1. Discrete time Markov chains

Discrete-Time Markov Chains

A sequence $(X_t)_{t\in\mathbb{Z}_+}$ of discrete random variables forms a DTMC if

$$\begin{aligned}&\mathbb{P}(X_{t+1} = x_{t+1} | X_t = x_t, X_{t-1} = x_{t-1}, \cdots, X_0 = x_0)\\ &= \mathbb{P}(X_{t+1} = x_{t+1} | X_t = x_t), \qquad (1)\end{aligned}$$

for all $t, x_0, \ldots, x_{t+1}$ such that the left hand side is well defined.

Let

$$p_{i,j}(t) = \mathbb{P}(X_{t+1} = j | X_t = i).$$

We will assume that (as long as they are well defined) the transition probabilities $p_{i,j}(t)$ do not depend on t, in which case the DTMC is called time homogeneous and we write $p_{i,j} = p_{i,j}(t)$.

The Markov property (1) above can then be rewritten as:

$$\mathbb{P}(X_{t+1} = j | X_t = i, X_{t-1} = x_{t-1}, \cdots, X_0 = x_0) = p_{i,j},$$

for all $t, i, j, x_0, \ldots, x_{t-1}$ such that the left hand side is well defined.

A more general picture

Henceforth all our DTMCs are time-homogeneous.

One can infer from the Markov property that

$$\mathbb{P}(\cap_{i=0}^{n}\{X_i = x_i\}) = \mathbb{P}(X_0 = x_0)\prod_{i=0}^{n-1} p_{x_i,x_{i+1}}.$$

If the left hand side is positive then this follows by recursive conditioning. If the left hand side is 0 then either $\mathbb{P}(X_0 = x_0) = 0$ or there is a smallest $m \in [1, n]$ such that $\mathbb{P}(\cap_{i=0}^{m}\{X_i = x_i\}) = 0$, and the recursive conditioning applies to this probability.

More generally the Markov property implies (assuming that these are well-defined)

$$\begin{aligned}&\mathbb{P}\big((X_{n+1}, X_{n+2}, \ldots, X_{n+k}) \in B \big| X_n = x, (X_{n-1}, \ldots, X_0) \in A\big)\\ =&\mathbb{P}\big((X_{n+1}, X_{n+2}, \ldots, X_{n+k}) \in B \big| X_n = x\big).\end{aligned}$$

I.e. if you know the present state, then receiving information about the past tells you nothing more about the future.

Transition matrix

If the *state space* $\mathcal{S}$ (i.e. the set of possible values that the elements of the sequence can take) has $m \in \mathbb{N}$ elements, then by relabelling the values if necessary, we may assume that $\mathcal{S} = \{1, 2, \ldots, m\}$.

For a DTMC, we define the (one step)-transition matrix to be a matrix with rows and columns corresponding to the states of the process and whose ij-th entry is $p_{i,j}$. So

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,m} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,m} \end{pmatrix}.$$

Transition matrix

For a transition matrix of a DTMC:

- Each entry is ≥ 0.
- Each row sums to 1.

Any square matrix having these two properties is called a stochastic matrix.

(If the state space is infinite we will still refer to the infinite matrix containing the $p_{i,j}$ as the transition matrix. It is also a stochastic matrix (albeit an infinite one).)

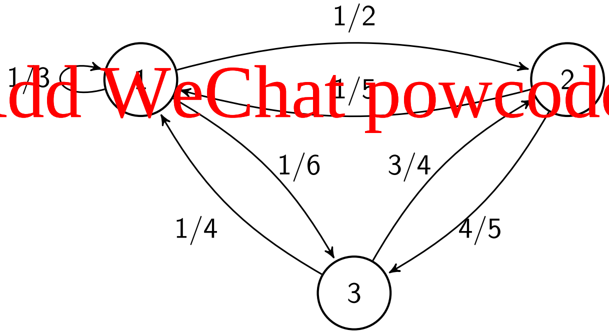
Transition diagram

We can associate a weighted "directed graph" (called the transition diagram) with a stochastic matrix by letting the nodes correspond to states and putting in an arc/edge jk with weight $p_{j,k} > 0$ on it.

Example: if

$$P = \begin{pmatrix} 1/3 & 1/2 & 1/6 \\ 1/5 & 0 & 4/5 \\ 1/4 & 3/4 & 0 \end{pmatrix},$$

then the transition diagram can be drawn as follows.



Examples

- Suppose that the $(X_t)_{t\in\mathbb{Z}_+}$ are i.i.d. random variables with $\mathcal{S} = \{1, \dots, k\}$ and $\mathbb{P}(X_t = i) = p_i$. What does the transition matrix look like?
- A communication system transmits the digits 0 and 1 at discrete times. Let X_t denote the digit transmitted at time t. At each time point, there is a probability p that the digit will not change and prob $1 - p$ it will change. Find the transition matrix and draw the transition diagram.

Examples

- Suppose that whether or not it rains tomorrow depends on previous weather conditions only through whether or not it is raining today. Suppose also that if it rains today, then it will rain tomorrow with probability p and if it does not rain today, then it will rain tomorrow with probability q. If we say that $X_t = 1$ if it rains on day t and $X_t = 0$ otherwise, then $(X_t)_{t \in \mathbb{Z}_+}$ is a two-state Markov chain.

- Let $(Y_i)_{i \in \mathbb{N}}$ be i.i.d. random variables with $\mathbb{P}(Y_i = 1) = p$ and $\mathbb{P}(Y_i = -1) = 1 - p$. Let $S_0 = 0$ and $S_n = \sum_{i=1}^{n} Y_i$ for each $n \in \mathbb{N}$. Then $(S_n)_{n \in \mathbb{Z}_+}$ is a Markov chain called a simple random walk. If $p = 1/2$ it is called a symmetric (or unbiased) simple random walk.

n-step transition probabilities

The n-step transition probabilities $\mathbb{P}(X_{t+n} = j | X_t = i)$ of a (time-homogeneous) DTMC do not depend on t. For $n = 1, 2, \cdots$, we denote them by

$$p_{i,j}^{(n)} = \mathbb{P}(X_{t+n} = j | X_t = i).$$

It is also convenient to use the notation

$$p_{i,j}^{(0)} := \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

Chapman-Kolmogorov equations

The Chapman-Kolmogorov equations show how we can calculate the n-step transition probabilities $p_{i,j}^{(n)}$ from smaller-step transition probabilities. For $n = 1, 2, \cdots$ and any $r = 1, 2, \cdots, n$,

$$p_{i,j}^{(n)} = \sum_{k \in \mathcal{S}} p_{i,k}^{(r)} p_{k,j}^{(n-r)}.$$

Interpretation: *In order to go from i to j in n steps, you have to be somewhere after r steps!*

n-step transition matrix

If we define the n-step transition matrix as

$$P^{(n)} = \begin{pmatrix} p_{1,1}^{(n)} & p_{1,2}^{(n)} & \ddots & \ddots \\ p_{2,1}^{(n)} & p_{2,2}^{(n)} & p_{2,3}^{(n)} & \ddots \\ \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

then the Chapman-Kolmogorov equations can be written in the matrix form

$$P^{(n)} = P^{(r)} P^{(n-r)}$$

with $P^{(1)} = P$. By mathematical induction, it follows that

$$P^{(n)} = P^n,$$

the nth power of P.

