Lecture 20 : Markov Chains

We consider stochastic processes. A process represents a system that evolves through incremental changes called transitions. When such changes occur randomly then a process is stochastic.

There are two components of a stochastic process: one is time and the other is a set of states that a system can be in at a point in time. In a simple categorization, both time and a space of states can be either discrete or continuous, which gives four possible cases. The use of “chain” rather than “process” usually indicates discreteness, typically discrete time.

We consider discrete-time processes, unless stated otherwise, which means that state transitions occur in discrete points in time. We refer to time as determined by a sequence of steps $0, 1, 2, 3, \ldots$

Regarding states, we consider discrete states, unless stated otherwise. A discrete set of states can be either finite or countably infinite. We use natural numbers $0, 1, 2, 3, \ldots$ as names for states, to simplify the notation.

20.1 Definitions and notations

The evolution of a discrete-time stochastic process is represented as a countably infinite sequence of random variables $X = (X_0, X_1, X_2, \ldots)$ taking states as values. In such a context, the following terminology is used.

If $X_0 = i$ then we say that $i$ is the initial state of process $X$.

If $X_n = j$ then process $X$ is in state $j$ in step $n$, for $n > 0$.

If $X_{n-1} = k$ and $X_n = \ell$ then process $X$ enters state $\ell$ from state $k$ in step $n$, for $n > 0$.

Let us consider two examples of a process $X = (X_0, X_1, X_2, \ldots)$ where $X_i \in \{0, 1\}$. One is when the random variables $X_0, X_1, X_2, \ldots$ are independent and determined by tosses of a coin. Another is when the random variables $X_0, X_1, X_2, \ldots$ are determined by time completely, for instance, $X_{2i} = 0$ and $X_{2i+1} = 1$, for all natural numbers $i$.

Markov chains model stochastic processes with limited dependence, which are somewhat between the two examples above. Such a limited dependence is achieved by restricting how past events affect transitions. This means how $X_n$ is related to the values of $X_0, X_1, \ldots, X_{n-1}$. The examples above represent two extreme cases: one is with a total dependence (the pattern $0, 1, 0, 1, 0, 1, \ldots$ is completely deterministic) and the other has no dependence (stochastic independence of coin tosses). Markov chains are about a situation when $X_n$ depends on $X_{n-1}$ only, with $X_0$ determined by some probability distribution. Next we give formal definitions.
A discrete-time discrete-state stochastic process \( X \) is a Markov chain when the probability of entering a state \( j \) in step \( n \) depends only on a state that \( X \) is in step \( n - 1 \). In notation, we want the following Markov identity to hold, for any time step \( n > 0 \) and for any states \( j \) and \( k \), and for all possible trajectories of the process to reach state \( k \) in step \( n - 1 \):

\[
\Pr(X_n = j \mid X_{n-1} = k, X_{n-2} = i, \ldots, X_0 = \ell) = \Pr(X_n = j \mid X_{n-1} = k).
\]

This identity means that once the process is in a state \( k \) then the future evolution of the process does not depend on how it arrived into this state \( k \).

Markov chains may have the property that the probabilities \( \Pr(X_n = j \mid X_{n-1} = k) \) are equal for all \( n > 0 \) and depend only on the states \( j \) and \( k \). This is called time homogeneity because index \( n \) in \( X_n \) represents time. We will consider only such time-homogeneous Markov chains and drop the “homogeneous” part of the term. Observe that the property of a Markov chain to be time homogeneous implies that the numbers \( \Pr(X_1 = j \mid X_0 = k) \) uniquely determine the probabilities of transitions among states in all steps.

For a given Markov chain, we will use the following notations for probabilities:

1. \( p_k(n) = \Pr(X_n = k) \) is the probability to be in state \( k \) in step \( n \). The distribution of \( p_k(0) \) is called initial.

2. \( p_{kj} = \Pr(X_n = j \mid X_{n-1} = k) \), for all \( n > 0 \), is the probability of entering state \( j \) directly from state \( k \). This is about one-step transitions and these numbers are transition probabilities.

