 Let \{a_k\}, \{u_k\}, \{b_k\} be sequences indexed by \(k = 0, \pm 1, \pm 2, \ldots\). Suppose that \(a_k \geq 0, \sum a_k = 1, \sum |k|a_k < \infty, \sum ka_k > 0, \sum |b_k| < \infty,\) and that the greatest common divisor of the integers \(k\) for which \(a_k > 0\) is 1. If the renewal equation

\[ u_n - \sum_{k=-\infty}^{\infty} a_{n-k}u_k = b_n \quad \text{for } n = 0, \pm 1, \pm 2, \ldots \]

is satisfied by a bounded sequence \(\{u_n\}\) of real numbers, then \(\lim_{n \to \infty} u_n\) and \(\lim_{n \to -\infty} u_n\) exist. Furthermore, if

\[ \lim_{n \to -\infty} u_n = 0, \text{ then } \lim_{n \to \infty} u_n = \frac{\sum_{k=-\infty}^{\infty} b_k}{\sum_{k=-\infty}^{\infty} ka_k} \]

In case \(\sum_{k=-\infty}^{\infty} ka_k = \infty\), the limit relations are still valid provided we interpret

\[ \frac{\sum_{k=-\infty}^{\infty} b_k}{\sum_{k=-\infty}^{\infty} ka_k} = 0 \]

The proof of this theorem in its general form as stated is beyond the scope of this book. Actually we will make use of this theorem only for the case where \(a_k, u_k, b_k\) vanish for negative values of \(k\), and \(b_k \geq 0\).

Remark In the case where \(a_{-k} = 0, b_{-k} = 0,\) and \(u_{-k} = 0,\) for \(k > 0\) the renewal equation becomes

\[ u_n - \sum_{k=0}^{n} u_{n-k}a_k = b_n \quad \text{for } n = 0, 1, 2, \ldots. \]

Remark (Reason for the term “renewal equation.”) Consider a light bulb whose lifetime, measured in discrete units, is a random variable \(\xi\) where

\[ P(\xi = k) = a_k, \quad \text{for } k = 0, 1, 2, \ldots, \quad a_k > 0, \quad \sum_{k=0}^{\infty} a_k = 1. \]
Let each bulb be replaced by a new one when it burns out. Suppose the first bulb lasts until time $\xi_1$, the second bulb until time $\xi_1 + \xi_2$, and the $n$-th bulb until time $\sum_{i=1}^n \xi_i$, where the $\xi_i$ are independent identically distributed random variables each distributed as $\xi$. Let $u_n$ denote the expected number of renewals (replacements) up to time $n$. If the first replacement occurs at time $k$ then the expected number of replacements in the remaining time up to $n$ is $u_n - k$, and summing over all possible values for $k$, we obtain

$$u_n = \sum_{k=0}^n (1 + u_{n-k})a_k + 0 \sum_{k=n+1}^\infty a_k$$

$$= \sum_{k=0}^n u_{n-k}a_k + \sum_{k=0}^n a_k$$

$$= \sum_{k=0}^n u_{n-k}a_k + b_n,$$

where $\sum_{k=0}^n a_k = b_n$. The reasoning behind (1.1) goes as follows. The term $1 + u_{n-k}$ is the expected number of replacements in time $n$ if the first bulb fails at time $k$ ($0 < k < n$), the probability of this event being $a_k$. The second sum is the probability that the first bulb lasts a duration exceeding $n$ time units. Taking account of the regenerative nature of the process, we may clearly evaluate $u_n$ by decomposing the possible realizations by the event of the time of the first replacement.

The following theorem describes the limiting behavior of $P_{ij}^n$, as $n \to \infty$ for all $i$ and $j$ in the case of an aperiodic recurrent Markov chain. The proof is a simple application of Theorem 3.1.1.

**Theorem 3.1.2 (The basic limit theorem of Markov chains.)**

Consider a recurrent irreducible aperiodic Markov chain.