Distribution of a DTMC

How do we determine the distribution of a DTMC?

To uniquely determine the distribution we need 2 ingredients:

- the *initial distribution* $\boldsymbol{\pi}^{(0)} = (\pi_i^{(0)})_{i \in \mathcal{S}}$, where for each $j \in \mathcal{S}$, $\pi_j^{(0)} = \mathbb{P}(X_0 = j)$, and
- the transition matrix P.

In principle, we can use these and the Markov property to derive the finite dimensional distributions, e.g.

$$\mathbb{P}(X_0 = x_0, X_{t_1} = x_1, X_{t_2} = x_2, \cdots, X_{t_k} = x_k) = \pi_{x_0}^{(0)} p_{x_0,x_1}^{(t_1)} \cdots p_{x_{k-1},x_k}^{(t_k - t_{k-1})},$$

although the calculations are often intractable.

Example:

Suppose $\mathbb{P}(X_0 = 1) = 1/3$, $\mathbb{P}(X_0 = 2) = 0$, $\mathbb{P}(X_0 = 3) = 1/2$, $\mathbb{P}(X_0 = 4) = 1/6$ and

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

- Find the distribution of X_1.
- Find $\mathbb{P}(X_{n+2} = 2 | X_n = 4)$.
- Find $\mathbb{P}(X_3 = 2, X_2 = 3, X_1 = 1)$.

A fundamental question

What proportion of time does the chain spend in each state in the long run?

(and when does this question even make sense?)

To answer this appropriately we need to introduce a lot of concepts!

Classification of states/chains

Here are some definitions.

- State j is accessible from state i, denoted by $i \to j$, if there exists an $n \geq 0$ such that $p_{i,j}^{(n)} > 0$. That is, either $j = i$ or we can get from i to j in a finite number of steps.
- If $i \to j$ and $j \to i$, then states i and j communicate, denoted by $i \leftrightarrow j$.
- A state j is an absorbing state if $p_{j,j} = 1$.

Example

Draw a transition diagram associated with the transition matrix below and determine which states communicate with each other. Are there any absorbing states?

$$P = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

The communication relation

The communication relation $\leftrightarrow$ has the properties:

- $j \leftrightarrow j$ (reflexivity),
- $j \leftrightarrow k$ if and only if $k \leftrightarrow j$ (symmetry), and
- if $j \leftrightarrow k$ and $k \leftrightarrow i$, then $j \leftrightarrow i$ (transitivity).

A relation that satisfies these properties is known as an equivalence relation.

Communicating classes and irreducibility

Consider a set $\mathcal{S}$ whose elements can be related to each other via any equivalence relation $\Leftrightarrow$.

Then $\mathcal{S}$ can be partitioned into a collection of disjoint subsets $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \ldots, \mathcal{S}_M$ (where M might be infinite) such that $j, k \in \mathcal{S}_m$ if and only if $j \Leftrightarrow k$.

So the state space $\mathcal{S}$ of a DTMC is partitioned into communicating classes by the communication relation $\leftrightarrow$.

If a DTMC has only one communicating class (i.e., all states communicate) then it is called an irreducible DTMC. Otherwise it is called reducible.

Hitting and return probabilities

Let

$$h_{i,j} = \mathbb{P}(j \in \{X_0, X_1, \ldots\} | X_0 = i),$$

which is the probability that we ever reach state j, starting from state i. The quantities $h_{i,j}$ are referred to as hitting probabilities.

Note that

- $h_{i,i} = 1$
- $h_{i,j} > 0$ if and only if $i \to j$.

Let

$$f_i = \mathbb{P}(i \in \{X_1, X_2, \ldots\} | X_0 = i),$$

which is the probability that we ever return to state i, starting from i. The quantities f_i are referred to as return probabilities.

Hitting and return times

Let $T(j) = \inf\{n \geq 0 : X_n = j\}$ denote the first hitting time of j (it is a stopping time). Then

$$m_{i,j} := \mathbb{E}[T(j)|X_0 = i],$$

is the expected time to reach state j starting from state i. By definition

- $m_{j,j} = 0$,
- if $h_{i,j} < 1$ then $\mathbb{P}(T(j) = \infty|X_0 = i) > 0$ so $m_{i,j} = \infty$,
- if $h_{i,j} = 1$ then $m_{i,j}$ may or may not be finite.

Let $T^+(i) = \inf\{n > 0 : X_n = i\}$. Then

$$\mu_i = \mathbb{E}[T^+(i)|X_0 = i],$$

is the expected time to return to state i.

The strong Markov property

The Markov property as defined, holds at each *fixed time*. We would like the same to be true at certain random times such as hitting times. The property that we need (which can be proved from the Markov property) is called the

Strong Markov property: Let $(X_t)_{t \in \mathbb{Z}_+}$ be a (time-homogeneous) DTMC, and T be a *stopping time* for the chain. Then

$$\begin{aligned}\mathbb{P}(X_{T+1} = j \mid T = t, X_0 = x_0, \ldots, X_T = i) \\ = \mathbb{P}(X_{T+1} = j \mid T < \infty, X_T = i) = p_{i,j}.\end{aligned}$$

This says that looking at the next step of a Markov chain at a stopping time is the same as starting the process from the random state X_T (provided that T is finite). As with the ordinary Markov property, this can be generalized to handle general future events etc.

Recurrence and transience

One can classify individual *states* as *recurrent* or *transient*. We will be mostly interested in applying such a classification to *irreducible chains* where we can avoid some technicalities.

Definition: An irreducible DTMC is recurrent if $h_{i,j} = 1$ for every $i, j \in \mathcal{S}$. Otherwise it is transient.

Characterizing recurrence

Let $\Delta_i(j)$ be the time between the $(i+1)$st and ith visit to state j, and let $N(j)$ be the number of visits to state j. Suppose that $X_0 = j$.

If $f_j = 1$ then each $\Delta_i(j)$ is finite since the chain is certain to return to j in finite time, and $N(j) = \infty$.

If $f_j < 1$ then with probability $1 - f_j > 0$ it will never return to j. From the Markov property we see that $N(j)$ has a geometric distribution.

Specifically, for $n \geq 1$,

$$\mathbb{P}(N(j) = n | X_0 = j) = f_j^{n-1}(1 - f_j).$$

This implies that $\gamma_j := \mathbb{E}[N(j) | X_0 = j] = \frac{1}{1-f_j} < \infty$.

