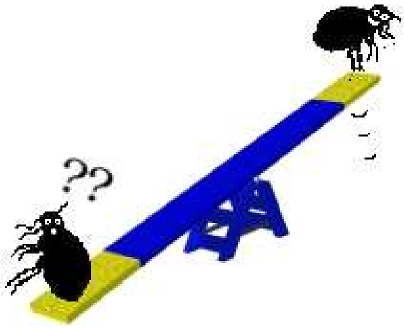
Chapter 9: Equilibrium



In Chapter 8, we saw that if $\{X_0, X_1, X_2, \ldots\}$ is a Markov chain with transition matrix P, then $X_t \sim \boldsymbol{\pi}^T \quad \Rightarrow \quad X_{t+1} \sim \boldsymbol{\pi}^T P$.

This raises the question: is there any distribution $\boldsymbol{\pi}$ such that $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$*?*

If $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, then

$$\begin{aligned} X_t \sim \boldsymbol{\pi}^T \quad &\Rightarrow \quad X_{t+1} \sim \boldsymbol{\pi}^T P = \boldsymbol{\pi}^T \\ &\Rightarrow \quad X_{t+2} \sim \boldsymbol{\pi}^T P = \boldsymbol{\pi}^T \\ &\Rightarrow \quad X_{t+3} \sim \boldsymbol{\pi}^T P = \boldsymbol{\pi}^T \\ &\Rightarrow \quad \ldots \end{aligned}$$

In other words, if $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, and $X_t \sim \boldsymbol{\pi}^T$, then

$$X_t \sim X_{t+1} \sim X_{t+2} \sim X_{t+3} \sim \ldots$$

Thus, once a Markov chain has reached a distribution $\boldsymbol{\pi}^T$ such that $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, *it will stay there.*

If $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, we say that the distribution $\boldsymbol{\pi}^T$ is an *equilibrium distribution.*

Equilibrium means a ***level position:*** there is *no more change* in the distribution of X_t as we wander through the Markov chain.

Note: Equilibrium does not mean that the **value** of X_{t+1} equals the **value** of X_t. It means that the **distribution** of X_{t+1} is the same as the **distribution** of X_t:

e.g. $\mathbb{P}(X_{t+1} = 1) = \mathbb{P}(X_t = 1) = \pi_1$*;*

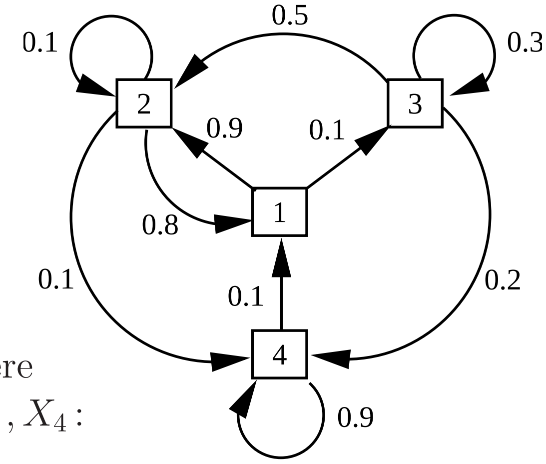
$\mathbb{P}(X_{t+1} = 2) = \mathbb{P}(X_t = 2) = \pi_2$, *etc.*

In this chapter, we will first see how to *calculate* the equilibrium distribution $\boldsymbol{\pi}^T$. We will then see the remarkable result that many Markov chains automatically *find their own way* to an equilibrium distribution as the chain wanders through time. This happens for many Markov chains, but not all. We will see the conditions required for the chain to find its way to an equilibrium distribution.

9.1 Equilibrium distribution in pictures

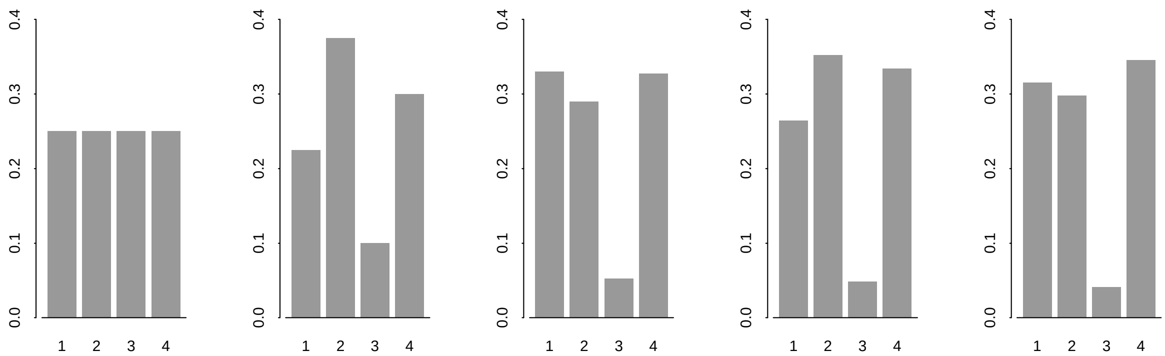
Consider the following 4-state Markov chain:

