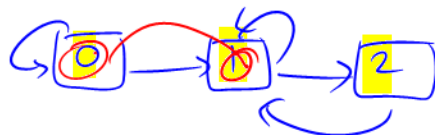


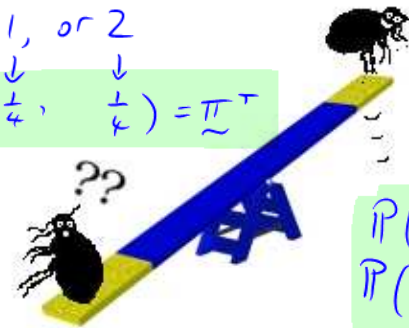
State space
 $= \{0, 1, 2\}$



$X_t = \text{box I'm in at step } t$

$X_t = 0, 1, \text{ or } 2$

$(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}) = \underline{\pi}^T$



$P(X_t=0) = \frac{1}{2}$
 $P(X_t=1) = \frac{1}{4}$
 $P(X_t=2) = \frac{1}{4}$

prob of being
in box 0, 1,
or 2 at time
 t .

Chapter 9: Equilibrium

In Chapter 8, we saw that if $\{X_0, X_1, X_2, \dots\}$ is a Markov chain with transition matrix P , then

$$X_t \sim \underline{\pi}^T \Rightarrow X_{t+1} \sim \underline{\pi}^T P$$

This raises the question: is there any distribution $\underline{\pi}$ such that

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

$$\underline{\pi}^T P = \underline{\pi}^T ?$$

$$\text{if } X_t \sim \underline{\pi}^T \Rightarrow X_{t+1} \sim \underline{\pi}^T P = \underline{\pi}^T$$

$$\Rightarrow X_{t+2} \sim \underline{\pi}^T P = \underline{\pi}^T$$

$$\Rightarrow X_{t+3} \sim \underline{\pi}^T P = \underline{\pi}^T \Rightarrow \dots \text{ for ever, } X_{t+n} \sim \underline{\pi}^T$$

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

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

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

If $\underline{\pi}^T P = \underline{\pi}^T$, we say that the distribution $\underline{\pi}^T$ 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 :

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

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

In this chapter, we will first see how to calculate the equilibrium distribution $\underline{\pi}^T$. 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.

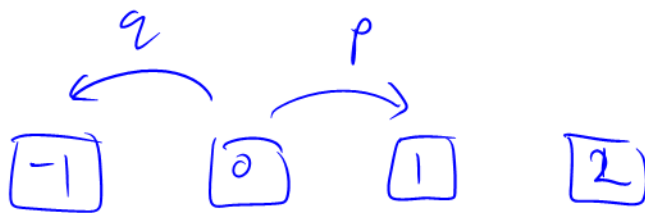
$\underline{\pi}$ = buried treasure.

buried treasure exists.

chain finds the treasure all by itself.

Q attempted.