Characterizing recurrence

Theorem: For an irreducible DTMC the following are equivalent:

(i) the chain is recurrent

(ii) $f_i = 1$ for every $i \in \mathcal{S}$

(iii) $f_i = 1$ for some $i \in \mathcal{S}$

(iv) $\gamma_i = \infty$ for some $i \in \mathcal{S}$

(v) $\gamma_i = \infty$ for every $i \in \mathcal{S}$

(Note that if an irreducible chain is transient, then it visits each state only a finite number of times, so the limiting proportion of time spent in each state is 0).

Characterizing recurrence

$(iii) \Leftrightarrow (iv)$: essentially done 2 slides ago.

$(i) \Rightarrow (iii)$: $f_i = \sum_{j \in \mathcal{S}} p_{i,j} h_{j,i} = \sum_{j \in \mathcal{S}} p_{i,j} 1 = 1$

$(iii) \Rightarrow (i)$: Suppose $f_i = 1$. Let $j \in \mathcal{S}$.
Since $i \to j$ we must have $h_{j,i} = 1$ (otherwise the chain could "escape" from i by visiting j). Similarly, since $f_i = 1$ and $i \to j$, every time the chain reaches state i it has a fixed (positive) probability of hitting j before returning to i, so $h_{i,j} = 1$.
Now for $j, k \in \mathcal{S}$ since we are guaranteed to hit i starting from j and guaranteed to hit k started from i, we are guaranteed to hit k starting from j, i,e. $h_{j,k} = 1$.

others: exercise

Irreducible finite-state chains are recurrent

(We will state a stronger result later)

Corollary: Irreducible finite-state chains are recurrent.

Proof: Let $\mathcal{S} = \{1, \ldots, k\}$. Let $N(j)$ denote the number of visits to j.

Now $\sum_{j=1}^{k} N(j) = \infty$, so for any i with $\mathbb{P}(X_0 = i) > 0$, we have $\mathbb{E}[\sum_{j=1}^{k} N(j) | X_0 = i] = \infty$. Thus there must exist $j \in \mathcal{S}$ such that $\mathbb{E}[N(j) | X_0 = i] = \infty$.

For this j, $\gamma_j \geq \mathbb{E}[N(j) | X_0 = i] = \infty$. $\square$

Simple Random Walk

Recall that $S_0 = 0$ and $S_n = \sum_{i=1}^{n} Y_i$, for $n \geq 1$, where $(Y_i)_{i \in \mathbb{N}}$ are i.i.d. random variables with $\mathbb{P}(Y_i = 1) = p = 1 - \mathbb{P}(Y_i = -1)$.

S_n is a sum of i.i.d. random variables. By the law of large numbers we have $S_n/n \to \mathbb{E}[Y_1] = 2p - 1$. Therefore:

If $p > 1/2$ then the random walk escapes to $+\infty$, and so the number of visits to any state is finite. This implies that the chain is transient (and similarly if $p < 1/2$).

Let us prove this another way (and also handle the case $p = 1/2$, for which the LLN does not give us the answer).

Simple Random Walk

Note that the number of visits to 0 satisfies

$$N(0) = \sum_{m=0}^{\infty} \mathbb{1}_{\{S_m=0\}},$$

so (since the walk starts at 0)

$$\mathbb{E}[N(0)] = \sum_{m=0}^{\infty} \mathbb{E}[\mathbb{1}_{\{S_m=0\}}] = \sum_{m=0}^{\infty} \mathbb{P}(S_m = 0) = \sum_{m=0}^{\infty} p_{0,0}^{(m)}.$$

Note that $p_{j,j}^{(m)} = 0$ if m is odd. If $m = 2n$ then

$$p_{j,j}^{(m)} = p_{j,j}^{(2n)} = \binom{2n}{n} p^n (1-p)^n.$$

Stirling's formula $n! \approx \sqrt{2\pi n} n^n e^{-n}$ gives us the fact that

$$p_{j,j}^{(2n)} \approx \frac{(4p(1-p))^n}{\sqrt{n\pi}},$$

and the series $\sum_{n=0}^{\infty} p_{j,j}^{(2n)}$

- diverges if $4p(1-p) = 1$ (i.e. $p = 1/2$), so the DTMC is recurrent
- converges if $4p(1-p) < 1$ (i.e. $p \neq 1/2$) (compare to geometric series), so the DTMC is transient.

We shall see another way of proving this same result later.

Periodicity

The simple random walk illustrates another phenomenon that can occur in DTMCs - periodicity.

Definition: For a DTMC, a *state* $i \in \mathcal{S}$ has period $d(i) \geq 1$ if $\{n \geq 1 : p_{i,i}^{(n)} > 0\}$ is non-empty and has greatest common divisor $d(i)$.

If state j has period 1, then we say that it is aperiodic (otherwise it is called periodic).

It turns out that if $i \leftrightarrow j$ then $d(i) = d(j)$, thus we can make the following definition:

Definition: An irreducible DTMC is periodic with period d if any (hence every) state has period $d > 1$. Otherwise it is aperiodic.

Examples

- The simple random walk is periodic with period $d = 2$.
- Which of the following are transition matrices for periodic chains?

$$P = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad P = \begin{pmatrix} 0 & 0 & 0.5 & 0.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

States in a communicating class have same period

Assume that state j has period $d(j)$ and $j \leftrightarrow k$. Then, as before, there must exist s and t such that $p_{j,k}^{(s)} > 0$ and $p_{k,j}^{(t)} > 0$. We know straight away that $d(j)$ divides $s + t$ since it is possible to go from j to itself in $s + t$ steps.

Now take a path from k to itself in r steps. If we concatenate our path from j to k in s steps, this r step path, and our path from from k to j in t steps, we have an $s + r + t$ step path from j to itself. So $d(j)$ divides $s + r + t$ which means that $d(j)$ divides r. So the $d(j)$ divides the period $d(k)$ of k.

Now we can switch j and k in the argument to conclude that $d(k)$ divides $d(j)$ which means that $d(j) = d(k)$, and all states in the same communicating class have a common period.

Computing hitting probabilities

For $i \in \mathcal{S}$ and $A \subset \mathcal{S}$, let $h_{i,A}$ denote the probability that the chain ever visits a state in A, starting from state i. If $A = \{j\}$ is a single state then we have seen this before: $h_{i,\{j\}} = h_{i,j}$.

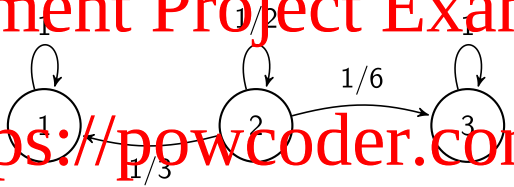
Let $T(A)$ denote the first time we reach a state in A. Then $h_{i,A} = \mathbb{P}(T(A) < \infty | X_0 = i)$.
We have

- if $i \in A$ then $h_{i,A} = 1$
- $h_{i,A} > 0$ if and only if $i \to j$ for some $j \in A$.
- $h_{i,A} = \sum_{j \in \mathcal{S}} p_{i,j} h_{j,A}$ if $i \notin A$.