(a) $\lim_{n \to \infty} P_{ii}^n = \frac{1}{\sum_{n=0}^\infty n f_{ii}}$

(b) $\lim_{n \to \infty} P_{ji}^n = \lim_{n \to \infty} P_{ii}^n$

**Proof.** (a) Note that (see Section 2.5)

$$P_{ii}^n - \sum_{k=0}^n f_{ii}^{n-k} P_{ii}^k = \begin{cases} 1, & n = 0; \\ 0, & n > 0. \end{cases}$$

Identify

$$u_n = \begin{cases} P_{ii}^n, & n \geq 0; \\ 0, & n < 0; \end{cases} \quad a_n = \begin{cases} f_{ii}^n, & n \geq 0; \\ 0, & n < 0; \end{cases} \quad b_n = \begin{cases} 1, & n = 0; \\ 0, & n \neq 0; \end{cases}$$

and then apply Theorem 3.1.1.

(b) We use the recursion relation

$$P_{ji}^n = \sum_{v=0}^n f_{ji}^v P_{ii}^{n-v} \quad i \neq j, \quad n \geq 0$$

More generally, let $y_n = \sum_{k=0}^n a_{n-k}x_k$, where $a_m \geq 0$, $\sum_{m=0}^\infty a_m = 1$, $\lim_{k \to \infty} x_k = c$. Under these circumstances we prove that $\lim_{n \to \infty} y_n = c$. (We will use this result several times in later chapters.) In fact,
\[
y_n - c = \sum_{k=0}^{n} a_{n-k}x_k - c \sum_{m=0}^{\infty} a_m = \sum_{k=0}^{n} a_{n-k}(x_k - c) - c \sum_{m=n+1}^{\infty} a_m
\]

For \(\epsilon > 0\) prescribed we determine \(K(\epsilon)\) so that \(|x_k - c| < \epsilon/3\) for all \(k \geq K(\epsilon)\). So

\[
|y_n - c| \leq \sum_{k=0}^{K(\epsilon)} a_{n-k}(x_k - c) + \sum_{k=K(\epsilon)+1}^{n} a_{n-k}(x_k - c) + c \sum_{m=n+1}^{\infty} a_m
\]

\[
\leq M \sum_{k=0}^{K(\epsilon)} a_{n-k} + \frac{\epsilon}{3} \sum_{k=K(\epsilon)+1}^{n} a_{n-k} + |c| \sum_{m=n+1}^{\infty} a_m
\]

where \(M = \max_{k \geq 0} |x_k - c|\). We choose \(N(\epsilon)\) so that \(|c| \sum_{m=n+1}^{\infty} a_m < \epsilon/3\) and

\[
\sum_{k=0}^{K(\epsilon)} a_{n-k} \equiv \sum_{m=n-K(\epsilon)}^{n} a_m < \frac{\epsilon}{3M} \text{ for } n \geq N(\epsilon)
\]

Then

\[
|y_n - c| \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \text{ for } n \geq N(\epsilon)
\]

Now setting

\[
y_n = P_{ji}^n, \quad a_n = f_{ji}^n, \quad x_n = P_{ii}^n
\]

and using Theorem 2.7.2 we have the desired result.

If \(\lim_{n \to \infty} P_{ii}^n = \pi_i > 0\), for one \(i\) in an aperiodic class, then we may show that \(\pi_j > 0\) for all \(j\) in the class of \(i\) \((P_{jj}^{m+n+v} \geq P_{jj}^{m}P_{ii}^nP_{ii}^v)\). So in this case, we call the class positive recurrent or strongly ergodic. If each \(\pi_i = 0\) and the class is recurrent we speak of the class as null recurrent or weakly ergodic.

**Theorem 3.1.3** In a positive recurrent aperiodic class with states \(j = 0, 1, 2, \cdots\),

\[
\lim_{n \to \infty} P_{jj}^n = \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad \sum_{i=0}^{\infty} \pi_i = 1
\]

and \(\pi\)'s are uniquely determined by the set of equations

\[
\pi_i \geq 0, \quad \sum_{i=0}^{\infty} \pi_i = 1, \quad \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}. \quad (1.2)
\]

Any set \((\pi_i)_{i=0}^{\infty}\) satisfying (1.2) is called a stationary probability distribution of the Markov chain.

