Chapter 4. Markov Chains
Example. (Weather Forecasting 2)

Suppose that whether it rains or not depends on previous weather conditions through the last two days. If it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2. Given that it rains on Monday and Tuesday, what is the probability that it will rains on Thursday?

Define the state space as follows:

- state 0 (rain, rain) — it rained today and yesterday
- state 1 (sun, rain) — it rained today but not yesterday
- state 2 (rain, sun) — it rained yesterday but not today
- state 3 (sun, sun) — it rained neither today nor yesterday

The one-step transition matrix is

\[
P = \begin{bmatrix}
0.7 & 0 & 0.3 & 0 \\
0.5 & 0 & 0.5 & 0 \\
0 & 0.4 & 0 & 0.6 \\
0 & 0.2 & 0 & 0.8 \\
\end{bmatrix}
\]
Other Examples

- **Queuing Systems**
  - number of customers waiting in line at a bank *at the beginning of each hour*
  - number of machines waiting for repair *at the end of each day*
  - number of patients waiting for a lung transplant *at beginning of each week*

- **Inventory Management**
  - number of units in stock *at beginning of each week*
  - number of backordered units *at end of each day*

- **Airline Overbooking**
  - number of coach seats reserved *at beginning of each day*

- **Finance**
  - price of a given security *when market closes each day*
Other Examples (con’t)

- Epidemiology
  - number of foot-and-mouth infected cows at the beginning of each day
- Population Growth
  - size of US population at end of each year
- Genetics
  - genetic makeup of your descendants in each subsequent generation
- Workforce Planning
  - number of employees in a firm with each level of experience at the end of each year
Introduction

- Consider a stochastic process \( \{X_n, n = 0, 1, 2, \ldots\} \) that takes on a finite or countable number of possible values denoted by the set of nonnegative integers \( \{0, 1, 2, \ldots\} \).
- If \( X_n = i \), then the process is said to be in state \( i \) at time \( n \).
- Suppose that whenever the process is in state \( i \), there is a fixed probability \( P_{ij} \) that it will next be in state \( j \). That is,

\[
P\{X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0\}
= P\{X_{n+1} = j|X_n = i\} \quad \text{Independent of past states}
= P_{ij} \quad \text{Independent of } n
\]

for all states \( i_0, i_1, \ldots, i_{n-1}, i, j \) and all \( n \geq 0 \).
\( \Rightarrow \) Such a stochastic process is known as a Markov chain.
**Introduction**

- *Markovian* property: For a Markov chain, the conditional distribution of any future state $X_{n+1}$, given the past states $X_0, X_1, ..., X_{n-1}$ and the present state $X_n$, is independent of the past states and depends only on the present state.

- The value $P_{ij}$ represents the probability that the process will, when in state $i$, next make a transition into state $j$.

- Since probabilities are nonnegative and since the process must make a transition into some state, we have that

  $$P_{ij} \geq 0, \quad i, j \geq 0; \quad \sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, ...$$
Introduction

- Let $P$ denote the matrix of one-step transition probabilities $P_{ij}$, so that

$$P = \begin{bmatrix}
    P_{00} & P_{01} & P_{02} & \ldots \\
    P_{10} & P_{11} & P_{12} & \ldots \\
    \vdots & \vdots & \vdots & \ddots \\
    P_{i0} & P_{i1} & P_{i2} & \ldots \\
    \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.$$
Transforming a Process into a Markov Chain

Example. (Weather Forecasting)
Suppose that whether it rains or not depends on previous weather conditions through the last two days. If it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

Question: How to model the process as a Markov chain? What is the one-step transition matrix?

Let \( X_n \) = raining or not at day \( n \)

Is \( X_n \) a Markov chain? **No!**

It doesn’t satisfy the Markov property since the conditional probability of rain at day \( n + 1 \) depends on the weather at day \( n \) and day \( n - 1 \).
Let the state of the process be todays and yesterdays weather and define the state space as follows:

- state 0 (rain, rain) it rained today and yesterday
- state 1 (sun, rain) it rained today but not yesterday
- state 2 (rain, sun) it rained yesterday but not today
- state 3 (sun, sun) it rained neither today nor yesterday

The transition matrix for this Markov chain is:

\[
P = \begin{bmatrix}
0.7 & 0 & 0.3 & 0 \\
0.5 & 0 & 0.5 & 0 \\
0 & 0.4 & 0 & 0.6 \\
0 & 0.2 & 0 & 0.8 \\
\end{bmatrix}
\]
Example 1. The $M/G/1$ Queue

- The $M/G/1/\infty$ Queue.
  - The customers arrive at a service center in accordance with a Poisson process with rate $\lambda$.
  - There is a single server and those arrivals finding the server free go immediately into service; all others wait in line until their service turn.
  - The service times of successive customers are assumed to be independent random variables having a common distribution $G$;
  - The service times are also assumed to be independent of the arrival process.
Example 1. The M/G/1 Queue

- Let $X(t)$ denote the number of customers in the system at $t$.

Does $\{X(t), t \geq 0\}$ possess the Markovian property??

The answer is No!!