This is a set of linear equations.

A simple example

Consider a Markov chain with $\mathcal{S} = \{1, 2, 3\}$, and $p_{11} = p_{33} = 1$ and $p_{22} = 1/2$, $p_{21} = 1/3$, $p_{23} = 1/6$. This Markov chain has transition diagram



The chain has (two) absorbing states. Find $h_{i,1}$ for each i.

Clearly $h_{1,1} = 1$, and $h_{3,1} = 0$. Finally,

$$h_{2,1} = \frac{1}{3}h_{1,1} + \frac{1}{2}h_{2,1} + \frac{1}{6}h_{3,1} = \frac{1}{3} + \frac{1}{2}h_{2,1}.$$

So $h_{2,1} = 2/3$.

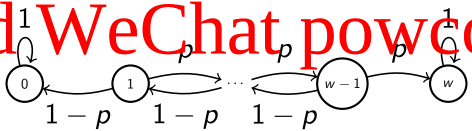
(There is another easy way to see this)

Example: Gambler's ruin

Starting with \$i, a gambler makes repeated bets of \$1 on a game of chance that she has probability p of winning on each attempt (independent of the past). She stops as soon as she reaches \$$w$ (where $w > i$) or \$0.

What is the probability that the gambler ends up with \$0?

Let $(X_n)_{n\geq 0}$ be a DTMC with state space $\mathcal{S} = \{0, 1, \ldots, w\}$, and transition probabilities $p_{i,i+1} = p$ and $p_{i,i-1} = 1 - p$ if $0 < i < w$ and $p_{0,0} = 1 = p_{w,w}$. The transition diagram is



Find the hitting probabilities $(h_{i,0})_{i\in\mathcal{S}}$.

Example: Gambler's ruin

We have $h_{0,0} = 1$, $h_{w,0} = 0$ and otherwise

$$h_{i,0} = p h_{i+1,0} + (1-p) h_{i-1,0}.$$

Rearrange to get

$$p(h_{i,0} - h_{i+1,0}) = (1-p)(h_{i-1,0} - h_{i,0}).$$

Let $u_i = h_{i-1,0} - h_{i,0}$, and $\alpha = (1-p)/p$. Then for $1 \le i < w$

$$u_{i+1} p = u_i (1-p).$$

Thus, $u_{i+1} = \alpha u_i$ and $u_1 = 1 - h_{1,0}$. It follows that for $k \le w$

$$u_k = \alpha^{k-1} u_1.$$

Example: Gambler's ruin

$$u_k = \alpha^{k-1} u_1.$$

On the other hand,

$$h_{i,0} = h_{0,0} + \sum_{m=1}^{i} (h_{m,0} - h_{m-1,0}) = 1 - u_1 \sum_{m=0}^{i-1} \alpha^m.$$

Since $h_{w,0} = 0$ we have that

$$u_1 = \left(\sum_{m=0}^{w-1} \alpha^m \right)^{-1},$$

and therefore

$$h_{i,0} = 1 - \frac{\sum_{r=0}^{i-1} \alpha^r}{\sum_{m=0}^{w-1} \alpha^m} = \frac{\sum_{r=i}^{w-1} \alpha^r}{\sum_{m=0}^{w-1} \alpha^m}.$$

Example: Gambler's ruin

So

$$h_{i,0}(w) = 1 - \frac{\sum_{r=0}^{i-1} \left(\frac{1-p}{p}\right)^r}{\sum_{m=0}^{w-1} \left(\frac{1-p}{p}\right)^m} = \frac{\sum_{r=i}^{w-1} \left(\frac{1-p}{p}\right)^r}{\sum_{m=0}^{w-1} \left(\frac{1-p}{p}\right)^m}.$$

If $p = 1/2$ we get

$$h_{i,0} = \frac{w-i}{w}.$$

Exercise: check that these do indeed satisfy the equations that we started with!

Exercise: find $\lim_{w \to \infty} h_{1,0}(w)$.

Example: Gambler's ruin with Martingales

When $p = 1/2$ in the Gambler's ruin problem, the chain is a *Martingale:*

$$\mathbb{E}[X_{n+1}|X_n] = X_n.$$

It is *bounded* because $0 \leq X_n \leq w$ for each n.
Let $T = T(\{0, w\})$ denote the first time that we hit 0 or w. Then since $(X_n)_{n\geq 0}$ is a bounded Martingale

$$\mathbb{E}[X_T] = \mathbb{E}[X_0].$$

If we start with $X_0 = i$ then $\mathbb{E}[X_0] = i$. Also, $\mathbb{P}(X_T = 0) = h_{i,0}$ and $\mathbb{P}(X_T = w) = 1 - h_{i,0}$ so

$$i = \mathbb{E}[X_T] = 0 \times h_{i,0} + w \times (1 - h_{i,0}).$$

Solving gives $h_{i,0} = \frac{w-i}{w}$ as before.

Example: Simple random walk

Consider the simple random walk, which is irreducible. Let $A = \{0\}$. Then $h_{0,0} = 1$ and $h_{i,0} = ph_{i+1,0} + (1-p)h_{i-1,0}$ for $i \neq 0$. Then $h_{i,0} = 1$ for all i satisfies these equations, regardless of p. But we know that the walk is transient unless $p = 1/2$, so how is this possible?

If $\mathcal{S}$ is infinite then there need not be a unique solution to these equations.

Above we have not found the solution that we are looking for (except when $p = 1/2$).

Example: Simple random walk

If $i > 0$ then

$$h_{i,0} = p h_{i+1,0} + (1-p) h_{i-1,0}.$$

Let $x = h_{1,0}$. Then $x = p h_{2,0} + 1 - p$. Also

$$h_{i,0} = h_{i,i-1} h_{i-1,0} = h_{1,0} h_{i-1,0}.$$

In particular with $i = 2$ we get $h_{2,0} = h_{1,0} h_{1,0} = x^2$. Therefore,

$$x = p x^2 + (1-p).$$

Solving the quadratic gives solutions

$$x = 1, \quad \text{and } x = (1-p)/p.$$

If $p \leq 1/2$ then $h_{1,0} = 1$ is the only possible value for $h_{1,0}$, and in this case $h_{i,0} = 1$ for every $i > 0$.

If $p > 1/2$ then the walk is transient (to the right), so $h_{i,0}$ cannot be 1 for $i > 0$. Thus $h_{1,0} = (1-p)/p$ and $h_{i,0} = ((1-p)/p)^i$.

Do the above expressions look familiar?