**Proof.** For every \(n\) and \(M\), \(1 = \sum_{i=0}^{\infty} P_{ij}^n \geq \sum_{i=0}^{M} P_{ij}^n\). Letting \(n \to \infty\), and using Theorem 3.1.2, we obtain \(1 \geq \sum_{i=0}^{M} \pi_j\) for every \(M\). Thus \(\sum_{i=0}^{\infty} \pi_j \leq 1\). Now \(P_{ij}^{n+1} \geq \sum_{k=0}^{M} P_{ik}^n P_{kj}\); if we let \(n \to \infty\), we obtain \(\pi_j \geq \sum_{k=0}^{M} \pi_k P_{kj}\). Next, since the left-hand side is independent of \(M\), \(M \to \infty\) gives

\[
\pi_j \geq \sum_{k=0}^{\infty} \pi_k P_{kj}. \quad (1.3)
\]
Multiplying by $P_{ji}$, then summing on $j$ and using (1.3), yields $\pi_j \geq \sum_{k=0}^{\infty} \pi_k P_{kj}^2$ and then generally $\pi_j \geq \sum_{k=0}^{\infty} \pi_k P_{kj}^n$ for any $n$. Suppose strict inequality holds for some $j$. Adding these inequalities with respect to $j$, we have

$$\sum_{j=0}^{\infty} \pi_j > \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \pi_k P_{kj}^n = \sum_{k=0}^{\infty} \pi_k \sum_{j=0}^{\infty} P_{kj}^n = \sum_{k=0}^{\infty} \pi_k,$$

a contradiction. Thus, $\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}^n$ for any $n$. Letting $n \to \infty$, since $\sum \pi_k$ converges and $P_{kj}^n$ is uniformly bounded, we conclude that

$$\pi_j = \sum_{k=0}^{\infty} \pi_k \lim_{n \to \infty} P_{kj}^n = \pi_j \sum_{k=0}^{\infty} \pi_k$$

for every $j$.

Thus $\sum_{k=0}^{\infty} \pi_k = 1$ since $\pi_j > 0$ by positive recurrence.

Suppose $x = \{x_n\}$ satisfies the relations (1.2). Then

$$x_k = \sum_{j=0}^{\infty} x_j P_{jk} = \sum_{j=0}^{\infty} x_j P_{jk},$$

and if we let $n \to \infty$ as before,

$$x_k = \sum_{j=0}^{\infty} x_j \lim_{n \to \infty} P_{jk}^n = \pi_k \sum_{j=0}^{\infty} x_j = \pi_k.$$

**Example** Consider the class of random walks whose transition matrices are given by

$$P = \begin{pmatrix}
0 & 1 & 0 & 0 & \cdots \\
q_1 & 0 & p_1 & 0 & \cdots \\
0 & q_2 & 0 & p_2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}$$

This Markov chain has period 2. Nevertheless we investigate the existence of a stationary probability distribution, i.e., we wish to determine the positive solutions of

$$x_i = \sum_{j=0}^{\infty} x_j P_{ji} = p_{i-1} x_{i-1} + q_{i+1} x_{i+1}, \quad i = 0, 1, \cdots, \quad (1.4)$$

under the normalization $\sum_{i=0}^{\infty} x_i = 1$, where $p_{-1} = 0$ and $p_0 = 1$, and thus $x_0 = q_1 x_1$.

Using Eq. (1.4) for $i = 1$, we can determine $x_2$ in terms of $x_0$. Eq. (1.4) for $i = 2$ determines $x_3$ in terms of $x_0$, etc. By induction,

$$x_i = x_0 \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}}, \quad i \geq 1.$$ 

Now since

$$1 = x_0 + \sum_{i=1}^{\infty} x_0 \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}},$$

we have

$$x_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}}}.$$
and so
\[ x_0 > 0 \iff \sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} < \infty. \]
In particular, if \( p_k = p \) and \( q_k = q = 1 - p \) for \( k \geq 1 \), the series
\[
\sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} = \frac{1}{p} \sum_{i=1}^{\infty} \left( \frac{q}{p} \right)^i
\]
converges only when \( p < q \).