Proof:

Assume the current time point is $t$. It can be seen that at the next time point $t'$, the probability of $k$ new arrivals is with probability $e^{-\lambda(t'-t)} \frac{(\lambda(t'-t))^{k}}{k!}$, indepdendent of $t$ as well as the state before $t$. However, the number of customers depart from $t$ to $t'$ depends on individual customer’ arrival time (which may be before $t$) so that the number of customers at $t'$ depends on past and doesn’t possess Markovian Property.
Example 1. The M/G/1 Queue

• For if we knew the number in the system at time $t$, then to predict future behavior
  
  – we would not care how much time had elapsed since the last arrival (since the arrival process is memoryless)
  – we would care how long the person in service had already been there (since the service distribution $G$ is arbitrary and therefore not memoryless)

• Let us only look at the system at moments when customers ”depart”. 
  
  – let $X_n$ denote the number of customers left behind by the $n_{th}$ departure, $n \geq 1$
  – let $Y_n$ denote the number of customers arriving during the service period of the $(n + 1)st$ customer
Example 1. The M/G/1 Queue

- When $X_n > 0$, the $n_{th}$ departure leaves behind $X_n$ customers - of which one enters service and the other $X_n - 1$ wait in line.

- Hence, at the next departure the system will contain the $X_n - 1$ customers that were in line in addition to any arrivals during the service time of the $(n + 1)$st customer. Since a similar argument holds when $X_n = 0$, we see that

$$X_{n+1} = \begin{cases} X_n - 1 + Y_n & \text{if } X_n > 0 \\ Y_n & \text{if } X_n = 0. \end{cases}$$
Example 1. The M/G/1 Queue

- Since $Y_n, n \geq 1$, represent the number of arrivals in nonoverlapping service intervals, it follows, the arrival process being a Poisson process, that they are independent and

$$P\{Y_n = j\} = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} dG(x), \quad j = 0, 1, 2, \ldots$$

- From the above, it follows that $\{X_n, n = 1, 2, \ldots\}$ is a Markov chain with transition probabilities given by

$$P_{0j} = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} dG(x), \quad j \geq 0,$$

$$P_{ij} = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^{j-i+1}}{(j-i+1)!} dG(x), \quad j \geq i - 1, i \geq 1,$$

$$P_{ij} = 0 \quad \text{otherwise.}$$
Example 2. The G/M/1 Queue

- The G/M/1 Queue.
  - The customers arrive at a single-server service center in accordance with an arbitrary renewal process having inter-arrival distribution $G$.
  - The service distribution is exponential with rate $\mu$.

- Let us only look at the system at moments when customers "arrive".
  - Let $X_n$ denote the number of customers in the system as seen by the $n_{th}$ arrival.
  - 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 $\mu t$. Therefore,

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

$\Rightarrow \{X_n, n = 1, 2, \ldots\}$ is a Markov chain.
Example 2. The G/M/1 Queue

- The above equation follows since if an arrival finds $i$ in the system, then the next arrival will find $i + 1$ minus the number served, and the probability that $j$ will be served is easily seen (by conditioning on the time between the successive arrivals) to equal the right-hand side.

- The formula for $P_{i0}$ is little different (it is the probability that at least $i + 1$ Poisson events occur in a random length of time having distribution $G$) and thus is given by

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

- Remark:
  Note that in the previous two examples we were able to discover an *embedded* Markov chain by looking at the process only at certain time points, and by choosing these time points so as to exploit the lack of memory of the exponential distribution.
Example 3. The General Random Walk

The general random walk: sums of independent, identically distributed random variable.

- Let $X_i, i \geq 1$ be independent and identically distributed with

  $$P\{X_i = j\} = a_j, \ j = 0, \pm 1, \ldots$$

- If we let

  $$S_0 = 0 \text{ and } S_n = \sum_{i=1}^{n} X_i$$

  then $\{S_n, n \geq 0\}$ is a Markov chain for which

  $$P_{ij} = a_{j-i}$$

- $\{S_n, n \geq 0\}$ is called the general random walk.
Example 4. The Simple Random Walk

- The random walk \( \{S_n, n \geq 1\} \), where \( S_n = \sum_{1}^{n} X_i \), is said to be a simple random walk if for some \( p, 0 < p < 1 \),
  \[
P\{X_i = 1\} = p \\
P\{X_i = -1\} = q \equiv 1 - p
\]

Thus in the simple random walk the process always either goes up one step (with probability \( p \)) or down one step (with probability \( q \)).

- Consider \(|S_n|\), the absolute value of the simple random walk. The process \( \{|S_n|, n \geq 1\} \) measures at each time unit the absolute distance of the simple random walk from the origin.

- To prove \( \{|S_n|\} \) is itself a Markov chain, we first show that if \(|S_n| = i\), then no matter what its previous values the probability that \( S_n \) equals \( i \) (as opposed to \(-i\)) is \( p^i/(p^i + q^i) \).
Example 4. The Simple Random Walk (con’t)

Proposition.
If \( \{S_n, n \geq 1\} \) is a simple random walk, then

\[
P\{S_n = i \mid |S_n| = i, |S_{n-1}| = i_{n-1}, \ldots, |S_1| = i_1\} = \frac{p^i}{p^i + q^i}.
\]

Proof.
If we let \( i_0 = 0 \) and define

\[
j = \max\{k : 0 \leq k \leq n : i_k = 0\},
\]

then, since we know the actual value of \( S_j \), it is clear that

\[
P\{S_n = i \mid |S_n| = i, |S_{n-1}| = i_{n-1}, \ldots, |S_1| = i_1\} = P\{S_n = i \mid |S_n| = i, \ldots, |S_{j+1}| = i_{j+1}, |S_j| = 0\}.
\]
Example 4. The Simple Random Walk (con’t)