The "correct" solution to the hitting probability equations

Theorem: The vector of hitting probabilities $(h_{i,B})_{i \in \mathcal{S}}$ is the unique minimal non-negative solution to the equations

$$h_{i,B} = \begin{cases} 1 & \text{if } i \in B \\ \sum_{j \in \mathcal{S}} p_{i,j} h_{j,B}, & \text{otherwise} \end{cases}$$

Proof: Let $h_{i,B}^{(n)} = \mathbb{P}(T(B) \leq n | X_0 = i)$. Let $(x_i)_{i \in \mathcal{S}}$ be a non-negative solution. We'll show that $h_{i,B}^{(n)} \leq x_i$ for each $n \in \mathbb{Z}_+$, by induction. This is sufficient since $h_{i,B} = \lim_{n \to \infty} h_{i,B}^{(n)}$.

For $n = 0$, note that $h_{i,B}^{(0)} = 1$ if $i \in B$ and $h_{i,B}^{(0)} = 0$ if $i \notin B$. Since $(x_i)_{i \in \mathcal{S}}$ are non-negative and equal 1 for $i \in B$, we have $h_{i,B}^{(0)} \leq x_i$ for all $i \in \mathcal{S}$.

Proceeding by induction, suppose that $h_{i,B}^{(n)} \leq x_i$ for all $i \in \mathcal{S}$. Then

$$\begin{aligned} h_{i,B}^{(n+1)} = \sum_{j \in \mathcal{S}} p_{i,j} h_{j,B}^{(n)} &\leq \sum_{j \in \mathcal{S}} p_{i,j} x_j \quad \text{(by the induction hypothesis)} \\ &= x_i \quad ((x_i)_{i \in \mathcal{S}} \text{ is a solution}). \quad \square \end{aligned}$$

Exercise:

Check that what this theorem says about the simple random walk agrees with what we already proved.

Difference and differential equations

The equation $h_{i,0} = ph_{i+1,0} + (1-p)h_{i-1,0}$ is a second-order linear difference equation with constant coefficients.

These can be solved in a similar way to second-order linear differential equations with constant coefficients, which you learned about in Calculus II or accelerated Mathematics II.

Recall that, to solve

$$a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + cy = 0,$$

we try a solution of the form $y = y(t) = e^{\lambda t}$ to obtain the Characteristic Equation

$$a\lambda^2 + b\lambda + c = 0.$$

Differential equations

If the characteristic equation has distinct roots, λ_1 and λ_2, the general solution has the form

$$y = Ae^{\lambda_1 t} + Be^{\lambda_2 t}.$$

If the roots are coincident, the general solution has the form

$$y = Ae^{\lambda_1 t} + Bte^{\lambda_1 t}.$$

In both cases, the values of the constants A and B are determined by the initial conditions.

Difference equations

The method for solving second-order linear difference equations with constant coefficients is similar. To solve

$$av_{j+1} + bv_j + cv_{j-1} = 0,$$

we try a solution of the form $v_j = z^j$ to obtain the Characteristic Equation

$$az^2 + bz + c = 0.$$

Difference equations

If this equation has distinct roots, z_1 and z_2, the general solution has the form

$$y = Az_1^j + Bz_2^j.$$

If the roots are coincident, the general solution has the form

$$y = Az_1^j + Bjz_1^j.$$

The values of the constants A and B need to be determined by boundary equations, or other information that we have.

Application to simple random walk

The characteristic equation of

$$h_{i,0} = p h_{i+1,0} + (1-p) h_{i-1,0}$$

is

$$pz^2 - z + (1-p) = 0$$

which (as we have seen before) has roots $z = 1$ and $z = (1-p)/p$.
If $(1-p)/p \neq 1$, the general solution for $j \geq 1$ is of the form

$$h_{i,0} = A + B\left(\frac{1-p}{p}\right)^i .$$

Application to simple random walk

If $(1-p)/p > 1$, then the general solution is

$$h_{j,0} = A + B\left(\frac{1-p}{p}\right)^j.$$

Similarly, if $(1-p)/p = 1$, the general solution is of the form

$$h_{j,0} = A + Bj.$$

In either case, these can only be probabilities if $B = 0$ and then notice

$$A = h_{1,0} = ph_{2,0} + (1-p) = pA + (1-p),$$

so $A = 1$. This makes sense because $p \leq 1/2$ and so we have a neutral or downward drift.

Expected hitting times in DTMCs

For $i \in \mathcal{S}$ and $A \subset \mathcal{S}$ let $m_{i,A}$ denote the expected time to reach A starting from i. Note that $m_{i,\{j\}} = m_{i,j}$ is a special case that we have already seen before.

By definition and the Markov property we have that

$$m_{i,A} = \begin{cases} 0, & \text{if } i \in A \\ 1 + \sum_{j \in \mathcal{S}} p_{i,j} m_{j,A}, & \text{otherwise.} \end{cases}$$

Theorem: The vector $(m_{i,A})_{i \in \mathcal{S}}$ of mean hitting times is the minimal non-negative solution to

$$m_{i,A} = \begin{cases} 0, & \text{if } i \in A \\ 1 + \sum_{j \in \mathcal{S}} p_{i,j} m_{j,A}, & \text{otherwise.} \end{cases}$$

Computing expected hitting times

Example: the symmetric simple random walk is recurrent. Now,

$$m_{1,0} = 1 + \frac{1}{2}m_{0,0} + \frac{1}{2}m_{2,0} = 1 + \frac{1}{2}m_{2,0}.$$

But $m_{2,0} = m_{2,1} + m_{1,0} = 2m_{1,0}$ so

$$m_{1,0} = 1 + m_{1,0}.$$

This has no finite solution, so $m_{1,0} = \infty$.

Positive recurrence

Recall that an irreducible DTMC is recurrent iff the return probabilities satisfy $f_j = 1$ for every $j \in \mathcal{S}$.

Recall that the expected return times are denoted by μ_j.

Definition: A recurrent DTMC is *positive recurrent* if $\mu_j < \infty$ for every $j \in \mathcal{S}$. Otherwise it is *null recurrent*.

Finite irreducible DTMCs are positive recurrent

Lemma: Any irreducible DTMC with finite state space is positive recurrent.

Proof: Let $j \in \mathcal{S}$. Then there exists $n_0 \in \mathbb{N}$ such that $\mathbb{P}(T(j) \leq n_0 | X_0 = k) > 0$ for every $k \in \mathcal{S}$. Therefore there exists some $\varepsilon > 0$ such that $\mathbb{P}(T(j) \leq n_0 | X_0 = k) \geq \varepsilon$ for every $k \in \mathcal{S}$. Starting from any state i, observe whether the chain has reached j within n_0 steps. This has probability at least ε. If it has not reached j then observe it for the next n_0 steps... The number of blocks of n_0 steps that we have to observe is dominated by a Geometric(ε) random variable. Therefore $m_{i,j} \leq n_0/\varepsilon$ for every i.

