Methods in AI Research: Markov models for multi-agent learning

Part IV: Behaviour of Markov processes

Gerard Vreeswijk

Intelligent Systems Group, Computer Science Department
Faculty of Sciences, Utrecht University, The Netherlands

October, 2014
Important questions

Suppose state $i$ is recurrent.
Important questions

Suppose state $i$ is recurrent.

1. If the process has left $i$, \textit{how many steps will it take}, on average, to return to $i$?
Important questions

Suppose state \( i \) is recurrent.

1. If the process has left \( i \), how many steps will it take, on average, to return to \( i \)?

2. How many times, on the average, will the process visit state \( i \)?
Important questions

Suppose state \( i \) is recurrent.

1. If the process has left \( i \), how many steps will it take, on average, to return to \( i \)?

2. How many times, on the average, will the process visit state \( i \)?

The first question addresses, what is called, the mean recurrence time.
Important questions

Suppose state \( i \) is recurrent.

1. If the process has left \( i \), *how many steps will it take*, on average, to return to \( i \)?

2. *How many times*, on the average, will the process visit state \( i \)?

The first question addresses, what is called, the *mean recurrence time*.

The second question addresses, what is called, the *empirical frequency*.
Important questions

Suppose state $i$ is recurrent.

1. If the process has left $i$, \textit{how many steps will it take}, on average, to return to $i$?

2. \textit{How many times}, on the average, will the process visit state $i$?

- The first question addresses, what is called, the \textit{mean recurrence time}.
- The second question addresses, what is called, the \textit{empirical frequency}.

It turns out that the answers to both questions are strongly related.
Empirical frequency

**Definition (Empirical frequency)**

Let $N_i(j)$ be the number of times a node is visited, starting at $i$. Then

$$F_{ij} = \text{Def} \lim_{n \to \infty} \frac{N_i(j)}{n}.$$
Empirical frequency

**Definition (Empirical frequency)**

Let $N_i(j)$ be the number of times a node is visited, starting at $i$. Then

$$F_{ij} = \text{Def} \lim_{n \to \infty} \frac{N_i(j)}{n}.$$

It is not a priori clear that empirical frequencies exist.
Empirical frequency

**Definition (Empirical frequency)**

Let $N_i(j)$ be the number of times a node is visited, starting at $i$. Then

$$F_{ij} \overset{\text{Def}}{=} \lim_{n \to \infty} \frac{N_i(j)}{n}.$$  

It is not a priori clear that empirical frequencies exist.

**Theorem**

*Empirical frequencies exist for every Markov chain.*
Empirical frequency

**Definition (Empirical frequency)**

Let \( N_i(j) \) be the number of times a node is visited, starting at \( i \). Then

\[
F_{ij} \overset{\text{Def}}{=} \lim_{n \to \infty} \frac{N_i(j)}{n}.
\]

It is not a priori clear that empirical frequencies exist.

**Theorem**

*Empirical frequencies exist for every Markov chain.*

What value does the empirical frequency of a state assume?
Empirical frequency

**Definition (Empirical frequency)**

Let $N_i(j)$ be the number of times a node is visited, starting at $i$. Then

$$F_{ij} = \text{Def} \lim_{n \to \infty} \frac{N_i(j)}{n}.$$ 

It is not a priori clear that empirical frequencies exist.

**Theorem**

*Empirical frequencies exist for every Markov chain.*

What value does the empirical frequency of a state assume?

1. The *long-run average of the $n$-step transition probabilities.*
**Empirical frequency**

**Definition (Empirical frequency)**

Let $N_i(j)$ be the number of times a node is visited, starting at $i$. Then

$$F_{ij} = \text{Def} \lim_{n \to \infty} \frac{N_i(j)}{n}.$$  

It is not a priori clear that empirical frequencies exist.

**Theorem**

Empirical frequencies exist for every Markov chain.

What value does the empirical frequency of a state assume?

1. The *long-run average of the $n$-step transition probabilities*.

2. One over the so-called *mean recurrence time* of that state.
The *long-run average* of a sequence \( a_1, a_2, a_3, \ldots \) is defined as

\[
\lim_{n \to \infty} \frac{a_1 + \cdots + a_n}{n}
\]
The *long-run average* of a sequence $a_1, a_2, a_3, \ldots$ is defined as

$$\lim_{n \to \infty} \frac{a_1 + \cdots + a_n}{n}$$

$\bar{p}_{ij} = \text{Def} \quad \text{long run average of } p^{(1)}_{ij}, p^{(2)}_{ij}, p^{(3)}_{ij}, \ldots$
The long-run average of a sequence $a_1, a_2, a_3, \ldots$ is defined as