Now there are two possible values of the sequence $S_{j+1}, \ldots, S_n$ for which $|S_{j+1}| = i_{j+1}, \ldots, |S_n| = i$. The first of which results in $S_n = i$ and has probability

$$p \frac{n-j}{2} + \frac{i}{2} q \frac{n-j}{2} - \frac{i}{2},$$

and the second results in $S_n = -i$ and has probability

$$p \frac{n-j}{2} - \frac{i}{2} q \frac{n-j}{2} + \frac{i}{2}. $$

Hence,

$$P\{S_n = i | |S_n| = i, \ldots, |S_1| = i_1\} = \frac{p \frac{n-j}{2} + \frac{i}{2} q \frac{n-j}{2} - \frac{i}{2}}{p \frac{n-j}{2} + \frac{i}{2} q \frac{n-j}{2} - \frac{i}{2} + p \frac{n-j}{2} - \frac{i}{2} q \frac{n-j}{2} + \frac{i}{2}}$$

$$= \frac{p^i}{p^i + q^i}$$

and the proposition is proven.
Example 4. The Simple Random Walk (con’t)

- From the proposition, it follows upon conditioning on whether $S_n = +i$ or $-i$ that

$$P\{ |S_{n+1}| = i + 1 | S_n = i, S_{n-1}, \ldots, S_1 \}$$

$$= P\{ S_{n+1} = i + 1 | S_n = i \} \frac{p^i}{p^i + q^i}$$

$$+ P\{ S_{n+1} = -(i + 1) | S_n = -i \} \frac{q^i}{p^i + q^i} = \frac{p^{i+1} + q^{i+1}}{p^i + q^i}.$$

- Hence, $\{ |S_n| , n \geq 1 \}$ is a Markov chain with transition probabilities

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

$$P_{01} = 1.$$
Other Examples

- **Queuing Systems**
  - number of customers waiting in line at a bank at the beginning of each hour
  - number of machines waiting for repair at the end of each day
  - number of patients waiting for a lung transplant at beginning of each week

- **Inventory Management**
  - number of units in stock at beginning of each week
  - number of backordered units at end of each day

- **Airline Overbooking**
  - number of coach seats reserved at beginning of each day

- **Finance**
  - price of a given security when market closes each day
Other Examples (con’t)

- Epidemiology
  - number of foot-and-mouth infected cows at the beginning of each day

- Population Growth
  - size of US population at end of each year

- Genetics
  - genetic makeup of your descendants in each subsequent generation

- Workforce Planning
  - number of employees in a firm with each level of experience at the end of each year
Chapman-Kolmogorov Equations

- $P_{ij}$ : the one-step transition probabilities

- Define the $n$-step transition probabilities $P_{ij}^n$ to be the probability that a process in state $i$ will be in state $j$ after $n$ additional transitions. That is,

$$P_{ij}^n = P\{X_{n+m} = j \mid X_m = i\}, \quad n \geq 0, \ i, j \geq 0$$

where, of course, $P_{ij}^1 = P_{ij}$.

- The *Chapman – Kolmogorov equations* provide a method for computing these $n$-step transition probabilities. These equations are

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \quad \text{for all } n, m \geq 0, \ \text{all } i, j,$$
Chapman-Kolmogorov Equations (con’t)

and are established by observing that

\[ P_{ij}^{n+m} = P\{X_{n+m} = j | X_0 = i\} \]

\[ = \sum_{k=0}^{\infty} P\{X_{n+m} = j, X_n = k | X_0 = i\} \]

\[ = \sum_{k=0}^{\infty} P\{X_{n+m} = j | X_n = k, X_0 = i\} P\{X_n = k | X_0 = i\} \]

\[ = \sum_{k=0}^{\infty} P_{kj}^m P_{ik}^n. \]
Chapman-Kolmogorov Equations (con’t)

- Let $P^{(n)}$ denote the matrix of $n$-step transition probabilities $P_{ij}^n$, then the Chapman-Kolmogorov equations assert that

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

where the dot represents matrix multiplication.

- Hence,

$$P^{(n)} = P \cdot P^{(n-1)} = P \cdot P \cdot P^{(n-2)} = \cdots = P^n,$$

and thus $P^{(n)}$ may be calculated by multiplying the matrix $P$ by itself $n$ times.

- State $j$ is said to be accessible from state $i$ if for some $n \geq 0$, $P_{ij}^n > 0$.

- Two states $i$ and $j$ accessible to each other are said to communicate and is denoted by $i \leftrightarrow j$. 
Example. (Weather Forecasting 1)

Suppose that the chance of rain tomorrow depends on whether or not it is raining today. Suppose also that if it rains today, then it will rain tomorrow with probability 0.7; and if it does not rain today, ten it will rain tomorrow with probability 0.4. Given that it is rain today, what is the probability that it will rain four days from today?

The one-step probability matrix is given by

\[
P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}
\]

Hence,

\[
P^2 = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix}
\]

\[
P^4 = (P^2)^2 = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix} \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix} = \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}
\]

⇒ The desired probability \(P_{0,0}^4 = 0.5749\).

**Note:** The probability it will rain four successive days from today \((P_{0,0})^4\).
Example. (Weather Forecasting 2)

Suppose that whether it rains or not depends on previous weather conditions through the last two days. If it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2. Given that it rains on Monday and Tuesday, what is the probability that it will rains on Thursday?