Thus, $\mu_j \leq 1 + n_0/\varepsilon < \infty$. $\square$

Simple symmetric random walk:

Recall that simple random walk with $p = 1/2$ is recurrent.

Recall that for this walks $m_{1,0} = \infty$.

Therefore $\mu_0 = \infty$ as well.

So simple symmetric random walk is null recurrent.

A question

Some properties of a Markov chain depend only on the transition matrix P, while others depend on the starting state of the chain as well.

We will now focus on the long run behaviour of DTMCs and one of the interesting questions to keep in mind as we do this is:

When does a property of the MC depend on the initial distribution?

Long run behaviour of DTMCs

Let $N_n(j) = \sum_{i=0}^{n-1} \mathbb{1}_{\{X_i = j\}}$ denote the number of visits to j before time n. Then $Y_n(j) = \frac{1}{n} N_n(j)$ is the proportion of time spent in state j before time n. Note that $Y_n(j)$ is a random variable.

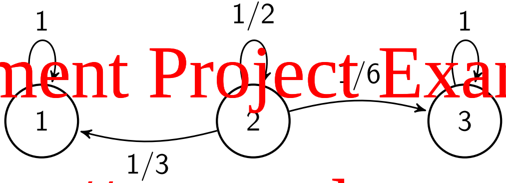
Fundamental question: what is the long run proportion of the time spent in state j? I.e. what is $Y(j) = \lim_{n\to\infty} Y_n(j)$?

- We have argued that this will be zero if the chain is irreducible and transient.
- For an irreducible DTMC we will get an answer to this question that does not depend on the initial distribution.
- If the chain is reducible then the answer to this question may be random, and may depend on the initial distribution.

A related fundamental question is: what is $\lim_{t\to\infty} \mathbb{P}(X_t = j)$?

A reducible example

Consider a Markov chain with transition diagram



Find the limiting proportion of time spent in state 1.

If $X_0 = 1$ then $X_t = 1$ for all t so $Y(1) = 1, Y(2) = 0, Y(3) = 0$.
If $X_0 = 3$ then $X_t = 3$ for all t so $Y(1) = 0, Y(2) = 0, Y(3) = 1$.
If $X_0 = 2$ then $Y(1)$ is random, with $\mathbb{P}(Y(1) = 1) = h_{2,1} = 2/3$
and $\mathbb{P}(Y(1) = 0) = 1 - h_{2,1} = 1/3$
So in general, $Y(1)$ is a Bernoulli random variable with

$$\mathbb{P}(Y(1) = 1) = \mathbb{P}(X_0 = 1) + \frac{2}{3} \cdot \mathbb{P}(X_0 = 2).$$

Long run behaviour of irreducible DTMCs

Recall that μ_j denotes the mean return time to state j.

Theorem: (∗) Let $(X_t)_{t\in\mathbb{Z}_+}$ be an irreducible (discrete time, time-homogeneous) Markov chain with state space $\mathcal{S}$. Then

$$\mathbb{P}\left(\frac{N_n(i)}{n} \to \frac{1}{\mu_i}\right) = 1, \tag{2}$$

where we interpret the limit $\frac{1}{\mu_i}$ as $0 = 1/\infty$ if the chain is not positive recurrent.

Idea of proof: If the chain is transient then $N_n(i)$ is bounded so $N_n(i)/n$ converges to 0. Otherwise the chain visits state i about once in every μ_i steps, so the proportion of time spent at i is $1/\mu_i$. (We make this rigorous using the law of large numbers).

Computing mean return times

Note that

$$\mu_i = 1 + \sum_{j \in \mathcal{S}} p_{i,j} m_{j,i}.$$

So if you can find $(m_{j,i})_{j \in \mathcal{S}}$ then you are done!

Stationary distribution

A vector $\boldsymbol{\pi} = (\pi_i)_{i \in \mathcal{S}}$ with non-negative entries is a stationary measure for a stochastic matrix P (or a DTMC with transition matrix P) if

$$\pi_i = \sum_{j \in \mathcal{S}} \pi_j p_{j,i}, \qquad \text{for every } i \in \mathcal{S}.$$

The above equations are called the full balance equations. In Matrix form this is $\boldsymbol{\pi} P = \boldsymbol{\pi}$ (with $\boldsymbol{\pi}$ as a row vector).

If $\boldsymbol{\pi}$ is a stationary measure with $\sum_{i \in \mathcal{S}} \pi_i = 1$ then $\boldsymbol{\pi}$ is called a stationary distribution for P.

Stationary distribution

Suppose that $\boldsymbol{\pi}$ is a stationary distribution for a DTMC $(X_n)_{n\in\mathbb{Z}_+}$ with transition matrix P, and suppose that $\mathbb{P}(X_0 = i) = \pi_i$ for each $i \in \mathcal{S}$. Then

$$\mathbb{P}(X_1 = i) = \sum_{j\in\mathcal{S}} \mathbb{P}(X_0 = j)p_{j,i} = \sum_{j\in\mathcal{S}} \pi_j p_{j,i} = \pi_i,$$

since $\boldsymbol{\pi}$ satisfies the full balance equations. This says that $\mathbb{P}(X_1 = i) = \pi_i$ too. By induction (exercise) we get:

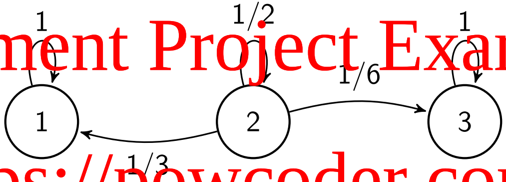
Lemma: Suppose that $(\pi_i)_{i\in\mathcal{S}}$ is a stationary distribution for P. Let $(X_n)_{n\in\mathbb{Z}_+}$ be DTMC with transition matrix P and initial distribution equal to $\boldsymbol{\pi}$ (i.e. $\mathbb{P}(X_0 = i) = \pi_i$ for each $i \in \mathcal{S}$), then

$$\mathbb{P}(X_n = i) = \pi_i, \quad \text{for every } i \in \mathcal{S}, n \in \mathbb{N}.$$

So, if your initial distribution is a stationary distribution then your distribution at any time is the same (hence the use of the term stationary)!

Examples:

Find all stationary distributions for a Markov chain with the following transition diagram:



The full balance equations are:

$$\pi_1 = \pi_1 + \frac{1}{3} \cdot \pi_2, \qquad \pi_2 = \frac{1}{2}\pi_2, \qquad \pi_3 = \pi_3 + \frac{1}{6} \cdot \pi_2.$$

The second equation gives $\pi_2 = 0$. The other equations then reduce to $\pi_1 = \pi_1$ and $\pi_3 = \pi_3$. Thus, any vector $(\pi_1, \pi_2, \pi_3) = (a, 0, b)$ with $a, b \geq 0$ is a stationary measure. To get a stationary distribution, we require that $a + b = 1$, so the set of stationary *distributions* is the set of vectors of the form $(a, 0, 1 - a)$ with $a \in [0, 1]$.