$$\lim_{n \to \infty} \frac{a_1 + \cdots + a_n}{n}$$

So

$$\bar{p}_{ij} \overset{\text{Def}}{=} \text{long run average of } p_{ij}^{(1)}, p_{ij}^{(2)}, p_{ij}^{(3)}, \ldots$$

$$\bar{P} = (P + P^2 + P^3 + \cdots + P^n)/n.$$
The *long-run average* of a sequence $a_1, a_2, a_3, \ldots$ is defined as

$$\lim_{n \to \infty} \frac{a_1 + \cdots + a_n}{n}$$

$$\bar{p}_{ij} = \text{Def} \, \text{long run average of } p_{ij}^{(1)}, p_{ij}^{(2)}, p_{ij}^{(3)}, \ldots$$

So

$$\bar{P} = (P + P^2 + P^3 + \cdots + P^n)/n.$$
The **long-run average** of a sequence \( a_1, a_2, a_3, \ldots \) is defined as

\[
\lim_{n \to \infty} \frac{a_1 + \cdots + a_n}{n}
\]

So

\[
\bar{p}_{ij} = \text{Def} \quad \text{long run average of } p_{ij}^{(1)}, p_{ij}^{(2)}, p_{ij}^{(3)}, \ldots
\]

Theorem (**Long-run average of n-step transition probabilities**)

The long-run average of \( n \)-step transition probabilities (exists and) equals the empirical frequencies:

\[
\bar{p}_{ij} = F_{ij}, \text{ for every } i, j.
\]
If

\[
\begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0.2 & 0 & 0 & 1 \\
0.8 & 0 & 0 & 1
\end{pmatrix}
\]
Example

If

then

\[ P = P^2 = \bar{P} = F = \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0.2 & 0.8 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \]
Some processes are periodic

For large values of $n$ the $P(n)$ keep cycling around three distributions:

$$P^*_0 = \lim_{n \to \infty} P_{3n},$$
$$P^*_1 = \lim_{n \to \infty} P_{3n+1},$$
$$P^*_2 = \lim_{n \to \infty} P_{3n+2}.$$
Some processes are periodic

For large values of $n$ the $P^{(n)}$ keep cycling around three distributions:

$$P_0^* = \lim_{n\to\infty} P^{3n}, \quad P_1^* = \lim_{n\to\infty} P^{3n+1}, \quad \text{and} \quad P_2^* = \lim_{n\to\infty} P^{3n+2}.$$
Some processes are periodic

\[
\lim_{n \to \infty} P^{3n} = \begin{pmatrix}
0.81 & 0.35 & 0.65 & 0.19 \\
0.35 & 0.35 & 0.65 & 0.65 \\
0.35 & 0.81 & 0.35 & 0.65 \\
0.81 & 0.35 & 0.65 & 0.19
\end{pmatrix},
\]

and

\[
\lim_{n \to \infty} P^{3n+1} = \begin{pmatrix}
0.35 & 0.35 & 0.65 & 0.65 \\
0.35 & 0.81 & 0.35 & 0.19 \\
0.81 & 0.35 & 0.65 & 0.19 \\
0.35 & 0.35 & 0.65 & 0.65
\end{pmatrix}
\]
Some processes are periodic

\[ \lim_{n \to \infty} P^{3n+2} = \begin{pmatrix} 
0.35 & 0.81 & 0.65 \\
0.35 & 0.35 & 0.65 \\
0.35 & 0.35 & 0.65 \\
0.35 & 0.81 & 0.65 \\
\end{pmatrix}, \]

while

\[ F = \bar{P} = \begin{pmatrix} 
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
\end{pmatrix}. \]
Period

Each period its own color:

- a (0)
- c (2)
- b (1)
- e (0)
- d (2)
- f (1)

Gerard Vreeswijk
MAIR: Markov models for multi-agent learning
Each period its own color:
Period

period 0

period 1

period 2

Gerard Vreeswijk
MAIR: Markov models for multi-agent learning
Theorem

Periodicity is a class property.

Proof:
See notes.
Results on periodicity

**Theorem**

Periodicity is a class property.

**Proof:** See notes.
Theorem

*Periodicity is a class property.*

**Proof:** See notes.

Theorem

*If a class contains a loop, it is a-periodic.*
Results on periodicity

Theorem

*Periodicity is a class property.*

**Proof:** See notes.

Theorem

*If a class contains a loop, it is a-periodic.*

**Proof:** The loop can be used to create cycles of any length.
Try to guess the period of this process.

Hint: first identify all cycles.
Try to guess the period of this process.