Define the state space as follows:

- **state 0** (rain, rain) it rained today and yesterday
- **state 1** (sun, rain) it rained today but not yesterday
- **state 2** (rain, sun) it rained yesterday but not today
- **state 3** (sun, sun) it rained neither today nor yesterday

The one-step transition matrix is

\[
P = \begin{bmatrix}
0.7 & 0 & 0.3 & 0 \\
0.5 & 0 & 0.5 & 0 \\
0 & 0.4 & 0 & 0.6 \\
0 & 0.2 & 0 & 0.8 \\
\end{bmatrix}
\]
Example. (Weather Forecasting 2) (con’t)

<table>
<thead>
<tr>
<th>Mon.</th>
<th>Tue.</th>
<th>Wed.</th>
<th>Thur.</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain</td>
<td>rain</td>
<td>?</td>
<td>rain</td>
</tr>
</tbody>
</table>

Yesterday & Today (state 0)

Today & Tomorrow

Note that rain on Thursday is equivalent to the process transforming to state 0 or state 1 on Thursday.

\[ P^2 = \begin{bmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{bmatrix} \begin{bmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.49 & 0.12 & 0.21 & 0.18 \\ 0.35 & 0.20 & 0.15 & 0.30 \\ 0.20 & 0.12 & 0.20 & 0.48 \\ 0.10 & 0.16 & 0.10 & 0.64 \end{bmatrix} \]

The desired probability is given by \( P^2_{0,0} + P^2_{0,1} = 0.49 + 0.12 = 0.61 \)
Classification of States

Proposition.
Communication is an equivalence relation. That is:

1. \( i \leftrightarrow i; \) (reflexive)

2. if \( i \leftrightarrow j, \) then \( j \leftrightarrow i; \) (symmetric)

3. if \( i \leftrightarrow j \) and \( j \leftrightarrow k, \) then \( i \leftrightarrow k. \) (transient)

Proof.
The first two parts follow trivially from the definition of communication. To prove 3., suppose that \( i \leftrightarrow j \) and \( j \leftrightarrow k; \) then there exists \( m, n \) such that \( P_{ij}^m > 0, P_{jk}^n > 0. \) Hence,

\[
P_{ik}^{m+n} = \sum_{r=0}^{\infty} P_{ir}^m P_{rk}^n \geq P_{ij}^m P_{jk}^n > 0.
\]

Similarly, we may show there exists an \( s \) for which \( P_{ki}^s > 0. \)
Classification of States (con’t)

- Two states that communicate are said to be in the same *class*; and by the above proposition, any two classes are either disjoint or identical.

\[ E = C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5 \]

**Ex:**
Consider a Markov chain consisting of three states 0, 1, 2, and 3 and having transition matrix as below. How many classes are in the Markov chain?

\[
P = \begin{bmatrix}
0.5 & 0.5 & 0 & 0 \\
0.5 & 0.5 & 0 & 0 \\
0.25 & 0.25 & 0.25 & 0.25 \\
0 & 0 & 0 & 1
\end{bmatrix}.
\]

=> Three classes \{0, 1\}, \{2\}, \{3\}
Classification of States (con’t)

- We say that the Markov chain is irreducible if there is only one class, i.e., all states communicate with each other.

**Ex:**
Consider a Markov chain consisting of three states 0, 1, 2, and 3 and having transition matrix

\[ P = \begin{bmatrix}
0.5 & 0.5 & 0 \\
0.5 & 0.25 & 0.25 \\
0 & 0.4 & 0.6
\end{bmatrix} \]

Is the Markov chain irreducible?

All states communicate with each other \(\Rightarrow\) It is irreducible.
Classification of States (con’t)

• Transient
  – A state is transient if it is possible to leave the state and never return.

• Periodic
  – A state is periodic if it is not transient, and if that state is returned to only on multiples of some positive integer greater than 1. This integer is known as the period of the state.

• Ergodic
  – A state is ergodic if it is neither transient nor periodic.
    • An absorbing state is a state that returns to itself immediately with probability 1

Tip1: Always find the communicating classes before classifying the states.
Tip2: Identify states in the order of the definition:
  transient states  -> periodic states second -> ergodic states.
Tip3: A state that has a loop cannot be periodic
Tip4: If a communicating class is periodic, each state has the same period
Classification of States (con’t)  \( g.c.d: \) Greatest Common Denominator

- State \( i \) is said to have **period** \( d \) if \( P_{ii}^n = 0 \) whenever \( n \) is not divisible by \( d \) and \( d \) is the greatest integer with this property.
  
  - Let \( d(j) \) denote the period of \( j \).
    \[
    d(j) = \text{g.c.d} \left\{ n \geq 1 : p_{jj}^n > 0 \right\}.
    \]

**Ex:** A Periodic Markov Chain

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

- If state \( j \) has period \( k \), then starting in state \( j \), you can only return to \( j \) after a multiple of \( k \) steps.
- If \( P_{ii}^n = 0 \) for all \( n > 0 \), then define the period of \( i \) to be infinite.

- A state with period 1 is said to be **aperiodic**.
• No. of communicating classes: 5
  \{ a, b, c \} : transient
  \{ d \} : transient
  \{ e \} : absorbing, ergodic
  \{ f, g, h, i, j \} : periodic w/ d=3
  \{ k, l, m \} : ergodic

1. f, g, h, i, j form a communicating class.
2. Once you enter, you never leave
   -> not transient.
3. Look at state h
   Once we leave it we will always return
   in 3 transitions
   -> period=3
   -> f, g, h, i, j all have period=3
   -> f, g, h, i, j are periodic
A Graph of Period Two.