3. \( p_{kj}(n) = \Pr(X_{m+n} = j \mid X_m = k) \), for all steps \( m \geq 0 \) and \( n \geq 0 \), is the probability of entering state \( j \) after \( n \) steps when starting from state \( k \). This is about \( n \)-step transitions. In particular, \( p_{ii}(0) = 1 \) and \( p_{ij}(0) = 0 \) for \( i \neq j \).

A matrix \( P = [p_{kj}] \) is called the transition matrix of \( X \). A row \( k \) consists of transition probabilities for the chain to move from state \( k \) to all other available states. It follows that a transition matrix has all non-negative entries that sum up to 1 in each row; such matrices are called stochastic. The matrix \( P_n = [p_{ij}(n)] \) is called \( n \)-step transition matrix of \( X \); it is also a stochastic matrix.

The following Chapman-Kolmogorov equations hold, for \( s, u \geq 0 \):

\[
p_{ij}(s + u) = \sum_k p_{ik}(s) \cdot p_{kj}(u) .
\]

They follow directly from the formula of total probability:

\[
\Pr(X_{s+u} = j \mid X_0 = i) = \sum_k \Pr(X_{s+u} = j \mid X_s = k) \cdot \Pr(X_s = k \mid X_0 = i) .
\]
In matrix notation, Chapman-Kolmogorov equations mean that \( P_{s+u} = P_s \cdot P_u \), in particular the \( n \)-step transition matrix is of the following form \( P_n = P^n \).

A probability distribution of a random variable \( Y \) taking on natural numbers as values can be represented as a sequence \( \left( \Pr(Y = i) \right)_{i \geq 0} \). If \( a = (a_1, a_2, \ldots) \) is an initial probability distribution of a Markov chain \( X \), which is the probability distribution of \( X_0 \) in the sense \( a_i = \Pr(X_0 = i) \), then the distribution of \( X_n \) is determined as the sequence \( a \cdot P^n \). This formula \( a \cdot P^n \) means that we treat \( a \) as a “horizontal” vector and we multiply it, using the dot-product, by each column, each such a multiplication providing a term of the outcome. The outcome is a probability distribution on the set of states and so is also a “horizontal” vector like the distribution \( a \).

For a given Markov chain, we will use the following notations for probabilities:

1. \( f_{ij}(n) \) is the probability of starting from state \( i \) and entering state \( j \) for the first time in step \( n \), where \( n > 0 \). In particular, when \( i = j \) then \( f_{ii}(n) \) denotes the probability of starting at state \( i \) and returning again to \( i \) for the first time after \( n \) steps and making at least one step. We extend this notation to stipulate that \( f_{ij}(0) = 0 \) for any pair of states \( i \) and \( j \), including the case \( i = j \).

2. \( f_{ij} = \sum_{n>0} f_{ij}(n) \) is the probability of the event that state \( j \) is eventually reached when starting from state \( i \) after at least one transition. In particular, the number \( f_{ii} = \sum_{n>0} f_{ii}(n) \) is the probability of return to state \( i \) when starting from \( i \).

The following equation is analogous to the one we considered for general recurrent events:

\[
p_{ik}(n) = \sum_{j=1}^{n} f_{ik}(j)p_{kk}(n-j) , \tag{1}
\]

for \( n > 0 \). It follows that the sequence \( (p_{ik}(n))_{n \geq 0} \), for any states \( i \) and \( k \), is “almost” a convolution of the sequences \( (f_{ik}(n))_{n \geq 0} \) and \( (p_{kk}(n))_{n \geq 0} \), but not exactly. Observe that the following identity holds when either \( n > 0 \) or \( i \neq k \):

\[
p_{ik}(n) = \sum_{j=0}^{n} f_{ik}(j)p_{kk}(n-j) , \tag{2}
\]

but (2) does not hold for \( n = 0 \) and \( i = k \), because \( p_{ii}(0) = 1 \) and \( f_{ii}(0) = 0 \). This can be represented concisely by generating functions. Let us consider the generating functions

\[ F_{ik}(z) = \sum_{n \geq 0} f_{ik}(n)z^n \quad \text{and} \quad P_{ik}(z) = \sum_{n \geq 0} p_{ik}(n)z^n. \]