Hint: first identify all cycles.

**Solution:** There is one cycle of 6, and one cycle of 10. Because \( \gcd(6, 10) = 2 \), the period is 2.
A Markov process is said to be ergodic if the long-run averages of the n-step transition probabilities do not depend on the starting state. Mixing occurs if the n-step probabilities converge. An ergodic and mixing process has a limit matrix $P^\ast$ that exists and has identical rows. If a process is ergodic but does not mix, it cycles in the neighbourhood of $d$ limit matrices $P^\ast_1, \ldots, P^\ast_d$, where $d$ is the period of the only recurrence class.
A Markov process is said to be **ergodic** if the long-run averages of the $n$-step transition probabilities do not depend on the starting state.
Definition (*Ergodicity, mixing*)

A Markov process is said to be

- **ergodic** if the long-run averages of the $n$-step transition probabilities do not depend on the starting state.
- **mixing** if the $n$-step probabilities converge.
Definition (Ergodicity, mixing)

A Markov process is said to be

- **ergodic** if the long-run averages of the n-step transition probabilities **do not depend on the starting state**.
- **mixing** if the n-step probabilities converge.

1. Ergodic and mixes $\Rightarrow$ limit matrix $P^*$ exists and has identical rows.
Ergodicity, mixing

**Definition (Ergodicity, mixing)**

A Markov process is said to be

- **ergodic** if the long-run averages of the \( n \)-step transition probabilities **do not depend on the starting state**.
- **mixing** if the \( n \)-step probabilities converge.

1. Ergodic and mixes \( \Rightarrow \) limit matrix \( P^* \) exists and has identical rows.

2. Ergodic but does not mix \( \Rightarrow \) process cycles in the neighbourhood of \( d \) limit matrices

\[
P_1^*, \ldots, P_d^*,
\]

where \( d \) is the period of the only recurrence class.
Ergodic, does not mix

Alternating matrices in the limit, so do not mix:

\[
P_{n+1} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},
\]

and

\[
\bar{P} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}.
\]
Ergodic, does not mix

Alternating matrices in the limit, so do not mix:

\[ P = P^{2n+1} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad P = P^{2n} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \]

and \[ F = \bar{P} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}. \]
Mixes, but not ergodic.

\[
\begin{pmatrix}
0.2 \\
0.8
\end{pmatrix}
\]
Mixes, but not ergodic.

Unequal rows so not ergodic:

\[
P = \begin{pmatrix} 0.2 & 0.8 \\ 1 & 1 \\ 0.7 & 0.3 \end{pmatrix}, \quad F = \bar{P} = P^* = \begin{pmatrix} 0.2 & 0.33 & 0.47 \\ 1 & 0.41 & 0.59 \\ 0.41 & 0.59 \end{pmatrix}.
\]
Conditions for ergodicity

Theorem

1. If a chain is uni-chain, it is ergodic.

2. If an irreducible chain is a-periodic, it mixes.

Definition (Regular)

A chain is called regular if, for some \( n \), every entry of \( P^n \) is positive. Regular \( \Rightarrow \) irreducible and mixing.

If finite, the converse holds.

Theorem (Erdős, Feller, Pollard)

If an irreducible chain is ergodic and mixes, then either:

- All states are positive recurrent, and \( P^\ast > 0 \). Moreover, the chain is regular.
- All states are null recurrent or transient, and \( P^\ast = 0 \).
If a chain is uni-chain, it is ergodic.
**Theorem**

1. *If a chain is uni-chain, it is ergodic.*

2. *If an irreducible chain is a-periodic, it mixes.*
Conditions for ergodicity

Theorem

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

Definition (Regular)

A chain is called regular if,
Theorem

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

Definition (Regular)

A chain is called regular if, for some $n$, every entry of $P^n$ is positive.

Regular $\Rightarrow$ irreducible and mixing. If finite, the converse holds.

Theorem (Erdős, Feller, Pollard)

If an irreducible chain is ergodic and mixes, then either

- All states are positive recurrent, and $P^* > 0$. Moreover, the chain is regular.
- All states are null recurrent or transient, and $P^* = 0$.
Conditions for ergodicity

**Theorem**

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

**Definition (Regular)**

A chain is called *regular* if, for some $n$, every entry of $P^n$ is positive.
Conditions for ergodicity

Theorem

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

Definition (Regular)

A chain is called regular if, for some $n$, every entry of $P^n$ is positive.

Regular $\Rightarrow$ irreducible and mixing.
Conditions for ergodicity

**Theorem**

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

**Definition (Regular)**

A chain is called regular if, for some $n$, every entry of $P^n$ is positive.