Existence and uniqueness

We have just seen an example of a DTMC without a unique stationary distribution. The following is our main existence and uniqueness result.

Theorem: ($**$) An irreducible (time-homogeneous) DTMC with countable state space $\mathcal{S}$ has a stationary measure. It has a unique stationary distribution if and only if the chain is positive recurrent, and in this case $\pi_i = 1/\mu_i$ for each $i \in \mathcal{S}$.

Combining Theorems ($*$) and ($**$) we see that for a positive recurrent irreducible DTMC, the long run proportion of time spent in state i is the stationary probability π_i.

Limiting distribution

A distribution $(a_i)_{i \in \mathcal{S}}$ is called the limiting distribution for a DTMC $(X_t)_{t \in \mathbb{Z}_+}$ if

$$\lim_{n \to \infty} \mathbb{P}(X_n = i) = a_i, \qquad \text{for each } i \in \mathcal{S}.$$

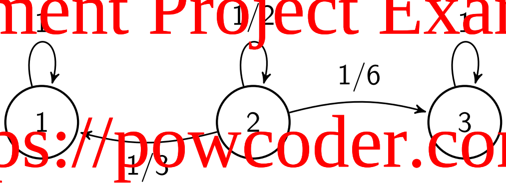
In general,

- A limiting distribution need not exist
- The limiting distribution (if it exists) is unique (by definition)
- The limiting distribution (if it exists) depends on both the initial distribution and P.

We have already seen that if $\boldsymbol{\pi}$ is a stationary distribution, and $\mathbb{P}(X_0 = i) = \pi_i$ for each $i \in \mathcal{S}$ then $\mathbb{P}(X_n = i) = \pi_i$ for each $i \in \mathcal{S}, n \in \mathbb{N}$, so in this case $\boldsymbol{\pi}$ is also the limiting distribution.

Examples:

Consider a DTMC with initial distribution (b_1, b_2, b_3) and transition diagram



Find the limiting distribution for the chain.

Note that if $b_1 = 1$ then the chain stays forever in state 1 so the limiting distribution is $(1, 0, 0)$ in this case. Similarly it is $(0, 0, 1)$ if $b_3 = 1$. If $b_2 = 1$ then we either hit state 1 (with probability $h_{2,3} = 2/3$) or 3 and then stay there, so the limiting distribution is $(2/3, 0, 1/3)$ in this case (see the next slide for details).

Examples:

To answer the question in general, note that if $X_k = 1$ then $X_n = 1$ for all $n \geq k$ and therefore $\{X_n = 1\} = \cup_{k=0}^{n}\{X_k = 1\}$. Thus,

$$\lim_{n\to\infty} \mathbb{P}(X_n = 1) = \lim_{n\to\infty} \mathbb{P}(\cup_{k=0}^{n}\{X_k = 1\}) = \mathbb{P}(\cup_{k=0}^{\infty}\{X_k = 1\}).$$

The last event $H_1 = \cup_{k=0}^{\infty}\{X_k = 1\}$ is the event that we ever reach state 1, so

$$\begin{aligned}
\mathbb{P}(H_1) &= \mathbb{P}(H_1 | X_0 = 1)\mathbb{P}(X_0 = 1) + \mathbb{P}(H_1 | X_0 = 2)\mathbb{P}(X_0 = 2) \\
&\quad + \mathbb{P}(H_1 | X_0 = 3)\mathbb{P}(X_0 = 3) \\
&= 1 \cdot \mathbb{P}(X_0 = 1) + \frac{2}{3} \cdot \mathbb{P}(X_0 = 2) + 0 \qquad (3) \\
&= \mathbb{P}(X_0 = 1) + \frac{2}{3} \cdot \mathbb{P}(X_0 = 2).
\end{aligned}$$

Example:

Consider a DTMC with $\mathcal{S} = \{1, 2\}$, $\mathbb{P}(X_0 = 1) = p$ and transition matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Find the limiting distribution (if it exists).

This chain is periodic (with period 2). Note that $\mathbb{P}(X_{2m} = 1) = \mathbb{P}(X_0 = 1) = p$ and $\mathbb{P}(X_{2m+1} = 1) = \mathbb{P}(X_0 = 2) = 1 - p$. Thus

$$\mathbb{P}(X_n = 1) = \begin{cases} p, & \text{if } n \text{ is even} \\ 1 - p, & \text{if } n \text{ is odd.} \end{cases}$$

This converges if and only if $p = 1 - p$.

In other words, for this example, a limiting distribution exists if and only if $p = 1/2$.

Example: simple random walk

Recall that $S_0 = 0$ and $p_{i,i+1} = p$ and $p_{i,i-1} = 1 - p$.

If $p > 1/2$ then $S_n \to \infty$ so $\mathbb{P}(S_n = i) \to 0$ as $n \to \infty$

If $p = 1/2$ the chain is recurrent, but we have seen (using Stirling's formula) that

$$\mathbb{P}(S_{2n} = 0) \approx \frac{C}{\sqrt{n}} \to 0,$$

while $\mathbb{P}(S_{2n+1} = 0) = 0$. So $\mathbb{P}(S_n = 0) \to 0$ as $n \to \infty$

Similarly one can show that $\mathbb{P}(S_n = i) \to 0$ for each i, so all of the limits exist, but they are all 0, so there is no limiting distribution in this example.

Limiting distribution results

Theorem: For a (time-homogeneous) DTMC, if a limiting distribution exists then it is a stationary distribution.

Theorem:$(***)$ Let $(X_n)_{n\in\mathbb{Z}_+}$ be an irreducible, aperiodic (time-homogeneous) DTMC with countable state space $\mathcal{S}$. Then for all $i, j \in \mathcal{S}$,

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_0 = i) \to \frac{1}{\mu_j}, \quad \text{as } n \to \infty.$$

Ergodicity

We say that a DTMC is ergodic if the limiting distribution exists and does not depend on the starting distribution. This is equivalent to saying that $a_j = \lim_{n\to\infty} p_{i,j}^{(n)}$ exists for each $i, j \in S$, does not depend on i, and $\sum_{j\in S} a_j = 1$.

Theorem: An irreducible (time-homogeneous) DTMC is ergodic if and only if it is aperiodic, and positive recurrent. For an ergodic DTMC, the limiting distribution is equal to the stationary distribution.

Doubly stochastic P:

An $m \times m$ stochastic matrix P is called doubly-stochastic if all the column sums are equal to one.

Suppose that a finite-state DTMC has doubly stochastic transition matrix.