$$P = \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix}$$



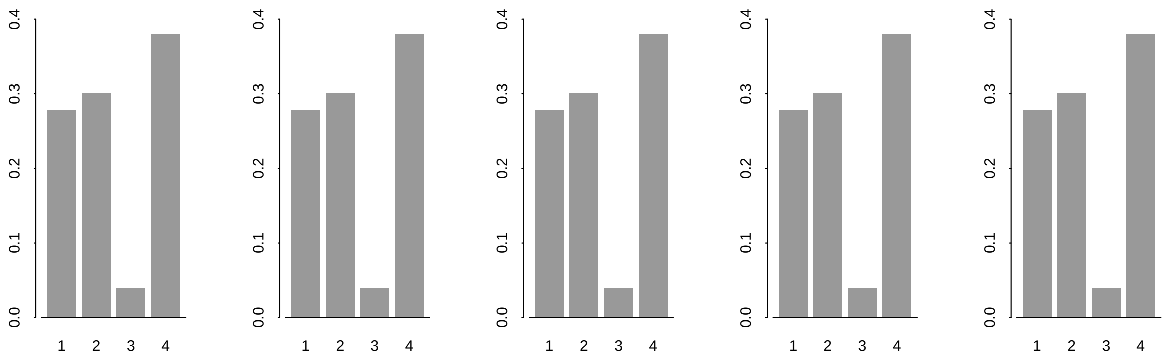
Suppose we start at time 0 with $X_0 \sim \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$: so the chain is equally likely to start from any of the four states. Here are pictures of the distributions of $X_0, X_1, \ldots, X_4$:

$\mathbb{P}(X_0 = x)$ $\mathbb{P}(X_1 = x)$ $\mathbb{P}(X_2 = x)$ $\mathbb{P}(X_3 = x)$ $\mathbb{P}(X_4 = x)$



The distribution starts off level, but quickly changes: for example the chain is least likely to be found in state 3. The distribution of X_t changes between each $t = 0, 1, 2, 3, 4$. Now look at the distribution of X_t 500 steps into the future:

$\mathbb{P}(X_{500} = x)$ $\mathbb{P}(X_{501} = x)$ $\mathbb{P}(X_{502} = x)$ $\mathbb{P}(X_{503} = x)$ $\mathbb{P}(X_{504} = x)$



The distribution has reached a steady state: it ***does not change*** between $t = 500, 501, \ldots, 504$. *The chain has reached equilibrium of its own accord.*

9.2 Calculating equilibrium distributions

Definition: Let $\{X_0, X_1, \ldots\}$ be a Markov chain with transition matrix P and state space S, where $|S| = N$ (possibly infinite). Let $\boldsymbol{\pi}^T$ be a row vector denoting a probability distribution on S: so each element π_i denotes the probability of being in state i, and $\sum_{i=1}^{N} \pi_i = 1$, where $\pi_i \geq 0$ for all $i = 1, \ldots, N$. The probability distribution $\boldsymbol{\pi}^T$ is an **equilibrium** distribution for the Markov chain if $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$.

That is, $\boldsymbol{\pi}^T$ is an equilibrium distribution if

$$\left(\boldsymbol{\pi}^T P\right)_j = \sum_{i=1}^{N} \pi_i p_{ij} = \pi_j \quad \textit{for all } j = 1, \ldots, N.$$

By the argument given on page 174, we have the following Theorem:

Theorem 9.2: Let $\{X_0, X_1, \ldots\}$ be a Markov chain with transition matrix P. Suppose that $\boldsymbol{\pi}^T$ is an equilibrium distribution for the chain. If $X_t \sim \boldsymbol{\pi}^T$ for any t, then $X_{t+r} \sim \boldsymbol{\pi}^T$ *for all* $r \geq 0$. □

Once a chain has hit an equilibrium distribution, *it stays there for ever.*

Note: There are several other names for an equilibrium distribution. If $\boldsymbol{\pi}^T$ is an equilibrium distribution, it is also called:

- **invariant:** *it doesn't change:* $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$*;*
- **stationary:** *the chain 'stops' here.*

Stationarity: the Chain Station

a BUS station is where a BUS stops

a train station is where a train stops

a **workstation** is where . . . ? ? ?

`a stationary distribution is where a Markov chain stops`

9.3 Finding an equilibrium distribution

Vector $\boldsymbol{\pi}^T$ is an equilibrium distribution for P if:

1. $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$;
2. $\sum_{i=1}^{N} \pi_i = 1$;
3. $\pi_i \geq 0$ *for all* i.

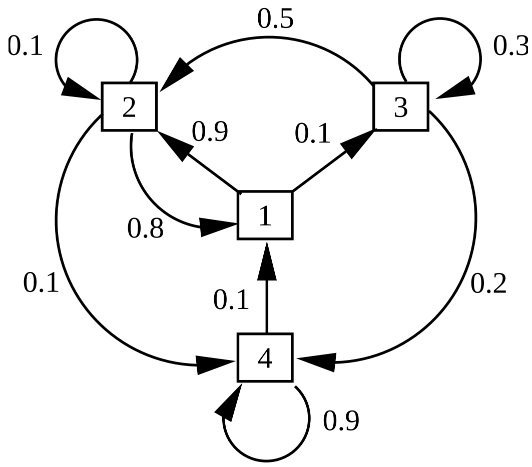
Conditions 2 and 3 ensure that $\boldsymbol{\pi}^T$ *is a genuine probability distribution.*

Condition 1 means that $\boldsymbol{\pi}$ is a row eigenvector of P.

Solving $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ by itself will just specify $\boldsymbol{\pi}$ up to *a scalar multiple.*
We need to include Condition 2 to scale $\boldsymbol{\pi}$ to a genuine probability distribution, and then check with Condition 3 that the scaled distribution is valid.