### 3.2 Absorption Probabilities

If \( T \) is the set of all transient states, then consider
\[
x_i^1 = \sum_{j \in T} P_{ij} \leq 1, \quad i \in T,
\]
and define
\[
x_i^n = \sum_{j \in T} P_{ij} x_j^{n-1}, \quad n \geq 2, \quad i \in T.
\]
Observe that \( x_i^n \) is just the probability that, starting from \( i \), the state of the process stays in \( T \) for the next \( n \) transitions. Since \( x_i^n \leq 1 \) for all \( n \geq 1 \) (they are probabilities), we may prove by induction that \( x_i^n \) is non-increasing as a function of \( n \). In fact,
\[
x_i^2 = \sum_{j \in T} P_{ij} x_j^1 \leq \sum_{j \in T} P_{ij} = x_i^1.
\]
Now assuming that \( x_j^n \leq x_j^{n-1} \) for all \( j \in T \) we have
\[
0 \leq x_i^{n+1} = \sum_{j \in T} P_{ij} x_j^n \leq \sum_{j \in T} P_{ij} x_j^{n-1} = x_i^n.
\]
Therefore, \( x_i^n \downarrow x_i \), i.e. \( x_i^n \) decreases to some limit \( x_i \), and
\[
x_i = \sum_{j \in T} P_{ij} x_j, \quad i \in T. \tag{2.5}
\]
It follows that if the only bounded solution of this set of equations is the zero vector \((0, 0, ...\), then starting from any transient state absorption into a recurrent class occurs with probability one. The reason is as follows. It is clear that \( x_i (i \in T) \) is the probability of never being absorbed into a recurrent class, starting from state \( i \). Since this sequence is a bounded solution of (2.5) it follows that \( x_i \) is zero for all \( i \).

Let \( C, C_1, C_2, \cdots \) denote recurrent classes. We define \( \pi_i(C) \) as the probability that the process will be ultimately absorbed into the recurrent class \( C \) if the initial state is the transient state \( i \). (Recall that once the process enters a recurrent class, it never leaves it. Otherwise, suppose \( i \) is in a recurrent class \( C \), \( j \) is not in \( C \). If the prob. from \( i \) to \( j \) is positive, then \( i \) is not a recurrent state since from \( j \) we can’t go back to \( C \) any more. )
Let \( \pi_i^n(C) \) = probability that the process will enter and thus be absorbed in \( C \) for the first time at the \( n \)-th transition, given that the initial state is \( i \in T \). Then

\[
\pi_i(C) = \sum_{n=1}^{\infty} \pi_i^n(C) \leq 1, \tag{2.6}
\]

\[
\pi_i^1(C) = \sum_{j \in C} P_{ij},
\]

\[
\pi_i^n(C) = \sum_{j \in T} P_{ij} \pi_j^{n-1}(C), \quad n \geq 2, \tag{2.7}
\]

noting that in (2.7), if from \( i \), we goes to \( j \) of other recurrent classes, then \( \pi_j^{n-1}(C) = 0 \).

Rewriting (2.6) using (2.7) gives

\[
\pi_i(C) = \pi_1^1(C) + \sum_{n=2}^{\infty} \pi_i^n(C) = \pi_1^1(C) + \sum_{n=2}^{\infty} \sum_{j \in T} P_{ij} \pi_j^{n-1}(C) = \pi_1^1(C) + \sum_{j \in T} P_{ij} \sum_{n=2}^{\infty} \pi_j^{n-1}(C),
\]

\[
\pi_i(C) = \pi_1^1(C) + \sum_{j \in T} P_{ij} \pi_j(C), \quad i \in T. \tag{2.8}
\]

**Theorem 3.2.1** Let \( j \in C \) which is an aperiodic recurrent class. Then for \( i \in T \), we have

\[
\lim_{n \to \infty} P_{ij}^n = \pi_i(C) \lim_{n \to \infty} P_{jj}^n = \pi_i(C) \pi_j.
\]

**Proof.** Clearly, \( \pi_i^n(C) = \sum_{k \in C} \pi_{ik}^n(C) \), where \( \pi_{ik}^n(C) \) is the prob. starting from \( i \) of being absorbed at the \( n \)-th transition into class \( C \) at state \( k \). We have

\[
\pi_i(C) = \sum_{v=1}^{\infty} \sum_{k \in C} \pi_{ik}^v(C) \leq 1.
\]

Therefore for any \( \varepsilon > 0 \) there exist a finite number of states \( C' \subset C \) and an integer \( N(\varepsilon) = N \) s.t.

\[
|\pi_i(C) - \sum_{v=1}^{n} \sum_{k \in C'} \pi_{ik}^v(C)| < \varepsilon, \quad \text{i.e.} \quad \left| \sum_{v=1}^{\infty} \sum_{k \in C} \pi_{ik}^v - \sum_{v=1}^{n} \sum_{k \in C'} \pi_{ik}^v \right| < \varepsilon \tag{2.9}
\]

for \( n > N(\varepsilon) \). Here we have abbreviated \( \pi_{ik}^v \) for \( \pi_{ik}^v(C) \).