The equations (1) and (2) yield the following equation involving these generating functions

\[
P_{ik}(z) = [i = k] + F_{ik}(z)P_{kk}(z) ,
\]

where \([i = k]\) is the Iverson’s bracket equal 1 when \( i = k \) and 0 otherwise.
20.2 A worked example

A Markov chain can be represented as a directed graph with labels attached to nodes and edges. Nodes represent states and edges direct transitions between states. We consider a transition worth representing when its probability is positive, and this probability is attached as a label to the respective edge.

Figure 1 is an example of such a graphic representation of a Markov chain with a finite set of states. Such a representation is called a transition graph.

An evolution of a Markov chain proceeds in such a manner that when the chain is in a state in a step then it will move to a new state in the next step traversing each outgoing edge with the probability equal to the label on that edge. It follows that the sum of all probabilities on outgoing edges, for any node, needs to be precisely 1. In Figure 1, we can see that this implies that $r + s + t = 1$, by examining states 1 and 2.

An advantage of a representation by way of a transition graph is that certain state properties are visible at a glance. For instance, state labeled by 3 has the property that when the chain enters this state then it cannot leave it; such states are rightly called absorbing.

Consider a Markov chain with a transition matrix $Q$ given in Figure 2. It specifies the same Markov chain as represented by a directed labeled graph in Figure 1. More precisely, the rows in $Q$, ordered from top to bottom, and the columns in $Q$, ordered from left to right, correspond to the states in Figure 2, ordered from left to right. The entries of $Q$ are interpreted as probabilities. The property of a transition graph that the sum of probabilities on outgoing edges from a node is exactly 1 has the matching property that the sum of entries in a row of a transition matrix is 1.

As another approach to the same Markov chain, it can be described by a story, which is as good as a graph or a matrix. Let us consider a particle wandering randomly among four points 0, 1, 2, 3 arranged on a line, such that the particle either moves from a node to one of its direct neighbors among the nodes 0, 1, 2, 3 or stays put in a node. Furthermore, assume...
A worked example

\[
Q = \begin{bmatrix}
0 & 1 & 0 & 0 \\
t & s & r & 0 \\
0 & t & s & r \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

Figure 2: A Markov chain with four states represented as a $4 \times 4$ transition matrix $Q$. All entries in $Q$ are nonnegative real numbers. The sum of entries in each row of $Q$ is 1.

that point 0 is a reflecting barrier and point 3 is an absorbing barrier, these terms are self-explanatory. When visiting a point 1 or 2, the particle moves to the right with probability $r$ and to the left with probability $t$, and stays put with probability $s$, where $r + s + t = 1$. It is the same finite Markov chain with four states 0, 1, 2, 3 represented by a transition graph in Figure 1 and by its transition matrix in Figure 2.

For this Markov chain, let us compute the probability $f_{1,0}$. Let us consider the first movement of the particle from state 1. By the formula for total probability, we obtain the following equation:

\[ f_{1,0} = t \cdot 1 + s \cdot f_{1,0} + r \cdot f_{2,0} . \]  

Similarly, the probability $f_{2,0}$ satisfies the following equation:

\[ f_{2,0} = t \cdot f_{1,0} + s \cdot f_{2,0} + r \cdot 0 . \]

We factor out $f_{2,0}$ to obtain $f_{2,0}(1 - s) = t \cdot f_{1,0}$. Substituting this to (3) and solving for $f_{1,0}$ yields

\[ f_{1,0} = t \cdot \left( r + t - \frac{rt}{r + t} \right)^{-1} . \]

Similarly, we can find probability $f_{ij}$ for this Markov chain for any states $i$ and $j$. In particular we find that $f_{i,3} = 1$ for $i = 0, 1, 2, 3$, which means that the event that the particle is eventually absorbed by state 3 is a sure one; see Exercise 2.