Example: Find an equilibrium distribution for the Markov chain below.

$$P = \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix}$$



Solution:

Let $\boldsymbol{\pi}^T = (\pi_1,\ \pi_2,\ \pi_3,\ \pi_4)$.
The equations are $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ *and* $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$.

$$\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T \quad \Rightarrow \quad (\pi_1\ \pi_2\ \pi_3\ \pi_4) \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix} = (\pi_1\ \pi_2\ \pi_3\ \pi_4)$$

$$\begin{aligned}
.8\pi_2 + .1\pi_4 &= \pi_1 && (1)\\
.9\pi_1 + .1\pi_2 + .5\pi_3 &= \pi_2 && (2)\\
.1\pi_1 + .3\pi_3 &= \pi_3 && (3)\\
.1\pi_2 + .2\pi_3 + .9\pi_4 &= \pi_4 && (4)\\
\textit{Also} \qquad \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1. && (5)
\end{aligned}$$

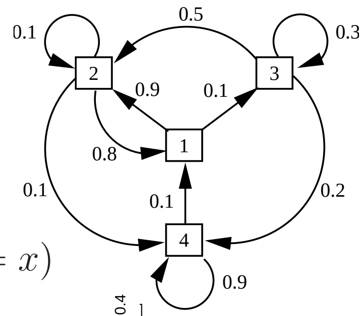
$$\begin{aligned}
(3) \quad \Rightarrow \quad \pi_1 &= 7\pi_3\\
\textit{Substitute in (2)} \quad \Rightarrow \quad .9(7\pi_3) + .5\pi_3 &= .9\pi_2\\
\Rightarrow \quad \pi_2 &= \frac{68}{9}\pi_3\\
\textit{Substitute in (1)} \quad \Rightarrow \quad .8\left(\frac{68}{9}\pi_3\right) + .1\pi_4 &= 7\pi_3\\
\Rightarrow \quad \pi_4 &= \frac{86}{9}\pi_3\\
\textit{Substitute all in (5)} \quad \Rightarrow \quad \pi_3\left(7 + \frac{68}{9} + 1 + \frac{86}{9}\right) &= 1\\
\Rightarrow \quad \pi_3 &= \frac{9}{226}
\end{aligned}$$

Overall:

$$\begin{aligned}
\boldsymbol{\pi}^T &= \left(\frac{63}{226},\ \frac{68}{226},\ \frac{9}{226},\ \frac{86}{226}\right)\\
&= (0.28,\ 0.30,\ 0.04,\ 0.38).
\end{aligned}$$

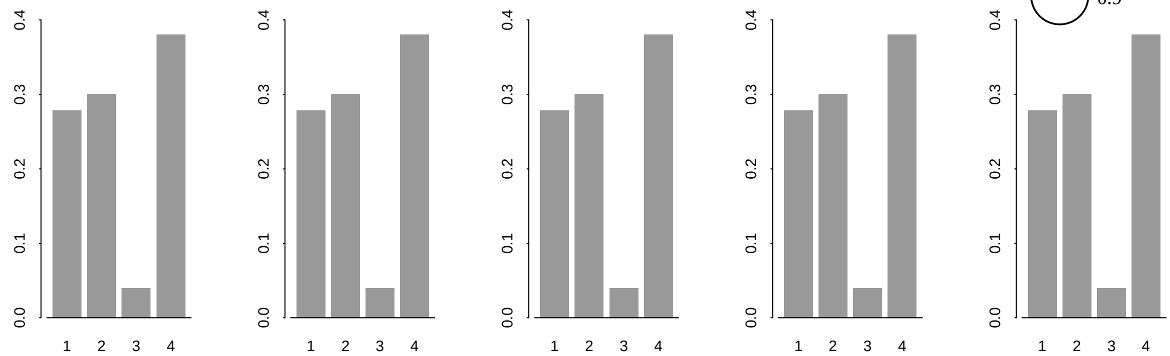
This is the distribution the chain converged to in Section 9.1.

9.4 Long-term behaviour



In Section 9.1, we saw an example where the Markov chain wandered of its own accord into its equilibrium distribution:

$\mathbb{P}(X_{500} = x)$ $\mathbb{P}(X_{501} = x)$ $\mathbb{P}(X_{502} = x)$ $\mathbb{P}(X_{503} = x)$



This will always happen for this Markov chain. In fact, the distribution it converges to (found above) does not depend upon the starting conditions: *for ANY value of X_0, we will always have* $X_t \sim (0.28,\ 0.30,\ 0.04,\ 0.38)$ *as* $t \to \infty$.

What is happening here is that *each row of the transition matrix P^t converges to the equilibrium distribution* $(0.28,\ 0.30,\ 0.04,\ 0.38)$ *as* $t \to \infty$*:*