For \( j \in C \) consider

\[
P_{ij}^n - \sum_{v=1}^{n} \sum_{k \in C'} \pi_{ik}^v \pi_j.
\]

Decomposing the events by the time of first entering some state in \( C \), we have

\[
P_{ij}^n = \sum_{v=1}^{n} \sum_{k \in C} \pi_{ik}^v P_{kj}^{n-v}, \quad i \in T, \ j \in C.
\]
Combining these relations, we have

\[ |P^n_{ij} - \left( \sum_{v=1}^{n} \sum_{k \in C'} \pi^v_{ik} \right) \pi_j | = \sum_{v=1}^{n} \sum_{k \in C'} \pi^v_{ik} (P^{n-v}_{kj} - \pi_j) + \sum_{v=1}^{n} \sum_{k \in C' \setminus C'} \pi^v_{ik} P^{n-v}_{kj} \]

\[ \leq \sum_{v=1}^{N} \sum_{k \in C'} \pi^v_{ik} (P^{n-v}_{kj} - \pi_j) | + \sum_{v=1}^{n} \sum_{k \in C' \setminus C'} \pi^v_{ik} P^{n-v}_{kj}. \]

But \( P^{n-v}_{kj} \leq 1, |P^{n-v}_{kj} - \pi_j| \leq 2, \) and \( \lim_{n \to \infty} P^{n-v}_{kj} = \pi_j \) if \( C \) is aperiodic and \( k \in C' \) (Theorems 3.1.2, 3.1.3). Hence there exists \( N' > N \) s.t. for \( n > N' \),

\[ |P^n_{ij} - \left( \sum_{v=1}^{n} \sum_{k \in C'} \pi^v_{ik} \right) \pi_j | \leq 2, \sum_{v=1}^{n} \sum_{k \in C'} \pi^v_{ik} + \sum_{v=1}^{n} \sum_{k \in C' \setminus C'} \pi^v_{ik} P^{n-v}_{kj}. \]

However, the choice of \( N \) and \( C \) assures us that the right-hand side is \( \leq 4 \varepsilon \). Then appealing to (2.9) and the above result, we obtain

\[ |P^n_{ij} - \pi_i(C) \pi_j | \leq 4 \varepsilon + \varepsilon \pi_j \text{ for } n > N', \]

and therefore

\[ \lim_{n \to \infty} P^n_{ij} = \pi_i(C) \lim_{n \to \infty} P^n_{jj} = \pi_i(C) \pi_j. \]

We emphasize the fact that if \( i \) is a transient state and \( j \) is a recurrent state, then the limit of \( P^n_{ij} \) depends on both \( i \) and \( j \). This is in sharp contrast with the case where \( i \) and \( j \) belong to the same recurrent class.

**Example** (The gambler’s ruin on \( n + 1 \) states).

\[
\mathbf{P} = \begin{pmatrix}
1 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
q & 0 & p & 0 & \ldots & \ldots & \ldots \\
0 & q & 0 & p & \ldots & \ldots & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & q & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{pmatrix}
\]

We shall calculate \( u_i = \pi_i(C_0) \) and \( v_i = \pi_i(C_n) \), the probabilities that starting from \( i \) the process ultimately enters the absorbing (and therefore recurrent) states 0 and \( n \), respectively. The system of equations (2.8) becomes

\[ u_1 = q + pu_2 \]

\[ u_i = qu_{i-1} + pu_{i+1}, \quad 2 \leq i \leq n - 2. \tag{2.10} \]

\[ u_{n-1} = qu_{n-2}. \]