A Non-periodic Graph.
Examples of Periodicity

Example: Gambler’s Ruin

\[ P = \begin{bmatrix}
    0.7 & 0 & 0.3 & 0 \\
    0.5 & 0 & 0.5 & 0 \\
    0 & 0.4 & 0 & 0.6 \\
    0 & 0.2 & 0 & 0.8 \\
\end{bmatrix} \]

\[ p_{1,1}^n > 0 \text{ for } n = 2, 4, 6, \ldots \Rightarrow d(1) = 2 \]
\[ p_{2,2}^n > 0 \text{ for } n = 2, 4, 6, \ldots \Rightarrow d(2) = 2 \]
\[ p_{3,3}^n > 0 \text{ for } n = 2, 4, 6, \ldots \Rightarrow d(3) = 2 \]
\[ p_{0,0}^n > 0 \text{ for } n = 1, 2, 3, \ldots \Rightarrow d(0) = 1 \]
\[ p_{4,4}^n > 0 \text{ for } n = 1, 2, 3, \ldots \Rightarrow d(4) = 1 \]
\[ \Rightarrow \text{States 0 and 4 are aperiodic.} \]

Example: Weather Prediction

\[ p_{0,0}^n > 0 \text{ for } n = 1, 2, 3, \ldots \Rightarrow d(0) = 1 \]
\[ p_{1,1}^n > 0 \text{ for } n = 2, 3, 4, \ldots \Rightarrow d(1) = 1 \]
\[ p_{2,2}^n > 0 \text{ for } n = 2, 3, 4, \ldots \Rightarrow d(2) = 1 \]
\[ p_{3,3}^n > 0 \text{ for } n = 1, 2, 3, \ldots \Rightarrow d(3) = 1 \]
\[ \Rightarrow \text{All states are aperiodic.} \]
Classification of States (con’t)

Proposition. *(Periodicity is a class property)*
If \( i \leftrightarrow j \), then \( d(i) = d(j) \).

Proof.

- Let \( m \) and \( n \) be such that \( P_{ij}^m \cdot P_{ji}^n > 0 \), and suppose that \( P_{ii}^s > 0 \). Then
  \[
P_{jj}^{n+m} \geq P_{ji}^n \cdot P_{ij}^m > 0 \quad \text{and} \quad P_{jj}^{n+s+m} \geq P_{ji}^n \cdot P_{ii}^s \cdot P_{ij}^m > 0.
  \]

- The second inequality follows, for instance, since the left-hand side represents the probability that starting in \( j \) the chain will be back in \( j \) after \( n + s + m \) transitions, whereas the right-hand side is the probability of the same event subject to the further restriction that the chain is in \( i \) both after \( n \) and \( n + s \) transitions.

- Hence, \( d(j) \) divides both \( n + m \) and \( n + s + m \); thus \( n + s + m - (n + m) = s \), whenever \( P_{ii}^s > 0 \). Therefore, \( d(j) \) divides \( d(i) \). A similar argument yields that \( d(i) \) divides \( d(j) \), thus \( d(i) = d(j) \).
Classification of States (con’t)

- For any states $i$ and $j$, define $f_{i,j}^n$ to be the probability that, starting in $i$, the first transition into $j$ occurs at time $n$.

$$f_{i,j}^0 = 0,$$

$$f_{i,j}^n = P\{X_n = j, X_k \neq j, k = 1, \ldots, n-1 | X_0 = i\}.$$  

$$f_{i,j}^{(n)} = Pr \text{[first passage time } i \rightarrow j \text{ is } n]$$

- Let

$$f_{i,j} = \sum_{n=1}^{\infty} f_{i,j}^n.$$  

$f_{i,j}$ denotes the probability of ever making a transition into state $j$, given that the process starts in $i$.

- For $i \neq j$, $f_{i,j}$ is positive if, and only if, $j$ is accessible from $i$.

- State $j$ is **recurrent** if $f_{j,j} = 1$, and **transient** if $f_{j,j} < 1$,

$$\left( \text{state } i \text{ is recurrent } \Leftrightarrow \sum_{n=1}^{\infty} f_{i,j}^{(n)} = 1 \right) \quad \left( \text{state } i \text{ is transient } \Leftrightarrow \sum_{n=1}^{\infty} f_{i,j}^{(n)} < 1 \right)$$
Classification of States (con’t)

- If a state is recurrent, it can be recurrent nonnull or recurrent null.

\[
\text{(recurrent state } i \text{ is recurrent null } \iff \sum_{n=1}^{\infty} (n f_{ij}^{(n)}) = \infty) \\
\text{(recurrent state } i \text{ is recurrent nonnull } \iff \sum_{n=1}^{\infty} (n f_{ij}^{(n)}) < \infty)
\]

Example:

\[
f_{00}^{(1)} = 0 \quad f_{00}^{(2)} = \frac{1}{2} \\
\frac{f_{00}^{(3)}}{2} = \frac{1}{3} = \frac{1}{6} \quad \frac{f_{00}^{(4)}}{2} = \frac{1}{3} = \frac{1}{12}
\]
The probability of eventually returning to state 0 is

$$\sum_{n=1}^{\infty} f_{00}^{(n)} = \sum_{n=2}^{\infty} \left( \frac{1}{n-1} \cdot \frac{1}{n} \right) = \sum_{n=1}^{\infty} \left( \frac{1}{n} \cdot \frac{1}{n+1} \right)$$

$$\sum_{n=1}^{\infty} \left( \frac{1}{n} \cdot \frac{1}{n+1} \right) = \frac{k - 1}{k}$$

$$\sum_{n=1}^{\infty} f_{00}^{(n)} = \lim_{k \to \infty} \frac{k - 1}{k} = 1$$

Thus, state 0 is recurrent.