1	
2	
3	
4	
5	
6	
7	



T = # steps to reach state 1

$$T = \begin{cases} 1 & \text{w.p. } p \\ 1 + T' + T'' & (q) \end{cases}$$

$\uparrow \quad \uparrow \quad \uparrow$
 $1 \quad 3 \quad 5$

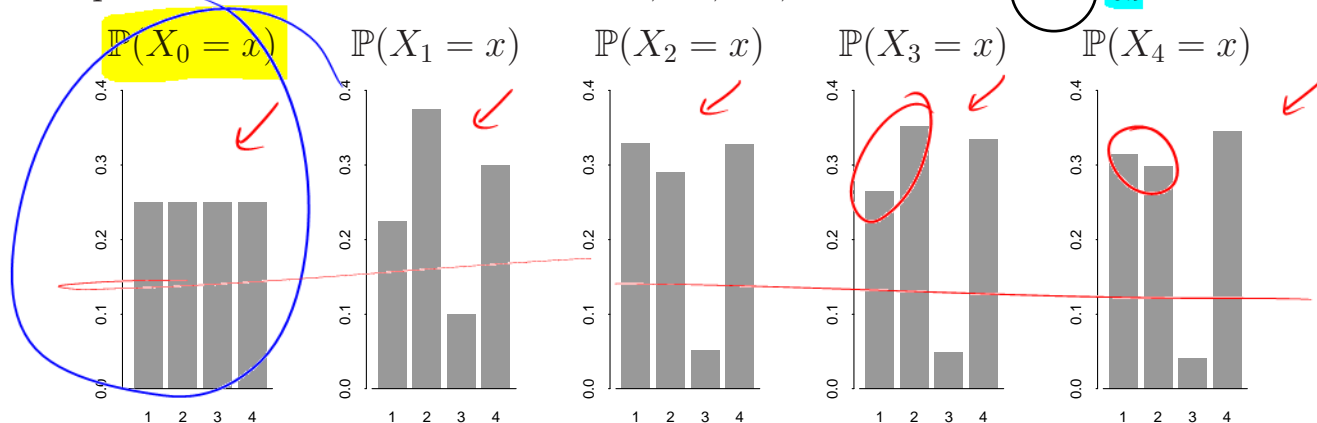
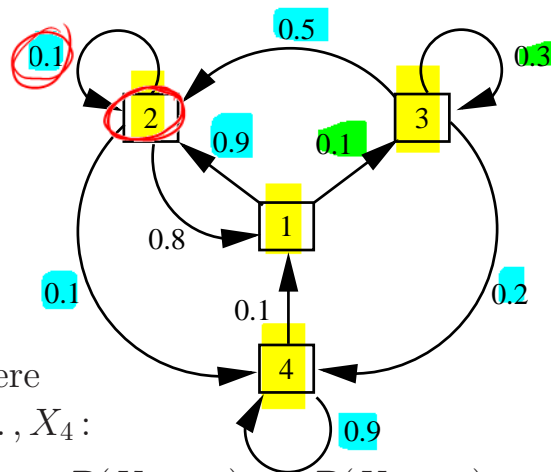
$$T' \sim T'' \sim T$$

9.1 Equilibrium distribution in pictures

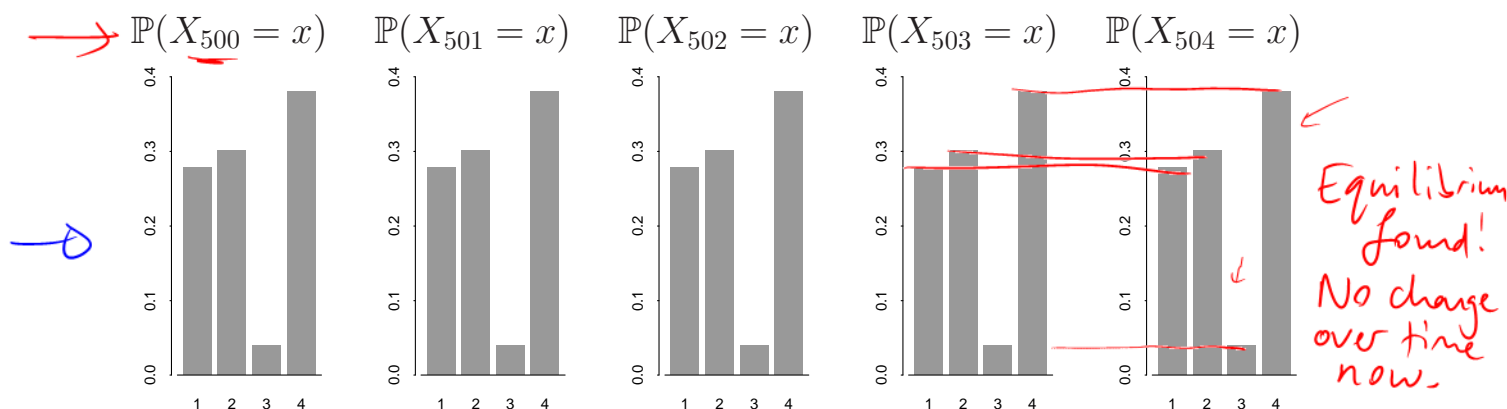
Consider the following 4-state Markov chain:

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

Suppose we start at time 0 with $X_0 \sim (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$: 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, \dots, X_4 :



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:



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

9.2 Calculating equilibrium distributions

Definition: Let $\{X_0, X_1, \dots\}$ be a Markov chain with transition matrix P and state space S , where $|S| = N$ (possibly infinite). Let π^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 $i = 1, \dots, N$. The probability distribution π^T is an equilibrium distribution for the Markov chain if $\pi^T P = \pi^T$.

That is, π^T is an equilibrium distribution if

$$(\pi^T P)_j = \sum_{i=1}^N \pi_i p_{ij} = \pi_j \quad \text{for all } j=1, \dots, N,$$

and $\sum_{i=1}^N \pi_i = 1$.

By the argument given on page 164, we have the following Theorem:

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

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

Note: There are several other names for an equilibrium distribution. If π^T is an equilibrium distribution, it is also called:

- invariant: it doesn't change: $\pi^T P = \pi^T$
- 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 π^T is an equilibrium distribution for P if:

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

Conditions 2 and 3 ensure that π^T really is a probability distribution.

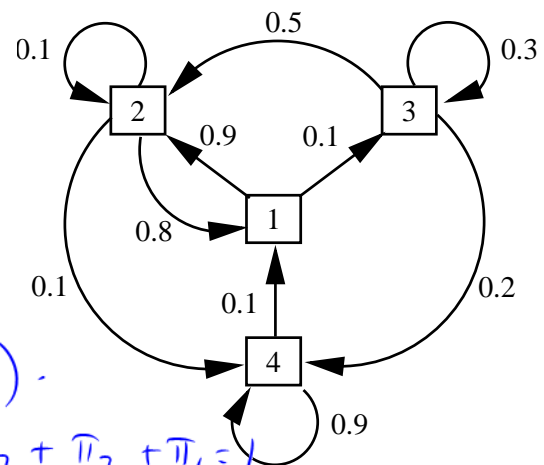
Condition 1 means that π is a row eigenvector of P .

Solving $\pi^T P = \pi^T$ by itself will just specify π up to a scalar multiple.

We need to include Condition 2 to scale π 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 $\pi^T = (\pi_1, \pi_2, \pi_3, \pi_4)$.

Solve $\pi^T P = \pi^T$ and $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$.

$$\pi^T P = \pi^T \Rightarrow (\pi_1 \ \pi_2 \ \pi_3 \ \pi_4) \begin{pmatrix} 0 & .9 & .1 & 0 \\ .8 & .1 & 0 & .1 \\ 0 & .5 & .3 & .2 \\ .1 & 0 & 0 & .9 \end{pmatrix} = (\pi_1 \ \pi_2 \ \pi_3 \ \pi_4)$$

Col. 1

$$.8 \pi_2 + .1 \pi_4 = \pi_1 \quad (1)$$

Col 2

$$.9 \pi_1 + .1 \pi_2 + .5 \pi_3 = \pi_2 \quad (2)$$

Col 3

$$.1 \pi_1 + .3 \pi_3 = \pi_3 \quad (3)$$

Col 4

Ignore : not linearly indept.

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \quad (5) \text{ needed.}$$

Bwied
treasure
exists!

$$(3) \Rightarrow \pi_1 = 7 \pi_3$$

$$\text{Subst in (2)} \Rightarrow .9 (7 \pi_3) + .5 \pi_3 = .9 \pi_2$$

$$\text{Subst in (1)} \Rightarrow .8 \left(\frac{68}{9} \pi_3 \right) + .1 \pi_4 = 7 \pi_3$$

$$\Rightarrow \pi_2 = \frac{68}{9} \pi_3$$

$$\Rightarrow \pi_4 = \frac{86}{9} \pi_3$$

Subst all in (5)