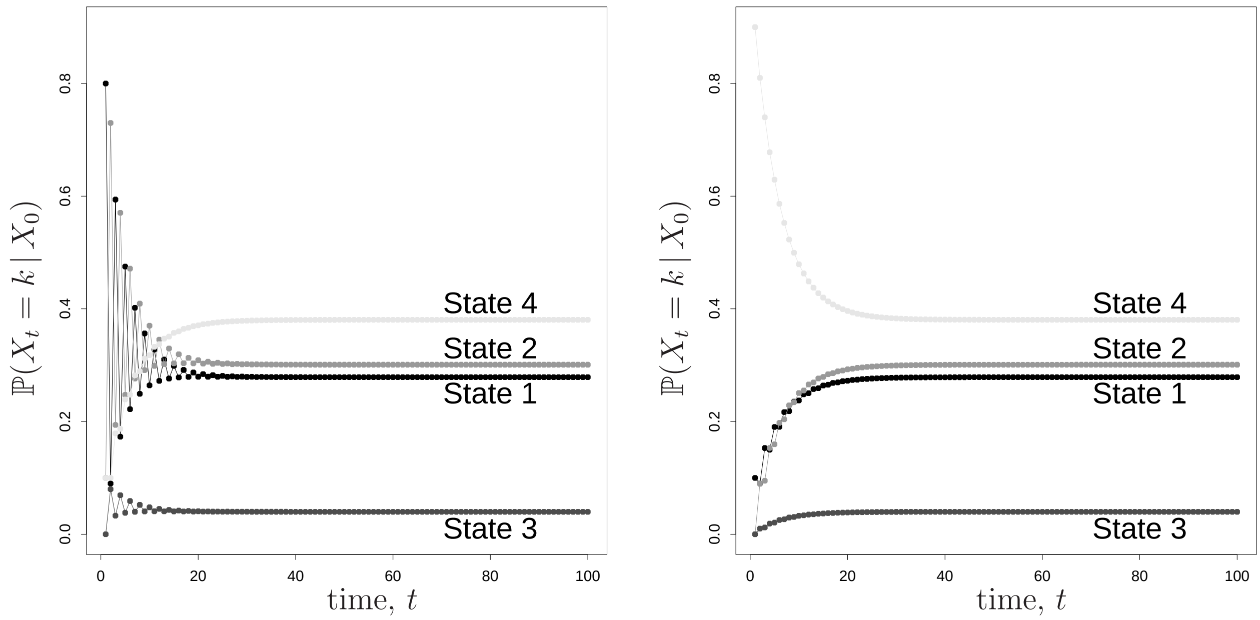
$$P = \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix} \quad \Rightarrow \quad P^t \to \begin{pmatrix} 0.28 & 0.30 & 0.04 & 0.38 \\ 0.28 & 0.30 & 0.04 & 0.38 \\ 0.28 & 0.30 & 0.04 & 0.38 \\ 0.28 & 0.30 & 0.04 & 0.38 \end{pmatrix} \quad \text{as } t \to \infty.$$

(If you have a calculator that can handle matrices, try finding P^t for $t = 20$ and $t = 30$: you will find the matrix is already converging as above.)

This convergence of P^t means that *for large t, no matter WHICH state we start in, we always have probability*

- about *0.28* of being in State *1* after t steps;
- about *0.30* of being in State *2* after t steps;
- about *0.04* of being in State *3* after t steps;
- about *0.38* of being in State *4* after t steps.

Start at $X_0 = 2$ **Start at $X_0 = 4$**



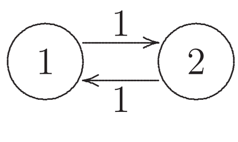
The **left graph** shows the probability of getting from state 2 to state k in t steps, as t changes: $(P^t)_{2,k}$ for $k = 1, 2, 3, 4$.

The **right graph** shows the probability of getting from state 4 to state k in t steps, as t changes: $(P^t)_{4,k}$ for $k = 1, 2, 3, 4$.

The ***initial behaviour*** differs greatly for the different start states.

The ***long-term behaviour*** (large t) is the same for both start states.

However, this does not always happen. Consider the two-state chain below:



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

As t gets large, P^t *does not converge:*

$$P^{500} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad P^{501} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad P^{502} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad P^{503} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ldots$$

For this Markov chain, *we never 'forget' the initial start state.*

General formula for P^t

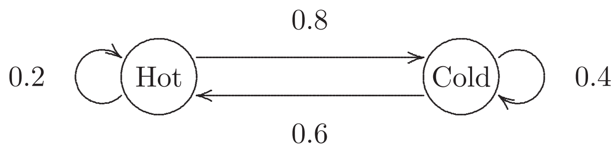
We have seen that we are interested in whether P^t converges to *a fixed matrix with all rows equal as $t \to \infty$.*

If it does, then the Markov chain will *reach an equilibrium distribution that does not depend upon the starting conditions.*

The equilibrium distribution is then given by *any row of the converged P^t.*

It can be shown that a general formula is available for P^t for any t, based on the eigenvalues of P. Producing this formula is beyond the scope of this course, but if you are given the formula, you should be able to recognise whether P^t is going to converge to a fixed matrix with all rows the same.

Example 1:



$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix}$$

We can show that the general solution for P^t is:

$$P^t = \frac{1}{7}\left\{\begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 4 & -4 \\ -3 & 3 \end{pmatrix}(-0.4)^t\right\}$$

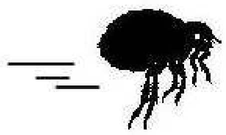
As $t \to \infty$, $(-0.4)^t \to 0$, so

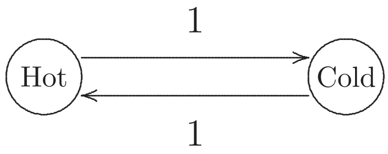
$$P^t \to \frac{1}{7}\begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{7} & \frac{4}{7} \\ \frac{3}{7} & \frac{4}{7} \end{pmatrix}$$

This Markov chain will therefore converge to the equilibrium distribution $\boldsymbol{\pi}^T = \left(\frac{3}{7}, \frac{4}{7}\right)$ as $t \to \infty$, regardless of whether the flea starts in state 1 or state 2.

Exercise: Verify that $\boldsymbol{\pi}^T = \left(\frac{3}{7}, \frac{4}{7}\right)$ is the same as the result you obtain from solving the equilibrium equations: $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ and $\pi_1 + \pi_2 = 1$.