Regular $\Rightarrow$ irreducible and mixing. If finite, the converse holds.
Conditions for ergodicity

**Theorem**

1. *If a chain is uni-chain, it is ergodic.*
2. *If an irreducible chain is a-periodic, it mixes.*

**Definition (Regular)**

A chain is called *regular* if, for some $n$, every entry of $P^n$ is positive.

Regular $\Rightarrow$ irreducible and mixing. If finite, the converse holds.

**Theorem (Erdös, Feller, Pollard)**

*If an irreducible chain is ergodic and mixes, then either*

- All states are positive recurrent, and $P^n > 0$. Moreover, the chain is regular.
- All states are null recurrent or transient, and $P^n = 0$. 

Gerard Vreeswijk

MAIR: Markov models for multi-agent learning
Conditions for ergodicity

Theorem

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

Definition (Regular)

A chain is called regular if, for some $n$, every entry of $P^n$ is positive.

Regular $\Rightarrow$ irreducible and mixing. If finite, the converse holds.

Theorem (Erdös, Feller, Pollard)

If an irreducible chain is ergodic and mixes, then either

- All states are positive recurrent, and $P^* > 0$. Moreover, the chain is regular.
Conditions for ergodicity

Theorem

1. If a chain is uni-chain, it is ergodic.
2. If an irreducible chain is a-periodic, it mixes.

Definition (Regular)

A chain is called regular if, for some $n$, every entry of $P^n$ is positive.

Regular $\Rightarrow$ irreducible and mixing. If finite, the converse holds.

Theorem (Erdös, Feller, Pollard)

If an irreducible chain is ergodic and mixes, then either
- All states are positive recurrent, and $P^* > 0$. Moreover, the chain is regular.
- All states are null recurrent or transient, and $P^* = 0$. 

Gerard Vreeswijk

MAIR: Markov models for multi-agent learning
The **first passage time** from \( i \) to \( j \), denoted by \( T_{ij} \), is defined as the number of steps to \( j \), after starting in \( i \), **with the obligation to leave** \( i \).
The **first passage time** from \(i\) to \(j\), denoted by \(T_{ij}\), is defined as the number of steps to \(j\), after starting in \(i\), with the obligation to leave \(i\).

Now:

\[
f_{ij}^{(n)} = \text{Def } P\{T_{ij} = n\}, \quad n \geq 1
\]
The first passage time from \( i \) to \( j \), denoted by \( T_{ij} \), is defined as the number of steps to \( j \), after starting in \( i \), with the obligation to leave \( i \).

Now:

\[
f_{ij}^{(n)} = \text{Def } P\{ T_{ij} = n \}, \quad n \geq 1
\]

The probabilities \( p_{ij}^{(n)} \) and \( f_{ij}^{(n)} \) are equally important and have similar intuitions:
The **first passage time** from $i$ to $j$, denoted by $T_{ij}$, is defined as the number of steps to $j$, after starting in $i$, *with the obligation to leave $i$*. 

Now:

$$f_{ij}^{(n)} = \text{Def } P\{T_{ij} = n\}, \quad n \geq 1$$

The probabilities $p_{ij}^{(n)}$ and $f_{ij}^{(n)}$ are equally important and have similar intuitions:

$p_{ij}^{(n)}$: The probability of passing through $j$, when starting in $i$ and taking $n$ steps.
The **first passage time** from \(i\) to \(j\), denoted by \(T_{ij}\), is defined as the number of steps to \(j\), after starting in \(i\), **with the obligation to leave \(i\)**.

Now:

\[
f_{ij}^{(n)} = \text{Def } P\{T_{ij} = n\}, \quad n \geq 1
\]

The probabilities \(p_{ij}^{(n)}\) and \(f_{ij}^{(n)}\) are equally important and have similar intuitions:

- \(p_{ij}^{(n)}\): The probability of passing through \(j\), when starting in \(i\) and taking \(n\) steps.
- \(f_{ij}^{(n)}\): The probability of passing through \(j\) **for the first time**, when starting in \(i\) and taking \(n\) steps.
Computing first passage time probabilities

Theorem (First passage time probability)

\[
 f_{ij}^{(n)} = \begin{cases} 
 p_{ij} & \text{if } n = 1, \\
 \sum_{k \neq j} p_{ik} f_{kj}^{(n-1)} & \text{otherwise.} 
\end{cases}
\]

Problem. Given the Markov process with transition matrix

\[
 P = \begin{pmatrix} 
 0.4 & 0.6 \\
 0.1 & 0.9 \\
 0.3 & 0.5 & 0.2 
\end{pmatrix}.
\]

Compute \( f_{13}^{(2)} \).
Computing first passage time probabilities

Theorem (First passage time probability)

\[ f_{ij}^{(n)} = \begin{cases} p_{ij} & \text{if } n = 1, \\ \sum_{k \neq j} p_{ik} f_{kj}^{(n-1)} & \text{otherwise.} \end{cases} \]

Problem. Given the Markov process with transition matrix

\[ P = \begin{pmatrix} 0.4 & 0.6 \\ 0.1 & 0.9 \\ 0.3 & 0.5 & 0.2 \end{pmatrix}. \]

Compute \( f_{13}^{(2)} \).