State 0 is recurrent nonnull or recurrent null?

$$\sum_{n=1}^{\infty} \left( n \cdot f_{00}^{(n)} \right) = \sum_{n=2}^{\infty} \left( n \cdot \frac{1}{n-1} \cdot \frac{1}{n} \right) = \sum_{n=1}^{\infty} \left( \frac{1}{n} \right) = \infty$$

The state must be recurrent null.
Classification of States (con’t)

Proposition.
State \( j \) is **recurrent** if, and only if,
\[
\sum_{n=1}^{\infty} P_{jj}^n = \infty
\]

Proof.
- State \( j \) is recurrent if, with probability 1, a process starting at \( j \) will eventually return.
- However, by the Markovian property it follows that the process probabilistically restarts itself upon returning to \( j \). Hence, with probability 1, it will return again to \( j \).
- Repeating this argument, we see that, with probability 1, the number of visits to \( j \) will be infinite and will thus have infinite expectation.
- On the other hand, suppose \( j \) is transient. Then each time the
Proof (con’t)

process returns to \( j \) there is a positive probability \( 1 - f_{jj} \) that it will never again return; hence the number of visits is geometric with finite mean \( 1/(1 - f_{jj}) \).

- By the above argument we see that state \( j \) is recurrent if, and only if,

\[
E[\text{number of visits to } j | X_0 = j] = \infty
\]

- But, letting \( I_n = \begin{cases} 1 & \text{if } X_n = j \\ 0 & \text{otherwise} \end{cases} \)

it follows that \( \sum_0^\infty I_n \) denotes the number of visits to \( j \). Since

\[
E \left[ \sum_{n=0}^\infty I_n | X_0 = j \right] = \sum_{n=0}^\infty E[I_n | X_0 = j] = \sum_{n=0}^\infty P_{jj}^n;
\]

the result follows.
Classification of States (con’t)

• The proposition also shows that a \textit{transient} state will only be visited a finite number of times (hence the name transient).

• This leads to the conclusion that in a \textit{finite – state} Markov chain not all states can be transient.
  
  – To see this, suppose the states are 0, 1, ..., \( M \) and suppose that they are all transient. Then after a finite amount of time (say after time \( T_0 \)) state 0 will never be visited, and after a time (say \( T_1 \)) state 1 will never be visited, and after a time (say \( T_2 \)) state 2 will never be visited, and so on.
  
  – Thus, after a finite time \( T = \max\{T_0, T_1, ..., T_M\} \) no states will be visited. But as the process must be in some state after time \( T \), we arrive at a contradiction, which shows that at least one of the states must be recurrent.

\[ \Rightarrow \text{A finite-state Markov must have at least one recurrent state.} \]
Examples of Recurrent and Transient States

Example:

\[ P = \begin{bmatrix}
0 & 0 & 1/2 & 1/2 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0
\end{bmatrix} \]

1. The chain is finite  \( \Rightarrow \) At least one recurrent state.
2. All states communicate \( \Rightarrow \) All states are recurrent!

Example:

\[ P = \begin{bmatrix}
1/2 & 1/2 & 0 & 0 & 0 \\
1/2 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & 0 \\
0 & 0 & 1/2 & 1/2 & 0 \\
1/4 & 1/4 & 0 & 0 & 1/2
\end{bmatrix} \]

Communicating classes: \{0,1\} \{2,3\} \{4\}

\[ f_{0,0} = \sum_{1}^{\infty} f_{0,0}^{n} = P\{X_1 = 0|X_0 = 0\} + P\{X_2 = 0|X_1 = 1, X_0 = 0\} + \ldots + P\{X_i = 0|X_j = 1, j = 1, \ldots, i-1, X_0 = 0\} + \ldots \]

\[ = \sum_{k=1}^{\infty} (1/2)^{k} = 1 \Rightarrow \text{States 0 and 1 are recurrent! Similarly, States 2 and 3 are recurrent!} \]

\[ f_{4,4} = \sum_{1}^{\infty} f_{4,4}^{n} = P\{X_1 = 4|X_0 = 4\} = 1/2 \leq 1 \Rightarrow \text{State 4 is transient!} \]
Corollary. (*Recurrent* is a class property)
If \( i \) is recurrent and \( i \leftrightarrow j \), then \( j \) is recurrent.

**Proof.**
Let \( m \) and \( n \) be such that \( P_{ij}^n > 0, P_{ji}^m > 0 \). Now for any \( s \geq 0 \)

\[
P_{jj}^{m+n+s} \geq P_{ji}^m P_{ii}^s P_{ij}^n
\]

and thus

\[
\sum_s P_{jj}^{m+n+s} \geq P_{ji}^m P_{ij}^n \sum_s P_{ii}^s = \infty
\]

and the result follows from the above proposition.
Example. The Simple Random Walk

- The Markov chain whose state space is the set of all integers and has transition probabilities

\[ P_{i,i+1} = p = 1 - P_{i,i-1}, \quad i = 0, \pm 1, \ldots, \]

where \( 0 < p < 1 \), is called the simple random walk.

- One interpretation of this process is that it represents the winnings of a gambler who on each play of the game either wins or loses one dollar.

- Since all states communicate, they are either all transient or all recurrent.