$$\Rightarrow \pi_3 \left(7 + \frac{68}{9} + 1 + \frac{86}{9} \right) = 1$$

$$\Rightarrow \pi_3 = \frac{9}{226}$$

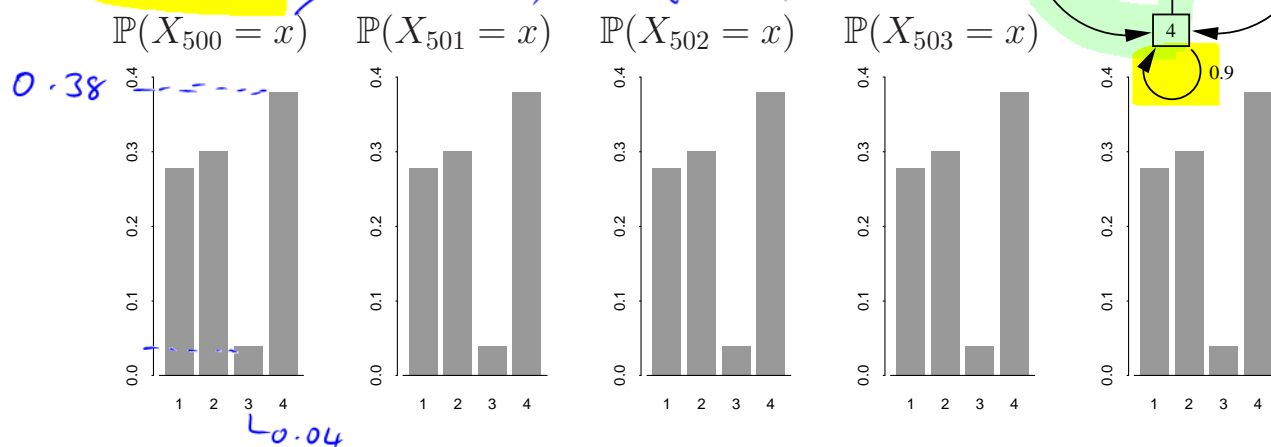
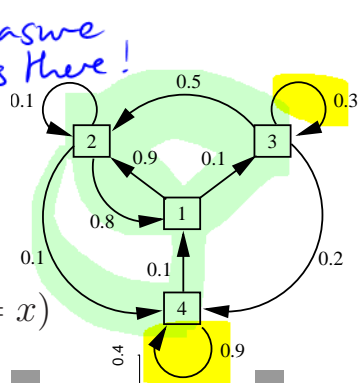
$$\begin{aligned} \text{Overall: } \underline{\pi}^T &= \left(\frac{63}{226}, \frac{68}{226}, \frac{9}{226}, \frac{86}{226} \right) \\ &= (0.28, 0.30, 0.04, 0.38). \end{aligned}$$

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

9.4 Long-term behaviour

*Finding the buried treasure
- chain wanders till it gets there!*

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



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*

$$X_t \sim (0.28, 0.30, 0.04, 0.38) \text{ as } t \rightarrow \infty.$$

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 $t \rightarrow \infty$.*

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

(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.)