Example 2: Purposeflea knows exactly what he is doing, so his probabilities are all 1:





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

We can show that the general solution for P^t is:

$$P^t = \frac{1}{2}\left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} (-1)^t \right\}$$

As $t \to \infty$, $(-1)^t$ does not converge to 0, so

$$P^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \textit{if } t \textit{ is even},$$

$$P^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \textit{if } t \textit{ is odd},$$

for all t.

In this example, P^t never converges to a matrix with both rows identical as t gets large. The chain never 'forgets' its starting conditions as $t \to \infty$.

Exercise: Verify that this Markov chain *does* have an equilibrium distribution, $\boldsymbol{\pi}^T = \left(\frac{1}{2}, \frac{1}{2}\right)$. However, the chain does not *converge* to this distribution as $t \to \infty$.

These examples show that some Markov chains forget their starting conditions in the long term, and ensure that X_t will have the same distribution as $t \to \infty$ regardless of where we started at X_0. However, for other Markov chains, the initial conditions are never forgotten. In the next sections we look for general criteria that will ensure the chain converges.

Target Result:

- If a Markov chain is ***irreducible*** and ***aperiodic***, and if an equilibrium distribution $\boldsymbol{\pi}^T$ ***exists***, then the chain converges to this distribution as $t \to \infty$, regardless of the initial starting states.

To make sense of this, we need to revise the concept of ***irreducibility***, and introduce the idea of ***aperiodicity.***

9.5 Irreducibility

Recall from Chapter 8:

Definition: A Markov chain or transition matrix P is said to be **irreducible** if $i \leftrightarrow j$ *for all* $i, j \in S$. *That is, the chain is irreducible if the state space* S *is a single communicating class.*

An irreducible Markov chain consists of a single class.

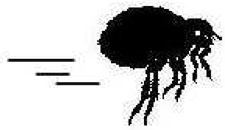
Irreducible *Not irreducible*

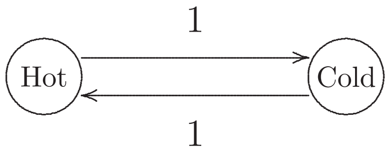
Irreducibility of a Markov chain is important for convergence to equilibrium as $t \to \infty$, because *we want the convergence to be independent of start state.*

This can happen if the chain is irreducible. When the chain is not irreducible, different start states might cause the chain to get stuck in different closed classes. In the example above, a start state of $X_0 = 1$ means that the chain is restricted to states 1 and 2 as $t \to \infty$, whereas a start state of $X_0 = 4$ means that the chain is restricted to states 4 and 5 as $t \to \infty$. A single convergence that 'forgets' the initial state is therefore not possible.

9.6 Periodicity

Consider the Markov chain with transition matrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.





Suppose that $X_0 = 1$.

Then $X_t = 1$ *for all even values of t, and $X_t = 2$ for all odd values of t.*

This sort of behaviour is called **periodicity:** *the Markov chain can only return to a state at particular values of t.*

Clearly, periodicity of the chain will interfere with convergence to an equilibrium distribution as $t \to \infty$. For example,

$$\mathbb{P}(X_t = 1 \,|\, X_0 = 1) = \begin{cases} 1 \text{ for even values of } t, \\ 0 \text{ for odd values of } t. \end{cases}$$

Therefore, the probability can not converge to any single value as $t \to \infty$.

Period of state i

To formalize the notion of periodicity, we define the **period** of a state i. Intuitively, *the period is defined so that the time taken to get from state i back to state i again is always a multiple of the period.*

In the example above, the chain can return to state 1 after *2 steps, 4 steps, 6 steps, 8 steps, ...*

The period of state 1 is therefore *2.*

In general, the chain can return from state i back to state i again in t steps if $(P^t)_{ii} > 0$. This prompts the following definition.

Definition: The **period** $d(i)$ of a state i is

$$d(i) = \gcd\{t \,:\, \left(P^t\right)_{ii} > 0\},$$

the greatest common divisor of the times at which return is possible.

Definition: The state i is said to be **periodic** if $d(i) > 1$.

> For a periodic state i, $(P^t)_{ii} = 0$ if t is not a multiple of $d(i)$.

Definition: The state i is said to be **aperiodic** if $d(i) = 1$.

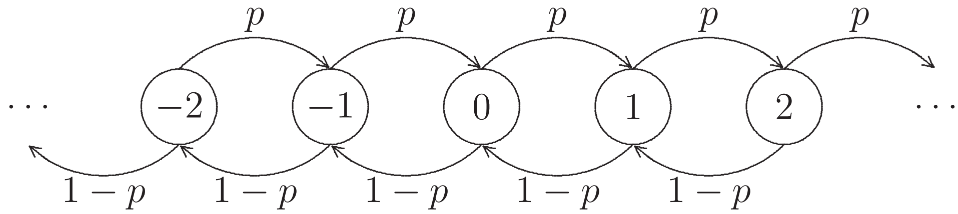
If state i is aperiodic, it means that *return to state i is not limited only to regularly repeating times.*

For convergence to equilibrium as $t \to \infty$, we will be interested only in *aperiodic states.*

The following examples show how to calculate the period for both aperiodic and periodic states.

Examples: Find the periods of the given states in the following Markov chains, and state whether or not the chain is irreducible.

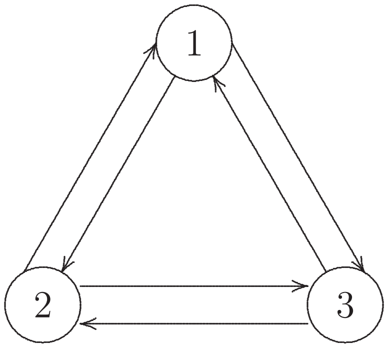
1. The simple random walk.



$$d(0) = \gcd\{2, 4, 6, \ldots\} = 2.$$

Chain is irreducible.