  - Consider state 0 and attempt to determine if \( \sum_{n=1}^{\infty} P_{00}^n \) is finite or infinite.

    If \( \sum_{n=1}^{\infty} P_{00}^n = \infty \), all states are recurrent.
Example. The Simple Random Walk (con’t)

- Since it is impossible to be even (using the gambling model interpretation) after an odd number of plays, we must have

\[ P^{2n+1}_{00} = 0, \quad n = 1, 2, \ldots \]

- On the other hand, the gambler would be even after \(2n\) trials if, and only if, he won \(n\) of these and lost \(n\) of these.

\[ \text{The desired probability is binomial with probability } p, \text{ i.e.,} \]

\[ P^{2n}_{00} = \binom{2n}{n} p^n (1 - p)^n \frac{(2n)!}{n!n!} (p(1 - p))^n, \quad n = 1, 2, 3, \ldots \]

Applying Stirling formula: \(n! \sim n^{n+1/2} e^{-n} \sqrt{2\pi}\)

\[ P^{2n}_{00} = \frac{(2n)!}{n!n!} (p(1 - p))^n = \frac{(2n)^{2n+1/2} e^{-2n} \sqrt{2\pi}}{(n^{n+1/2} e^{-n} \sqrt{2\pi})^2} p^n (1 - p)^n \sim \frac{(4p(1 - p))^n}{\sqrt{\pi n}} \]

\[ \Rightarrow \sum_{n=1}^{\infty} P^n_{00} = \sum_{n=1}^{\infty} P^{2n}_{00} \sim \sum_{n=1}^{\infty} \frac{(4p(1 - p))^n}{\sqrt{\pi n}} \]
Example. The Simple Random Walk (con’t)

When will \( \sum_{n=1}^{\infty} \frac{(4p(1-p))^n}{\sqrt{\pi n}} \) converge?

Applying the ratio test with \( a_i = \sum_{n=1}^{\infty} \frac{(4p(1-p))^n}{\sqrt{\pi n}} \) such that

\[
\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \lim_{n \to \infty} \frac{(4p(1-p))^{n+1}/\sqrt{\pi (n+1)}}{(4p(1-p))^n/\sqrt{\pi n}} = 4p(1-p) \lim_{n \to \infty} \frac{\sqrt{n}}{\sqrt{n+1}} = 4p(1-p)
\]

If \( p \neq 1/2, 0 \leq p \leq 1, 4p(1-p) < 1 \)
\[\Rightarrow \sum_{n=1}^{\infty} P_{00}^n \sim \sum_{n=1}^{\infty} \frac{(4p(1-p))^n}{\sqrt{\pi n}} \text{ converges to a fixed value } < \infty \]
\[\Rightarrow \text{All states are transient!} \]

If \( p = 1/2, 4p(1-p) = 1 \)
\[\Rightarrow \sum_{n=1}^{\infty} P_{00}^n \sim \sum_{n=1}^{\infty} \frac{(4p(1-p))^n}{\sqrt{\pi n}} = \infty \]
\[\Rightarrow \text{All states are recurrent!} \]

\[\Rightarrow \text{All states are recurrent if } p = 1/2 \]
\[\text{and all states are transient if } p \neq 1/2 \]
Example. The Simple Random Walk (con’t)

Remark.

• When \( p = \frac{1}{2} \), the above process is called a *symmetric random walk*. We could also look at symmetric random walks in more than one dimension.

• For instance, in the two-dimensional symmetric random walk the process would, at each transition, either take one step to the left, right, up, or down, each having probability \( p = \frac{1}{4} \).

• Similarly, in three dimensions the process would, with probability \( \frac{1}{6} \) make a transition to any of the six adjacent points.

• By using the same method as in the one-dimensional random walk it can be shown that the two-dimensional symmetric random walk is recurrent, but all higher-dimensional random walks are transient.
Classification of States (con’t)

Corollary.
If \( i \leftrightarrow j \) and \( j \) is recurrent, then \( f_{ij} = 1 \).

Proof.
- Suppose \( X_0 = i \), and let \( n \) be such that \( P_{ij}^n > 0 \).
- Say that we miss opportunity 1 if \( X_n \neq j \). If we miss opportunity 1, then let \( T_1 \) denote the next time we enter \( i \) (\( T_1 \) is finite with probability 1 by the previous corollary).
- Say that we miss opportunity 2 if \( X_{T_1+n} \neq j \). If opportunity 2 is missed, let \( T_2 \) denote the next time we enter \( i \) and say that we miss opportunity 3 if \( X_{T_2+n} \neq j \) and so on.
- It is easy to see that the opportunity number of the first success is a geometric random variable with mean \( 1/P_{ij}^n \), and is thus finite with probability 1. The result follows since \( i \) being recurrent implies that the number of potential opportunities is infinite.
Classification of States (con’t)

Remark.

- Let $N_j(t)$ denote 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 $j$, it follows that $\{N_j(t), t \geq 0\}$ is a renewal process with interarrival distribution $\{f_{jj}^n, n \geq 1\}$.
- If $X_0 = i, i \leftrightarrow j$, and $j$ is recurrent, then $\{N_j(t), t \geq 0\}$ is a delayed renewal process with initial interarrival distribution $\{f_{ij}^n, n \geq 1\}$.
Limit Theorems

- It is easy to show that if state $j$ is transient, then

$$
\sum_{n=1}^{\infty} P_{ij}^n < \infty \quad \text{for all } i,
$$

meaning that, starting in $i$, the expected number of transitions into state $j$ is finite. As a consequence it follows that for $j$ transient $P_{ij}^n \to 0$ as $n \to \infty$.