To continue with this example, suppose that the particle starts in state $X_0$, where random variable $X_0$ has some distribution $a = (a_0, a_1, a_2, a_3)$, in the sense that

\[ a_0 = \Pr(X_0 = 0), \quad a_1 = \Pr(X_0 = 1), \quad a_2 = \Pr(X_0 = 2), \quad a_3 = \Pr(X_0 = 3) . \]

If the particle starts at state $i$ then it will be absorbed by state 3 after the expected time $m_i = \sum_{n>0} n \cdot f_{i,3}(n)$. The initial position is given by $a$, so that the expected time of absorption is $m = \sum_{0\leq i \leq 3} a_i \cdot m_i$. Observe that when the particle starts at state 3 then it takes one step for it to be absorbed by this very state 3, according to $f_{3,3}(1) = 1$, so $m_3 = 1$. 
We may find \( m \) without summing the series \( \sum_{n \geq 0} n f_{i,3}(n) \). Namely, the following system of equations

\[
\begin{align*}
m_0 &= m_1 + 1 \\
m_1 &= tm_0 + sm_1 + rm_2 + 1 \\
m_2 &= tm_1 + sm_2 + rm_3 + 1 \\
m_3 &= 1
\end{align*}
\]
determines \( m_i \) and makes the problem purely algebraic. Intuitively, the fact that \( 1 \) occurs in each line of these equations follows from the fact that we consider what happens after one step and we need to account for this one step by adding \( 1 \) to the expectation. The general shape of each equation follows from the formula on total probability. As an example, here is a formal derivation of the equation \( m_1 = tm_0 + sm_1 + rm_2 + 1 \). By the formula for total probability, the following equations hold for every \( n \geq 0 \):

\[
f_{1,3}(n + 1) = tf_{0,3}(n) + sf_{1,3}(n) + rf_{2,3}(n) .
\]

We sum these equations over all \( n \geq 0 \) to obtain

\[
t \cdot m_0 + s \cdot m_1 + r \cdot m_2 = \sum_{n \geq 0} nf_{1,3}(n + 1) = (n + 1)f_{1,3}(n + 1) - \sum_{n \geq 0} f_{1,3}(n + 1) = m_1 - f_{1,3} = m_1 - 1 .
\]

The other equations can be justified similarly.

\section{20.3 Properties of states and chains}

If a discrete-time stochastic process starts at a state \( i \), that is \( X_0 = i \), then a return to \( i \) can be interpreted as a recurrent event. The properties of such an event, like being persistent or transient, can be associated with this state \( i \) to simplify the terminology. We will refer to states as persistent or transient accordingly. Similarly, rather than the period of the event to return to \( i \) we may say the period of state \( i \).

A state \( i \) is persistent when \( f_{ii} = 1 \) and it is transient otherwise. It follows from general properties of recurrent events that a state \( i \) is persistent when \( \sum_n p_{ii}(n) = \infty \) and it is transient when \( \sum_n p_{ii}(n) < \infty \). Observe that if \( i \) is transient then \( \lim_{n \to \infty} p_{ii}(n) = 0 \).

A state \( j \) is accessible from state \( i \) when \( f_{ij} > 0 \). States \( i \) and \( j \) that are mutually accessible, which holds when \( f_{ij} \cdot f_{ji} > 0 \), are said to communicate. The relation of communicating is an equivalence one. It classes of abstraction are called communicating classes.
Consider two states $i$ and $j$ that communicate. Let $s$ and $u$ be numbers such that $p_{ij}(s) > 0$ and $p_{ji}(u) > 0$. By the Chapman-Kolmogorov’s equations, it follows that for each integer $n \geq 0$ we have $p_{ii}(s + n + u) \geq p_{ij}(s) \cdot p_{jj}(n) \cdot p_{ji}(u)$. So the following inequality holds
\[
p_{ii}(n + s + u) \geq c \cdot p_{jj}(n) ,
\]
for some $c > 0$. Similarly
\[
p_{jj}(n + s + u) \geq c \cdot p_{ii}(n) .
\]
These inequalities imply the following equivalences:
\[
\lim_{n \to \infty} p_{jj}(n) = 0 \quad \text{if and only if} \quad \lim_{n \to \infty} p_{ii}(n) = 0 ;
\]
\[
\sum_{n \geq 0} p_{jj}(n) = \infty \quad \text{if and only if} \quad \sum_{n \geq 0} p_{ii}(n) = \infty .
\]
It follows from (7) that if two states communicate, then either both are transient or both are persistent. Similarly, they have the same period. Namely, let $t_i$ and $t_j$ be the periods of states $i$ and $j$, respectively. (Let $a \mid b$ denote that integer $a$ divides integer $b$.) By (4) and (5) it follows that $t_i \mid (s + u)$ and $t_j \mid (s + u)$, and therefore $t_i \mid t_j$ and $t_j \mid t_i$, so that $t_i = t_j$.