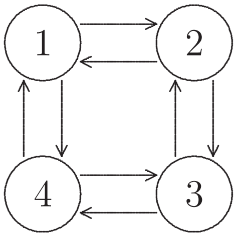
2.



$$d(1) = \gcd\{2, 3, 4, \ldots\} = 1.$$

Chain is irreducible.

3.



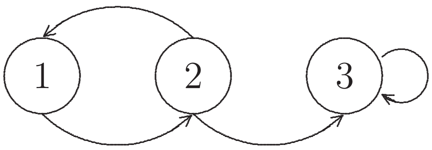
$$d(1) = \gcd\{2, 4, 6, \ldots\} = 2.$$

Chain is irreducible.

4.

$$d(1) = \gcd\{2, 4, 6, \ldots\} = 2.$$

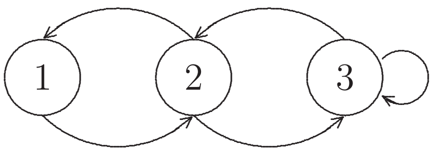
Chain is NOT irreducible (i.e. Reducible).



5.

$$d(1) = \gcd\{2, 4, 5, 6, \ldots\} = 1.$$

Chain is irreducible.



9.7 Convergence to Equilibrium

We now draw together the threads of the previous sections with the following results.

Fact: If $i \leftrightarrow j$, then i and j have the same period. (Proof omitted.)

This leads immediately to the following result:

> If a Markov chain is ***irreducible*** and has ***one*** aperiodic state, then ***all*** states are aperiodic.

We can therefore talk about an **irreducible, aperiodic chain,** *meaning that all states are aperiodic.*

Theorem 9.7: Let $\{X_0, X_1, \ldots\}$ be an **irreducible and aperiodic** Markov chain with transition matrix P. Suppose that there ***exists*** an equilibrium distribution $\boldsymbol{\pi}^T$. Then, from ***any*** starting state i, and for any end state j,

$$\mathbb{P}(X_t = j \mid X_0 = i) \to \pi_j \quad \text{as } t \to \infty.$$

In particular,

$$\left(P^t\right)_{ij} \to \pi_j \quad \text{as } t \to \infty, \text{ for all } i \text{ and } j,$$

so P^t converges to a matrix with all rows identical and equal to $\boldsymbol{\pi}^T$. □

> *For an irreducible, aperiodic Markov chain,*
>
> *with finite or infinite state space,*
>
> *the existence of an equilibrium distribution $\boldsymbol{\pi}^T$ ensures*
>
> *that the Markov chain will converge to $\boldsymbol{\pi}^T$ as $t \to \infty$.*

Note: If the state space is infinite, it is not guaranteed that an equilibrium distribution $\boldsymbol{\pi}^T$ exists. See Example 3 below.

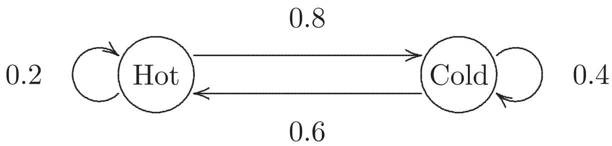
Note: If the chain converges to an equilibrium distribution $\boldsymbol{\pi}^T$ as $t \to \infty$, then *the long-run proportion of time spent in state k is π_k.*

9.8 Examples

A typical exam question gives you a Markov chain on a finite state space and asks if it converges to an equilibrium distribution as $t \to \infty$. An equilibrium distribution will always exist for a finite state space. You need to check whether the chain is irreducible and aperiodic. If so, it will converge to equilibrium. If the chain is irreducible but *periodic*, it cannot converge to an equilibrium distribution that is independent of start state. If the chain is *reducible*, it may or may not converge.

The first two examples are the same as the ones given in Section 9.4.

Example 1: State whether the Markov chain below converges to an equilibrium distribution as $t \to \infty$.

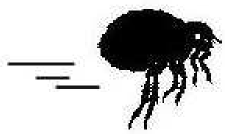


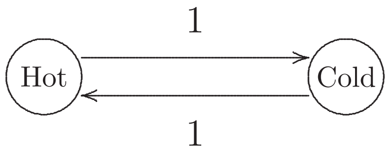
$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix}$$

The chain is irreducible and aperiodic, and an equilibrium distribution will exist for a finite state space. So the chain does converge.

(From Section 9.4, the chain converges to $\boldsymbol{\pi}^T = \left(\frac{3}{7}, \frac{4}{7}\right)$ as $t \to \infty$.)

Example 2: State whether the Markov chain below converges to an equilibrium distribution as $t \to \infty$.





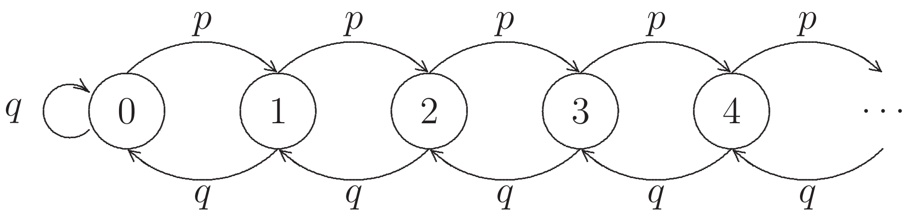
$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The chain is irreducible, but it is NOT aperiodic: period = 2.