- Let $\mu_{jj}$ denote the expected number of transitions needed to return to state $j$. That is,

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

- By interpreting transitions into state $j$ as being renewals, we obtain the following theorem from Chapter 3.

**Theorem.**

If $i$ and $j$ communicate, then:

1. $P \left\{ \lim_{t \to \infty} \frac{N_j(t)}{t} = 1/\mu_{jj} \mid X_0 = i \right\} = 1$ \hspace{1cm} \text{Strong Law of Renewal Process}

2. $\lim_{n \to \infty} \sum_{k=1}^{n} \frac{P_{ij}^k}{n} = 1/\mu_{jj}$ \hspace{1cm} \text{Elementary Renewal Theorem}

3. If $j$ is aperiodic, then $\lim_{n \to \infty} P_{ij}^n = 1/\mu_{jj}$ \hspace{1cm} \text{Blackwell Theorem}

4. If $j$ has period $d$, then $\lim_{n \to \infty} P_{ij}^{nd} = d/\mu_{jj}$
Limit Theorems

- If state $j$ is recurrent, then it is positive recurrent if $\mu_{jj} < \infty$ and null recurrent if $\mu_{jj} = \infty$.

- Let

$$\pi_j = \lim_{n \to \infty} P_{jj}^{nd(j)},$$

stationary probability, limiting probabilities, equilibrium probabilities, or balance probabilities

A recurrent state $j$ is positive recurrent if $\pi_j > 0$ and null recurrent if $\pi_j = 0$.

**Proposition.**
Positive (null) recurrence is a class property.

- A positive recurrent, aperiodic state is called **ergodic**.

- Before presenting a theorem that shows how to obtain the limiting probabilities in the ergodic case, we need the following definition.
Limit Theorems

Definition.
A probability distribution $\{P_j, j \geq 0\}$ is said to be *stationary* for the Markov chain if

$$P_j = \sum_{i=0}^{\infty} P_i P_{ij}, \quad j \geq 0.$$ 

- If the probability distribution of $X_0$, i.e., $P_j = \{X_0 = j\}, j \geq 0$, is a stationary distribution, then

$$P\{X_1 = j\} = \sum_{i=0}^{\infty} P\{X_1 = j | X_0 = i\} P\{X_0 = i\}$$

$$= \sum_{i=0}^{\infty} P_i P_{ij} = P_j$$
Limit Theorems

and, by induction,

\[ P\{X_n = j\} = \sum_{i=0}^{\infty} P\{X_n = j | X_{n-1} = i\} P\{X_{n-1} = i\} \]

\[ = \sum_{i=0}^{\infty} P_{ij} P_i = P_j. \]

- Hence, if the initial probability distribution is the stationary distribution, then \(X_n\) will have the same distribution for all \(n\).
- **Stationary assumption:** \(P\{X_{n+1} = j | X_n = i\}\) is independent of \(n\).
- **Stationary transition probability:** \(P_{ij} = P\{X_{n+1} = j | X_n = i\} = P\{X_1 = j | X_0 = i\}\).
- In fact, as \(\{X_n, n \geq 0\}\) is a Markov chain, it easily follows from this that for each \(m \geq 0, X_n, X_{n+1}, \ldots, X_{n+m}\) will have the same joint distribution for each \(n\); in other words, \(\{X_n, n \geq 0\}\) will be a stationary process.
Limit Theorems

Theorem.
An irreducible aperiodic Markov chain belongs to one of the following two cases:

1. All states are transient (or null recurrent);
in this case, $P^n_{ij} \rightarrow 0$ as $n \rightarrow \infty$ for all $i, j$ and there exists
no stationary distribution.

Note that for a finite Markov chain, there must be at least one recurrent state

=> There are infinite transient states!

2. All states are positive recurrent, that is,

$$\pi_j = \lim_{n \rightarrow \infty} P^n_{ij} > 0$$

In this case, $\{\pi_j, j = 0, 1, 2, \ldots\}$ is a stationary distribution
and there exists no other stationary distribution.
Proof.

We will first prove 2. To begin, note that

$$P_{ij}^{n+1} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj} \geq \sum_{k=0}^{M} P_{ik}^n P_{kj} \quad \text{for all } M.$$  

Letting $n \to \infty$ yields

$$\pi_j \geq \sum_{k=0}^{M} \pi_k P_{kj} \quad \text{for all } M,$$

implying that

$$\pi_j \geq \sum_{k=0}^{\infty} \pi_k P_{kj} \quad j \geq 0. \quad \text{(1)}$$

To show that equation (1) is actually an equality, suppose that the inequality is strict for some $j$. Then upon adding these inequalities
Proof. (con’t)

we obtain

\[ \sum_{j=0}^{\infty} \pi_j > \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \pi_k P_{k,j} = \sum_{k=0}^{\infty} \pi_k \sum_{j=0}^{\infty} P_{k,j} = \sum_{k=0}^{\infty} \pi_k, \]

Contradiction!!

Therefore,

\[ \pi_j = \sum_{k=0}^{\infty} \pi_k P_{k,j}, \quad j = 0, 1, 2, \ldots \]

balance equations

Putting \( P_j = \pi_j / \sum_{0}^{\infty} \pi_k \), we see that \( \{P_j, j = 0, 1, 2, \ldots\} \) is a stationary distribution, and hence at least one stationary distribution exists.

Next, we show that there is only one stationary distribution.