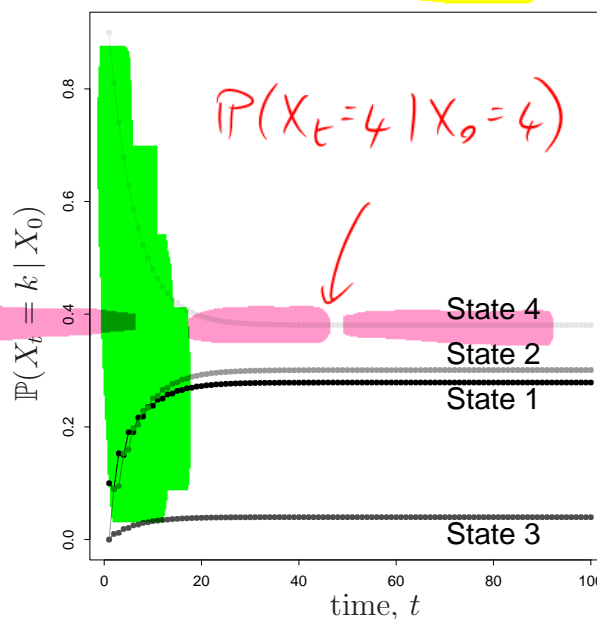
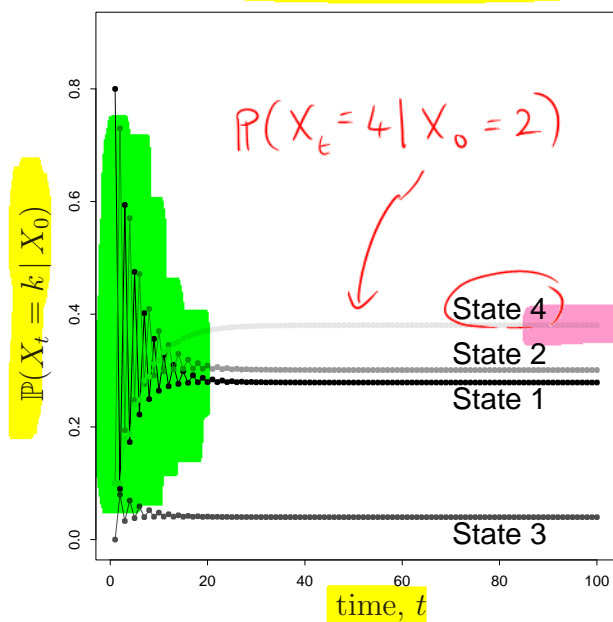
This convergence of P^t means that *for large t , no matter WHICH state we start in, we always have probability:*

- 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 $X_0 = 4$

← start states



destination states

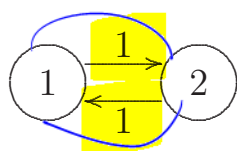
The **left graph** shows the probability of getting from state 2 to state k in t steps, as t changes: $(P^t)_{2,k}$ for $k = 1, 2, 3, 4$.

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:



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

$$P(X_{501} = 1) = 1$$

$$P(X_{502} = 1) = 0$$

As t gets large, P^t does not converge in this case:

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

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

General formula for P^t

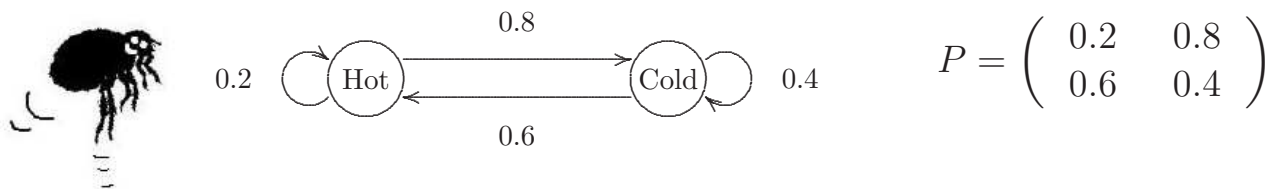
We have seen that we are interested in whether P^t converges to a fixed matrix with all rows equal as $t \rightarrow \infty$.

If it does, then the Markov chain will reach an equilibrium dist. that does not depend on 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:



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

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

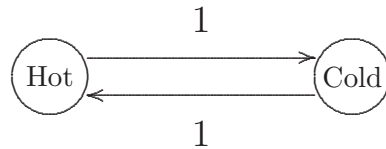
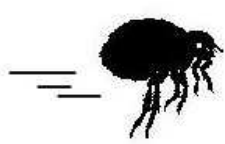
As $t \rightarrow \infty$, $(-0.4)^t \rightarrow 0$, so

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

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

Exercise: Verify that $\underline{\pi}^T = \left(\frac{3}{7}, \frac{4}{7} \right)$ is the same as the result you obtain from solving the equilibrium equations: $\underline{\pi}^T P = \underline{\pi}^T$ 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 $t \rightarrow \infty$, $(-1)^t$ does not converge to anything, \uparrow

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

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

for all t .