A set of states $C$ is closed when the probability of leaving this set is zero. This means in notation that $p_{ij} = 0$ for each pair of states $i \in C$ and $j \not\in C$. A set of states $C$ is irreducible if $C$ is closed and $C$ has no proper closed subset.

Let us revisit the Markov chain in Figure 1. All states are accessible from states 0, 1, and 2, but only 3 is accessible from 3. There are two communicating classes $\{0, 1, 2\}$ and $\{3\}$. The set of all states is closed and the singleton set $\{3\}$ is closed. It follows that $\{3\}$ is the only irreducible set of states.

The set of states of a Markov chain can be partitioned uniquely into disjoint sets of states $T \cup C_1 \cup C_2 \cup \ldots$, where $T$ is the set of transient states, possibly empty, and each $C_i$ is an irreducible set of persistent states, where $C_1 \cup C_2 \cup \ldots$ is possibly empty. Indeed, suppose there are persistent states and let $C_1 \cup C_2 \cup \ldots$ be the partitioning of these states into communication classes. Suppose there are some two states $i$ and $j$ from different classes in $C_1 \cup C_2 \cup \ldots$ for which $p_{ij} > 0$. Then $f_{ji} = 0$ because the classes are different. Observe that
\[
f_{ii} = \sum_k p_{ik} f_{ki} \leq \sum_{k \neq j} p_{ik} < 1 ,
\]
due to $p_{ij} > 0$. So the state $i$ is not persistent, which is a contradiction.

This fact can be used to explain how a Markov chain evolves with time. If the chain starts in a transient state then either it stays forever among such states or eventually enters a persistent state and then stays forever in the respective irreducible set of states.
When a chain is in an irreducible set of states then this set of states determines a Markov chain with the inherited transition probabilities. This motivates considering Markov chains in which all states form an irreducible set of states; such Markov chains are called irreducible.

A persistent state is null when \( \lim_{n \to \infty} p_{kk}(n) = 0 \), otherwise it is non-null or positive. When two persistent states communicate then either both are null or both are positive, which follows from (6). In an irreducible Markov chain either all states are transient or all are null persistent or all are positive persistent. Let us revisit the Markov chain in Figure 1: the states 0, 1, 2 are transient and 3 is positive persistent.

We apply the convention to call the whole irreducible chain by the property that all its state have. Hence an irreducible Markov chain is called transient or null persistent or positive persistent when all its states are such, respectively.

Let us observe that a finite irreducible Markov chain has to contain a positive persistent state and so the chain is positive persistent. Indeed, suppose otherwise that either all the states are transients or all are null persistent. Then \( \lim_{n \to \infty} p_{ij}(n) = 0 \), for each \( i \) and \( j \), see Exercise 10. Matrix \( P^n \) has a finite number of columns, so a row cannot have the sum of terms equal to 1 and converge pointwise to 0.

The mean time of return for a state \( k \) is defined as \( \mu_k = \sum_{n \geq 0} n f_{kk}(n) \), when \( k \) is persistent. If \( k \) is transient then with a positive probability the state never returns to \( k \), which can be interpreted as an infinite time of return with positive probability. For such a state \( k \) we stipulate that \( \mu_k = \infty \). For persistent states \( k \), the probability of return is 1 but the mean time of return \( \mu_k \) can be either finite or infinite. The categorization of persistent states into positive and null reflects the mean time of return as follows: a persistent state is positive if and only if its mean time of return is finite. We will show this fact later.