We try a solution of the form \( u_r = x^r \). Substituting in the middle equations and removing common factors leads to \( px^2 + q = x \). There are two solutions, \( x = 1 \) and \( x = q/p \). Thus
\( u_r = A + B(q/p)^r, \ r = 1, 2, \ldots, n - 1, \) satisfy the middle equations of (2.10) for any \( A \) and \( B \). We now determine \( A \) and \( B \) so that the first and last equations are fulfilled. (If \( q = p \), the solution \( x = 1 \) is a double root of \( px^2 + q = x \), and one then has to replace \((q/p)^r\) by \( r\).) In the case this leads to the conditions

\[
A + B \frac{q}{p}^r = q + p \left( A + B \frac{q^2}{p^2} \right)
\]
or simplifying,

\[
A = 1 - B
\]

and

\[
A + B \left( \frac{q}{p} \right)^{n-1} = q \left( A + B \left( \frac{q}{p} \right)^{n-2} \right) \quad \text{or} \quad p^n A + q^n B = 0
\]

Solving, we get

\[
A = \frac{q^n}{q^n - p^n}, \quad B = \frac{-p^n}{q^n - p^n}
\]

Combining, we have

\[
u_r = \frac{(q/p)^n - (q/p)^r}{(q/p)^n - 1} \quad \text{if} \quad \frac{q}{p} \neq 1
\]

If \( q = p \) we find similarly that \( A = 1, B = -1/n \) so that

\[
u_r = \frac{n - r}{n} \quad \text{when} \quad p = q
\]

A similar calculation shows that

\[
u_i = 1 - u_i
\]

which is to be expected, since it is evident that absorption into one of the classes \( C_0, C_n \) is certain.

Consider the gambler’s ruin with an infinitely rich adversary. The equations for the probability of the gambler’s ruin (absorption into 0) become

\[
u_1 = q + pu_2, \quad (2.11)
\]

\[
u_i = qu_{i-1} + pu_{i+1}, \quad i \geq 2 \quad (2.12)
\]

Again we find

\[
u_i = A + B \left( \frac{q}{p} \right)^i \quad (q \neq p) \quad \text{and} \quad \nu_i = A + Bi \quad (q = p = \frac{1}{2})
\]

If \( q \geq p \) then the condition that \( \nu_i \) is bounded requires that \( B = 0 \) and the first equation of (2.12) shows that \( \nu_i \equiv 1 \). If \( q < p \) we find that \( \nu_i = (q/p)^i \). In fact, a simple passage to the limit from the finite state gambler’s ruin yields \( \nu_1 = q/p \) and then it readily follows that \( \nu_i = (q/p)^i \).
3.3 Criteria for Recurrence

We prove two theorems which will be useful in determining whether a given Markov chain is recurrent or transient.

**Theorem 3.3.1** Let $B$ be an irreducible Markov chain whose state space is labeled by the nonnegative integers. Then a necessary and sufficient condition that $B$ be transient (i.e., each state is a transient state) is that the system of equations

$$\sum_{j=0}^{\infty} P_{ij} y_j = y_i, \quad i \neq 0, \quad (3.13)$$

have a bounded nonconstant solution.

**Proof.** Let the transition matrix for $B$ be

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

and associate with it the new transition matrix

$$\tilde{P} = \left(\tilde{P}_{ij}\right) = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (3.14)$$

We denote the Markov chain with transition probability matrix (3.14) by $\tilde{B}$. For the necessity, we shall assume that the process is transient and then exhibit a nonconstant bounded solution of (3.13).

Let $f_{i0}^* =$ probability of entering state 0 in some finite time, given that $i$ is the initial state ($f_{i0}^* = \sum_{n=1}^{\infty} f_{i0}^n$). Since the process $B$ is transient $f_{j0}^* < 1$ for some $j \neq 0$. (If $f_{j0}^* = 1$ for all $j \neq 0$, then $f_{00}^* = \sum_{n=1}^{\infty} f_{00}^n = f_{00} + \sum_{n=2}^{\infty} \sum_{j=1}^{\infty} P_{0j} f_{j0}^{n-1} = P_{00} + \sum_{j=1}^{\infty} P_{0j} f_{j0}^* = \sum_{j=0}^{\infty} P_{0j} = 1$)

For the process $\tilde{B}$ clearly $\tilde{\pi}_0(C_0) = 1, \tilde{\pi}_j(C_0) = f_{j0}^* < 1$ for some $j \neq 0$, and $\tilde{\pi}_i(C_0) = \sum_{j=0}^{\infty} \tilde{P}_{ij} \tilde{\pi}_j(C_0)$ for all $i$. Hence $\tilde{\pi}_i(C_0) = \sum_{j=0}^{\infty} \tilde{P}_{ij} \tilde{\pi}_j(C_0)$ for $i \neq 0$ and thus $y_j = \tilde{\pi}_j(C_0)(j = 0, 1, 2, \ldots)$ is the desired bounded nonconstant solution.