\therefore the chain does not forget its starting state,
does not converge to equilibrium as $t \rightarrow \infty$.

treasure exists...

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

\hookrightarrow ...but we don't get there.

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 $t \rightarrow \infty$ 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:

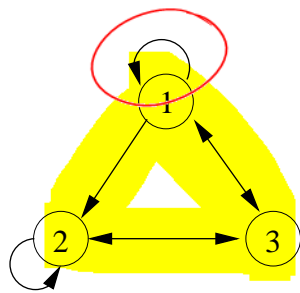
- it will be finite state space
- If a Markov chain is **irreducible** and **aperiodic**, and if an equilibrium distribution π^T exists, then the chain converges to this distribution as $t \rightarrow \infty$, regardless of the initial starting states.
- To make sense of this, we need to revise the concept of **irreducibility**, and introduce the idea of **aperiodicity**.

9.5 Irreducibility

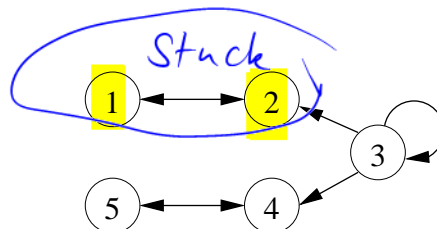
Recall from Chapter 8:

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

An irreducible Markov chain consists of a single class.



Irreducible



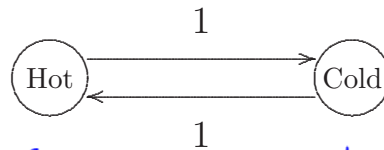
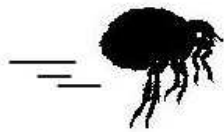
Not irreducible.

Irreducibility of a Markov chain is important for convergence to equilibrium as $t \rightarrow \infty$, because we want 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 $t \rightarrow \infty$, whereas a start state of $X_0 = 4$ means that the chain is restricted to states 4 and 5 as $t \rightarrow \infty$. 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 , Period = 2
 $X_t = 2$ for all odd values of t .

This sort of behaviour is called periodicity: the 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 $t \rightarrow \infty$. For example,

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

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

Period of state i

To formalize the notion of periodicity, we define the period of a state i .

Intuitively, the period is such 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, etc....

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

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

 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 can happen at any time, as t gets large. It's not restricted to certain values of t .

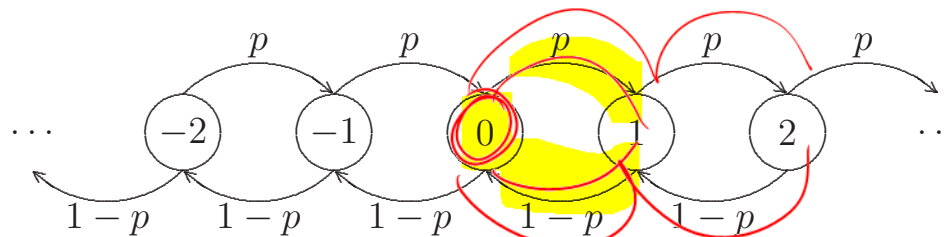
For convergence to equilibrium as $t \rightarrow \infty$, we will be interested only in aperiodic states.

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.

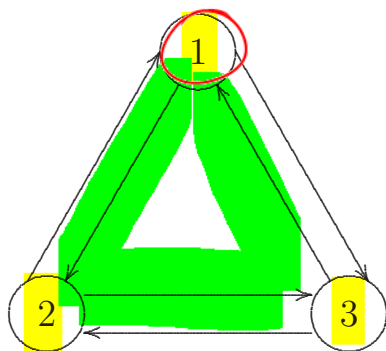
1. The simple random walk.

Irreducible.



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

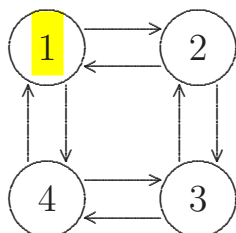
2.