Thus the chain does NOT converge to an equilibrium distribution as $t \to \infty$*.*

It is important to check for aperiodicity, because the existence of an equilibrium distribution does NOT ensure convergence to this distribution if the matrix is not aperiodic.

Example 3: Random walk with retaining barrier at 0.



Find whether the chain converges to equilibrium as $t \to \infty$, and if so, find the equilibrium distribution.

The chain is irreducible and aperiodic, so <u>if</u> an equilibrium distribution exists, then the chain will converge to this distribution as $t \to \infty$*.*

However, the chain has an infinite state space, so we cannot guarantee that an equilibrium distribution exists.

Try to solve the equilibrium equations:

$\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ *and* $\sum_{i=0}^{\infty} \pi_i = 1$.

$$P = \begin{pmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ 0 & q & 0 & p & \dots \\ \vdots & & & & \end{pmatrix}$$

$$\begin{aligned} q\pi_0 + q\pi_1 &= \pi_0 \qquad (\star) \\ p\pi_0 + q\pi_2 &= \pi_1 \\ p\pi_1 + q\pi_3 &= \pi_2 \\ &\vdots \\ p\pi_{k-1} + q\pi_{k+1} &= \pi_k \quad \text{for} \quad k = 1, 2, \dots \end{aligned}$$

From ($\star$), *we have* $p\pi_0 = q\pi_1$,

$$\text{so} \quad \pi_1 = \frac{p}{q}\pi_0$$

$$\Rightarrow \pi_2 = \frac{1}{q}(\pi_1 - p\pi_0) = \frac{1}{q}\left(\frac{p}{q}\pi_0 - p\pi_0\right) = \frac{p}{q}\left(\frac{1-q}{q}\right)\pi_0 = \left(\frac{p}{q}\right)^2 \pi_0.$$

We suspect that $\pi_k = \left(\frac{p}{q}\right)^k \pi_0$. *Prove by induction.*

The hypothesis is true for $k = 0, 1, 2$. *Suppose that* $\pi_k = \left(\frac{p}{q}\right)^k \pi_0$. *Then*

$$\begin{aligned} \pi_{k+1} &= \frac{1}{q}(\pi_k - p\pi_{k-1}) \\ &= \frac{1}{q}\left\{\left(\frac{p}{q}\right)^k \pi_0 - p\left(\frac{p}{q}\right)^{k-1} \pi_0\right\} \\ &= \frac{p^k}{q^k}\left(\frac{1}{q} - 1\right)\pi_0 \\ &= \left(\frac{p}{q}\right)^{k+1} \pi_0. \end{aligned}$$

The inductive hypothesis holds, so $\pi_k = \left(\frac{p}{q}\right)^k \pi_0$ *for all* $k \geq 0$.

We now need $\displaystyle\sum_{i=0}^{\infty} \pi_i = 1,$ *i.e.* $\displaystyle\pi_0 \sum_{k=0}^{\infty} \left(\frac{p}{q}\right)^k = 1.$

The sum is a Geometric series, and converges only for $\left|\frac{p}{q}\right| < 1$*. Thus when* $p < q$*, we have*

$$\pi_0 \left(\frac{1}{1-\frac{p}{q}}\right) = 1 \quad \Rightarrow \quad \pi_0 = 1 - \frac{p}{q}.$$

If $p \geq q$*, there is no equilibrium distribution.*

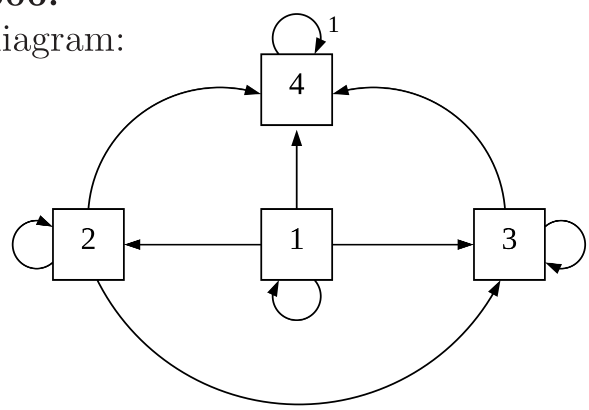
Solution:

If $p < q$*, the chain converges to an equilibrium distribution* $\boldsymbol{\pi}$*, where* $\pi_k = \left(1-\frac{p}{q}\right)\left(\frac{p}{q}\right)^k$ *for* $k = 0, 1, \ldots$*.*

If $p \geq q$*, the chain does not converge to an equilibrium distribution as* $t \to \infty$*.*

Example 4: **Sketch of Exam Question 2006.**

Consider a Markov chain with transition diagram:



(a) Identify all communicating classes. For each class, state whether or not it is closed.

Classes are:
$\{1\}, \{2\}, \{3\}$ *(each not closed);*
$\{4\}$ *(closed).*

(b) State whether the Markov chain is irreducible, and whether or not all states are aperiodic.

Not irreducible: there are 4 classes.
All states are aperiodic.

(c) The equilibrium distribution is $\boldsymbol{\pi}^T = (0, 0, 0, 1)$. Does the Markov chain converge to this distribution as $t \to \infty$, regardless of its start state?