Suppose that $\mathcal{S}$ contains exactly m elements. Then (exercise)

$$(1/m, 1/m, \ldots, 1/m)P = (1/m, 1/m, \ldots, 1/m).$$

It follows that the uniform distribution

$$\boldsymbol{\pi} = (1/m, 1/m, \ldots, 1/m),$$

is a stationary distribution.

Conversely if the uniform distribution on $\mathcal{S} = \{1, \ldots, m\}$ is a stationary distribution for P then for each i,

$$\frac{1}{m} = \sum_{j \in \mathcal{S}} \frac{1}{m} p_{j,i} = \frac{1}{m} \sum_{j \in \mathcal{S}} p_{j,i},$$

so P is doubly stochastic.

Example:

Let $(X_n)_{n\in\mathbb{Z}_+}$ be a DTMC with state space $\mathcal{S} = \{1, 2, 3\}$, and transition matrix

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

If the initial distribution is $x = (x_1, x_2, x_3)$, find:

(i) the limiting distribution

(ii) the limiting proportion of time spent in state 2

Reversibility

An irreducible DTMC is called reversible if there exists a probability distribution $\boldsymbol{\pi} = (\pi_i)_{i \in \mathcal{S}}$ such that

$$\pi_i p_{i,j} = \pi_j p_{j,i}, \quad i, j \in \mathcal{S}.$$

The above equations are called the detailed balance equations.

Any solution to these equations is a stationary distribution since if $\boldsymbol{\pi}$ satisfies the detailed balance equations then

$$\sum_{j \in \mathcal{S}} \pi_j p_{j,i} = \sum_{j \in \mathcal{S}} \pi_i p_{i,j} = \pi_i \left(\sum_{j \in \mathcal{S}} p_{i,j} \right) = \pi_i.$$

Kolmogorov's reversibility criterion

A (time-homogeneous) irreducible DTMC with state space $\mathcal{S}$ is reversible if and only if it has a stationary distribution and P satisfies

$$p_{j_n j_1} \prod_{i=1}^{n-1} p_{j_i j_{i+1}} = p_{j_1 j_n} \prod_{i=1}^{n-1} p_{j_{i+1} j_i},$$

for every n and every $\{j_1, j_2, \ldots, j_n\}$

Interpretation: Suppose I show you a short video clip of a Markov chain that starts and ends at the same state, and I choose to show it either forwards or in reverse. I tell you what P is. Then what you observe contains no information about whether I showed the process forwards or in reverse *if and only if the process is reversible*.

Example

Let P be an irreducible (so it has a stationary measure) stochastic matrix with $p_{i,j} = 0$ unless $j \in \{i-1, i, i+1\}$. (A Markov chain with such a stochastic matrix is called a *birth and death chain*.)

Then for any sequence of n transitions taking us from state i to state i: every occurrence of a transition $j \to j+1$ has a corresponding transition $j+1$ to j and vice versa.

Thus, reversing this sequence of n transitions we see exactly the same number of transitions from j to $j+1$ as the forward sequence. This is true for all j and therefore the Kolmogorov criterion is always satisfied for such a stochastic matrix, provided it also has a stationary distribution.

Thus a positive recurrent birth and death chain is always reversible.

Example

Let $(X_n)_{n\in\mathbb{Z}_+}$ be a DTMC with $\mathcal{S} = \mathbb{Z}_+$ and transition probabilities $p_{i,i+1} = p_i \in (0,1)$ for each $i \in \mathcal{S}$, and $p_{0,0} = 1 - p_{0,1}$ and $p_{i,i-1} = 1 - p_i$ for $i \geq 1$.

This is a birth and death chain, so if it has a stationary distribution then it satisfies the detailed balance equations:

$$\pi_i p_i = \pi_{i+1}(1 - p_{i+1}), \qquad \text{i.e.}$$

$$\pi_{i+1} = p_i/(1 - p_{i+1})\pi_i.$$

Letting $\rho_i = p_i/(1 - p_{i+1})$, it follows that

$$\pi_i = x \prod_{j=0}^{i-1} \rho_j,$$

gives a solution to the detailed balance equations.

Thus if $\sum_{i=0}^{\infty} \left(\prod_{j=0}^{i-1} \rho_j \right) < \infty$ then there is a stationary distribution (otherwise not).

Example: Random walk with one barrier

Continuing the example above, if $p_i = p$, then $\rho_i = \rho := p/(1-p)$ for every i, and the sum is finite if and only if $\rho < 1$ (i.e. $p < 1/2$).

When $p < 1/2$ we have a stationary distribution, and the chain is irreducible and aperiodic (due to the boundary), so also ergodic. Otherwise we have no stationary distribution.

The limiting proportion of time spent in state 0 is π_0. We find this by

$$1 = \sum_{i=0}^{\infty} \pi_i = \sum_{i=0}^{\infty} \pi_0 \rho^i = \pi_0/(1-\rho).$$

In other words $\pi_0 = 1 - \rho$.

We will see something similar when we study simple queues later in the course.

Summary:

We have introduced the Markov property, and notions of communicating classes (and irreducibility). We have seen how to calculate path probabilities, hitting probabilities and expected hitting times. We have discussed transience, positive-recurrence and null-recurrence, and periodicity.

For an irreducible, aperiodic and positive-recurrent DTMC, the stationary distribution π has a number of interpretations. It can be seen as:

- the initial distribution for which the process is a stationary process
- the limiting probability of being in each state
- the limiting proportion of time spent in each state

Tricks of the trade

Sometimes we have a process that is not a Markov chain, yet we can still use Markov chain theory to analyse its behaviour by being clever.

Consider the following example:

Let $(S'_t)_{t\in\mathbb{Z}_+}$ denote a random process with state space $\mathcal{S}' = \{1,2,3,4\}$, $S'_0 = 1$ and $S'_1 = 2$, such that whenever the process is at i, having just come from j, it chooses uniformly at random from $\mathcal{S}' \setminus \{i,j\}$. This process is a kind of non-backtracking random walk (on the set $\mathcal{S}$).

Exercise: Show that S' is not Markovian by showing that $\mathbb{P}(S'_4 = 4 | S'_3 = 1) \neq \mathbb{P}(S'_4 = 4 | S'_3 = 1, S'_2 = 4)$

If we put $X_t = (S'_t, S'_{t+1})$ for $t \in \mathbb{Z}_+$ then $(X_t)_{t\in\mathbb{Z}_+}$ is a Markov chain on a state space $\mathcal{S} = \{(i,j) : i,j \in \mathcal{S}',\ i \neq j\}$ with 12 elements (can relabel them 1 to 12 if you wish), with

$$p_{(i_1,j_1),(i_2,j_2)} = \begin{cases} \frac{1}{2}, & \text{if } i_2 = j_1 \text{ and } j_2 \notin \{i_1, j_1\} \\ 0, & \text{otherwise.} \end{cases}$$