Irreducible

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

3.



Irreducible.

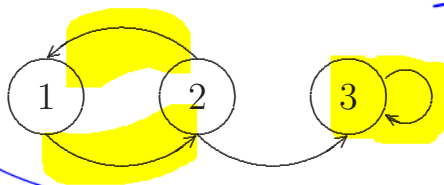
$$d(1) = \gcd \{2, 4, 6, \dots\} = 2.$$

4.

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

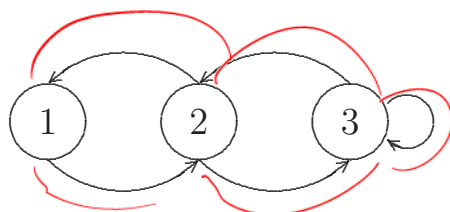
Not irreducible.

$$\underline{\pi}^T = (0, 0, 1)$$



5.

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



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

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

In particular,

$$(P^t)_{ij} \rightarrow \pi_j \text{ as } t \rightarrow \infty, \text{ for all } i, j.$$

So P^t converges to a matrix with all rows identical and equal to π^T .

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

Note: If the state space is infinite, it is not guaranteed that an equilibrium distribution π^T exists. See Example 3 below.

Note: If the chain converges to an equilibrium distribution π^T as $t \rightarrow \infty$, then the long-run proportion of time spent in state k is π_k .
 Ass 5.

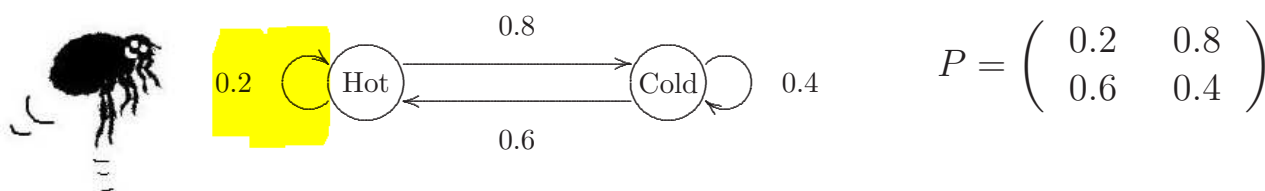
9.8 Examples

IP(success)
 P(wet)

A typical exam question gives you a Markov chain on a finite state space and asks if it converges to an equilibrium distribution as $t \rightarrow \infty$. 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 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 $t \rightarrow \infty$.



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

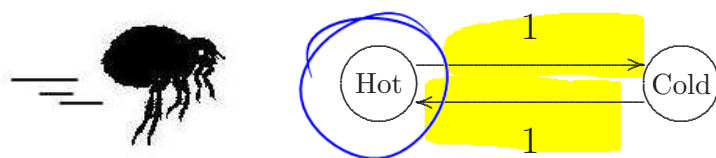
The chain is irreducible and aperiodic.

An equilibrium distn will exist for a finite state space.

So the chain does converge.

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

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



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

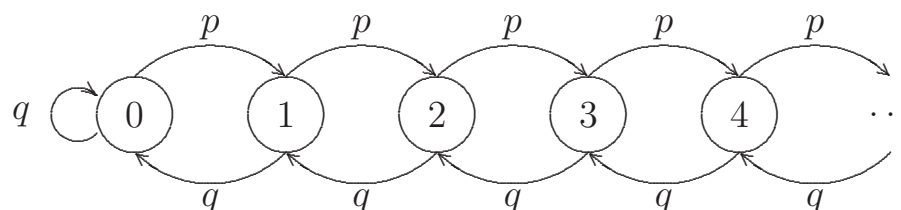
The chain is irreducible, but it is NOT aperiodic:
period = 2.

So the chain does NOT converge to equilibrium
as $t \rightarrow \infty$.

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.

Filled in on
handout.



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

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,

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

- If the state is positive recurrent, then

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

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/>.