Yes, it clearly will converge to $\boldsymbol{\pi}^T = (0, 0, 0, 1)$*, despite failure of irreducibility.*

Note: Equilibrium results also exist for chains that are ***not*** aperiodic. Also, states can be classified as **transient** (return to the state is not certain), **null recurrent** (return to the state is certain, but the expected return time is infinite), and **positive recurrent** (return to the state is certain, and the expected return time is finite). For each type of state, the long-term behaviour is known:

- If the state k is **transient** or **null-recurrent**,

$$\mathbb{P}(X_t = k \mid X_0 = k) = \left(P^t\right)_{kk} \to 0 \text{ as } t \to \infty.$$

- If the state is **positive recurrent**, then

$$\mathbb{P}(X_t = k \mid X_0 = k) = \left(P^t\right)_{kk} \to \pi_k \text{ as } t \to \infty, \text{ where } \pi_k > 0.$$

  The expected return time for the state is $1/\pi_k$.

A detailed treatment is available at
`http://www.statslab.cam.ac.uk/~james/Markov/`.

9.9 Special Process: the Two-Armed Bandit

A well-known problem in probability is called the **two-armed bandit** problem. The name is a reference to a type of gambling machine called the two-armed bandit. The two arms of the two-armed bandit offer different rewards, and the gambler has to decide which arm to play without knowing which is the better arm.

One-armed bandit

A similar problem arises when doctors are experimenting with two different treatments, without knowing which one is better. Call the treatments A and B. One of them is likely to be better, but we don't know which one. A series of patients will each be given one of the treatments. We aim to find a strategy that ensures that as many as possible of the patients are given the *better* treatment — though we don't know which one this is.

Suppose that, for any patient, treatment A has $\mathbb{P}(\text{success}) = \alpha$, and treatment B has $\mathbb{P}(\text{success}) = \beta$, and all patients are independent. Assume that $0 < \alpha < 1$ and $0 < \beta < 1$.

First let's look at a simple strategy the doctors might use:

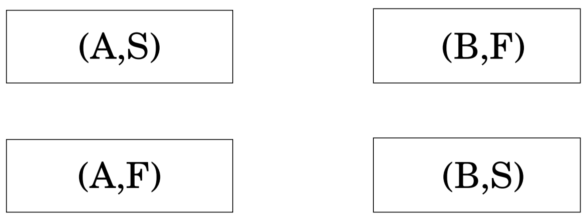
- The **random strategy** for allocating patients to treatments A and B is to choose from the two treatments at random, each with probability 0.5, for each patient.
- Let p_R be the overall probability of **success** for each patient with the random strategy. Show that $p_R = \frac{1}{2}(\alpha + \beta)$.

The **two-armed bandit strategy** is more clever. For the first patient, we choose treatment A or B at random (probability 0.5 each). If patient n is given treatment A and it is *successful*, then we use treatment A again for patient $n+1$, for all $n = 1, 2, 3, \ldots$. If A is a failure for patient n, we switch to treatment B for patient $n + 1$. A similar rule is applied if patient n is given treatment B: if it is successful, we keep B for patient $n+1$; if it fails, we switch to A for patient $n + 1$.

Define the **two-armed bandit process** to be a Markov chain with state space $\{(A, S), (A, F), (B, S), (B, F)\}$, where (A, S) means that patient n is given treatment A and it is successful, and so on.

Transition diagram:

Exercise: Draw on the missing arrows and find their probabilities in terms of α and β.



Transition matrix:

$$\begin{array}{c} \\ AS \\ AF \\ BS \\ BF \end{array} \begin{array}{c} \begin{array}{cccc} AS & AF & BS & BF \end{array} \\ \left(\begin{array}{cccc} \quad & \quad & \quad & \quad \\ & & & \\ & & & \\ & & & \end{array} \right) \end{array}$$

Probability of success under the two-armed bandit strategy

Define p_T to be the long-run probability of **success** using the two-armed bandit strategy.

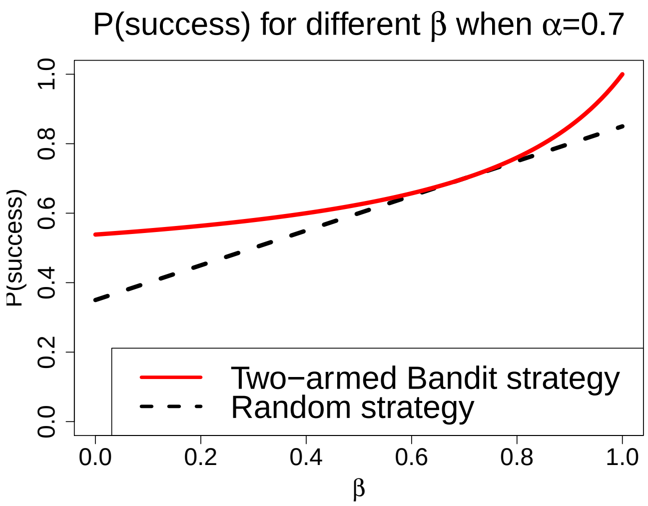
Exercise: Find the equilibrium distribution $\boldsymbol{\pi}$ for the two-armed bandit process. Hence show that the long-run probability of success for each patient under this strategy is:

$$p_T = \frac{\alpha + \beta - 2\alpha\beta}{2 - \alpha - \beta}.$$

Which strategy is better?

Exercise: Prove that $p_T - p_R \geq 0$ always, regardless of the values of α and β.

This proves that the two-armed bandit strategy is always better than, or equal to, the random strategy. It shows that we have been able to construct a strategy that gives all patients an increased chance of success, even though we don't know which treatment is better!



The graph shows the probability of success under the two different strategies, for $\alpha = 0.7$ and for $0 \leq \beta \leq 1$. Notice how $p_T \geq p_R$ for all possible values of β.