Now assume that we have a nonconstant bounded solution $\{y_i\}$ of (3.13). Then

$$\sum_{j=0}^{\infty} \tilde{P}_{ij} y_j = y_i, \text{ for all } i \geq 0,$$

and iterating we have for all $i > 0$ and all $n \geq 1$

$$\sum_{j=0}^{\infty} \tilde{P}_{ij}^n y_j = y_i$$

If the chain $B$ is recurrent, then

$$\lim_{n \to \infty} \tilde{P}_{i0}^n = 1$$
(since in $B$, $\sum_{n=1}^{\infty} f^n_{i0} = 1$, and $\tilde{P}^m_{i0} \geq \sum_{n=1}^{m} f^n_{i0}$) and

$$\sum_{j \neq 0} \tilde{P}^n_{0j} y_j \leq M (1 - \tilde{P}^n_{00}) \to 0 \text{ as } n \to \infty$$

where $M$ is a bound for $\{y_j\}$. Hence

$$y_i = \sum_{j \neq 0} \tilde{P}^n_{0j} y_j + \tilde{P}^n_{00} y_0 \to y_0.$$ 

Thus $y_i = y_0$ for all $i$ and $\{y_j\}$ is constant.

**Theorem 3.3.2** In an irreducible Markov chain a sufficient condition for recurrence is that there exists a sequence $\{y_i\}$ such that

$$\sum_{j=0}^{\infty} P_{ij} y_j \leq y_i \text{ for } i \neq 0 \text{ with } y_i \to \infty. \quad (3.15)$$

**Proof.** Using the same notation as in the previous theorem, we have

$$\sum_{j=0}^{\infty} \tilde{P}_{ij} y_j \leq y_i \text{ for all } i.$$ 

Since $z_i = y_i + b$ satisfies (3.15), we may assume $y_i > 0$ for all $i \geq 0$. Iterating the preceding inequality, we have

$$\sum_{j=0}^{\infty} \tilde{P}^m_{ij} y_j \leq y_i$$

Given $\varepsilon > 0$ we choose $M(\varepsilon)$ such that $1/y_i \leq \varepsilon$ for $i \geq M(\varepsilon)$. Now

$$\sum_{j=0}^{M-1} \tilde{P}^m_{ij} y_j + \sum_{j=M}^{\infty} \tilde{P}^m_{ij} y_j \leq y_i$$

and so

$$\sum_{j=0}^{M-1} \tilde{P}^m_{ij} y_j + \min_{r \geq M} \{y_r\} \sum_{j=M}^{\infty} \tilde{P}^m_{ij} y_j \leq y_i.$$ 

Since

$$\sum_{j=0}^{\infty} \tilde{P}^m_{ij} = 1$$

we have

$$\sum_{j=0}^{M-1} \tilde{P}^m_{ij} y_j + \min_{r \geq M} \{y_r\} \left(1 - \sum_{j=0}^{M-1} \tilde{P}^m_{ij}\right) \leq y_i.$$
Suppose $B$ is transient, noting in $\tilde{B}$, 0 is an absorbing state, we see that
\[
\lim_{n \to \infty} \tilde{P}_ij^n \leq \lim_{n \to \infty} P_{ij}^n = 0 \quad for \quad j \neq 0.
\]
Thus, passing to the limit as $m \to \infty$, we obtain for each fixed $i \neq 0$
\[
\tilde{\pi}_i(C_0)y_0 + \min_{r \geq M}\{y_r\}(1 - \tilde{\pi}_i(C_0)) \leq y_i
\]
\[
1 - \tilde{\pi}_i(C_0) \leq \frac{1}{\min_{r \geq M}\{y_r\}}(y_i - \tilde{\pi}_i(C_0)y_0) \leq \varepsilon K
\]
where $K = y_i - \tilde{\pi}_i(C_0)y_0$. Letting $\varepsilon \to 0$, $\tilde{\pi}_i(C_0) = 1$ for each $i \neq 0$. The original process is recurrent.