\section{20.4 Random walks in discrete Euclidean space}

Consider a random walk of a particle that wanders between two reflecting barriers: one at point 0 and the other at a point of integer coordinate \( r > 0 \). Let \( p > 0 \) be the probability that the particle moves to the right when at a point \( x < r \) and \( q = 1 - p > 0 \) be the probability of moving to the left when the particle is at a point \( x > 0 \). This is a finite irreducible Markov chain with the coordinates of points with integer coordinates between and including the barriers as states. This is a positive persistent Markov chain.

When we change the behavior of the right barrier from reflecting to absorbing then what we obtain is a Markov chain similar to that depicted in Figure 1. Such a chain is no longer irreducible: all states except for the absorbing one are transient and the absorbing state makes a singleton irreducible set.

Let us change this chain by removing the right barrier altogether so this is a random walk on a positive-side of a discrete line with a reflecting barrier at point 0. Again, \( p \) is the
20.4  Random walks in discrete Euclidean space

probability to move to the right and \( q \) to the left. The transition matrix \( P \) of this Markov chain is infinite:

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

This Markov chain is again irreducible. Is it transient or null persistent or positive persistent? This depends on how \( p \) and \( q \) are related. When \( p < \frac{1}{2} \) then this is a transient chain. When \( p = \frac{1}{2} \) then this is a null recurrent chain. We verified that the expected time to return to 0 is infinite when considering symmetric random walk on a discrete unbounded line. When \( p > \frac{1}{2} \) then this is positive recurrent chain.

Let us modify this chain further by removing the barrier altogether to obtain random walk on the whole infinite line with \( p \) and \( q \) the probabilities as above. Does this change the categorization of chains depending on how \( p \) and \( q \) are related? Again, when \( p < \frac{1}{2} \) then this is a transient chain. Also similarly, when \( p = \frac{1}{2} \) then this is a null recurrent chain. But when \( p > \frac{1}{2} \) then this is not a positive recurrent chain but a transient chain because this is the same situation as \( p < \frac{1}{2} \) just with the roles of \( p \) and \( q \) reversed.

Next, let us consider symmetric random walk if \( d \)-dimensional Euclidean space. States are defined to be points with all integer coordinates. In a step, the particle changes each of its \( d \) coordinates by 1 with probability \( 1/2 \) and by \(-1\) with probability \( 1/2 \), independently over the coordinates and time steps. This is an irreducible Markov chain.

We show that for \( d \leq 2 \) the Markov chain is null persistent and for \( d \geq 3 \) it is transient. Let us first consider the case of one dimension \( d = 1 \), that is, symmetric random walk on a discrete line.

Let \( g_n \) be the probability that the particle returns to the 0 point at step \( n \). Clearly we have

\[
g_{2n} = \binom{2n}{n} 2^{-2n} = \frac{(2n)!}{(n!)^2} \cdot 2^{-2n}.
\]

We use the Stirling’s approximation of factorials in the following form:

\[
n! = \Theta\left(\sqrt{n} \left(\frac{n}{e}\right)^n\right).
\]

This gives

\[
g_{2n} = \frac{\Theta\left(\sqrt{n} \left(\frac{2n}{e}\right)^{2n}\right)}{\Theta\left(n \left(\frac{n}{e}\right)^{2n}\right)} \cdot 2^{-2n} = \Theta\left(n^{-\frac{1}{2}}\right).
\]

Let \( g_k^{(d)} \) be the probability that the particle is in the start point of all coordinates zero in step \( k \) when the random walk is in \( d \) dimensions. It follows from independence among the
dimensions that

\[ g_{2n}^{(d)} = (g_{2n})^d = \Theta(n^{-d/2}) . \]

When \( d \leq 2 \) then \( \sum_n g_{2n}^{(d)} = \infty \), so the Markov chain is persistent. It is null because \( g_{2n}^{(d)} \to 0 \) for \( n \to \infty \). When \( d \geq 3 \) then \( \sum_n g_{2n}^{(d)} < \infty \) so the chain is transient.

We can summarize this fact as follows: the probability to return to the starting point is 1 when \( d \leq 2 \) and it is less than 1 for \( d \geq 3 \).

Apparently, with three dimensions, there is so much room in the Euclidean space that with positive probability the particle never returns to the starting position, while this is not the case with either one or two dimensions.

We showed directly that when \( d = 1 \) then the mean time of return is infinite. This also is the case for \( d = 2 \), which follows from general properties of null persistent states, which we will show later.

**Exercises**

1. Is a Markov chain that consists of only one communicating class irreducible in general?

2. Show that for the Markov chain in Figure 1 we have \( f_{i,j} = 1 \) for \( 0 \leq i < j \leq 3 \) and that \( f_{i,j} < 1 \) for any \( 0 \leq j \leq i \leq 3 \).

3. Consider a Markov chain with the following transition matrix:

\[
\begin{bmatrix}
1/3 & 1/3 & 0 & 1/3 \\
0 & 1/2 & 1/2 & 0 \\
0 & 1/3 & 2/3 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

Give a decomposition of the set of states into transient states and closed sets of persistent states.

4. Let a Markov chain has states 0, 1, 2 and the transition matrix \( P \):

\[
P = \begin{bmatrix}
1/2 & 1/4 & 1/4 \\
1/3 & 1/3 & 1/3 \\
0 & 0 & 1 \\
\end{bmatrix}
\]

Compute the probabilities \( f_{i,j} \) for each \( 0 \leq i \leq 2 \) and \( 0 \leq j \leq 2 \).

5. Show that \( f_{ij} = 1 \) for any two states \( i \) and \( j \) in an irreducible persistent chain.

6. Consider a Markov chain with the transition matrix

\[
\begin{bmatrix}
1/4 & 3/4 \\
2/3 & 1/3 \\
\end{bmatrix}
\]

Compute the mean return times \( \mu_i = \sum_{n \geq 0} nf_{ii}(n) \) directly from this formula.
7. Consider an infinite sequence of Bernoulli trial with probability of success $p$. Let $S$ denote success and $F$ failure. What is the expected number of trials until the first occurrence of the sequence $SFFS$?

8. We toss a fair coin over and over. Given a particular sequence $S_1S_2S_3$, where each $S_i$, for $1 \leq i \leq 3$, is either $H$ (representing heads) or $T$ (representing tails), we may consider the question: what is the expected number of tosses until the first three consecutive $S_1S_2S_3$ appear? Denote this expected number of tosses by $L(S_1S_2S_3)$.

   (a) Compute $L(HTH)$.
   
   (b) Compute $L(HHT)$.

   **Hint:** You may consider a Markov chain with three transient states representing the prefixes of the string $S_1S_2S_3$ and one absorbing state. For each state, compute the expected time until absorption when starting from this state. These expected times are a solution of a system of linear equations with constant coefficients. Surprisingly enough, the answers to (a) and (b) are different!

9. Let an infinite sequence $(c_n)$ be a convolution of the sequences $(a_n)$ and $(b_n)$, where $\sum_{n \geq 0} a_n < \infty$ and $b_n \to 0$. Show that $c_n \to 0$.

10. Show that if a state $k$ is transient or null persistent and $i$ is an arbitrary state then $\lim_{n \to \infty} p_{ik}(n) = 0$.

   **Hint:** Use (1) and Exercise 9.

11. Consider a random walk in $d$-dimensional Euclidean space in which in a step one coordinate is chosen randomly with probability $1/d$ independently in different steps and the particle changes only this one coordinate by 1 or $-1$ with probability $\frac{1}{2}$. Show that this is a null persistent chain for $d \leq 2$ and transient for $d \geq 3$.

12. Estimate the probability that the particle returns to the origin in a symmetric random walk in discrete Euclidean space $E^3$. 

---

20.4  Random walks in discrete Euclidean space 11