Solution:

\[ f_{13}^{(2)} = p_{11} f_{13}^{(1)} + p_{12} f_{23}^{(1)} \]
\[ = p_{11} p_{13} + p_{12} p_{23} \]
\[ = 0.0 \times 0.6 + 0.4 \times 0.9 = 0.36. \]
The hitting time from $i$ to $j$, denoted by $H_{ij}$, is defined as the number of steps to $j$, after starting in $i$, **without the obligation to leave $i$.**
The hitting time from $i$ to $j$, denoted by $H_{ij}$, is defined as the number of steps to $j$, after starting in $i$, without the obligation to leave $i$.

**Theorem**

The mean hitting times are given by the minimal non-negative solution of

$$h_{ij} = \begin{cases} 0 & \text{if } i = j, \\ 1 + \sum_{\text{all } k} p_{ik} h_{kj} & \text{otherwise}. \end{cases}$$
Compute the expected hitting time $h_{13}$. **Solution:**
Problem

Compute the expected hitting time $h_{13}$. **Solution:**

\[
\begin{pmatrix}
0.9 \\
0.1 \\
0.1 \\
0.9 \\
0.1 \\
0.9
\end{pmatrix}
\]
Compute the expected hitting time $h_{13}$. **Solution:**

$$
\begin{pmatrix}
0 & h_{12} & h_{13}
\end{pmatrix}
$$
Compute the expected hitting time $h_{13}$. Solution:

$$
\begin{pmatrix}
0 & h_{12} & h_{13} \\
h_{21} & 0 & h_{23}
\end{pmatrix}
$$
Compute the expected hitting time $h_{13}$. Solution:

$$
\begin{pmatrix}
0 & h_{12} & h_{13} \\
h_{21} & 0 & h_{23} \\
h_{31} & h_{32} & 0
\end{pmatrix}
$$
Compute the expected hitting time $h_{13}$. **Solution:**

$$
\begin{pmatrix}
0 & h_{12} & h_{13} \\
h_{21} & 0 & h_{23} \\
h_{31} & h_{32} & 0 \\
\end{pmatrix}
= \\
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
$$
Compute the expected hitting time $h_{13}$. **Solution:**

\[
\begin{pmatrix}
0 & h_{12} & h_{13} \\
h_{21} & 0 & h_{23} \\
h_{31} & h_{32} & 0
\end{pmatrix} = \\
\begin{pmatrix}
0 & 1 + 0.9h_{12} & 1 + 0.9h_{13} + 0.1h_{23}
\end{pmatrix}
\]
Compute the expected hitting time $h_{13}$. **Solution:**

\[
\begin{pmatrix}
0 & h_{12} & h_{13} \\
h_{21} & 0 & h_{23} \\
h_{31} & h_{32} & 0
\end{pmatrix} =
\begin{pmatrix}
0 & 1 + 0.9h_{12} & 1 + 0.9h_{13} + 0.1h_{23} \\
1 + 0.9h_{21} + 0.1h_{31} & 0 & 1 + 0.9h_{23}
\end{pmatrix}
\]
Problem

Compute the expected hitting time $h_{13}$. **Solution:**

\[
\begin{pmatrix}
0 & h_{12} & h_{13} \\
1 + 0.9h_{21} + 0.1h_{31} & 0 & 1 + 0.9h_{23} \\
1 + h_{31} & 1 + h_{32} & 0
\end{pmatrix}
\]
Solution, continued
Solution, continued

\[
\begin{align*}
    h_{12} &= 1 + 0.9 h_{12} \\
    h_{13} &= 1 + 0.9 h_{13} + 0.1 h_{23} \\
    h_{21} &= 1 + 0.9 h_{21} + 0.1 h_{31} \\
    h_{23} &= 10 h_{23} \\
    h_{31} &= 1 + h_{31} \\
    h_{32} &= 1 + h_{32} \\
    h_{32} &= \infty \Rightarrow h_{31} = \infty
\end{align*}
\]
Solution, continued

\[
\begin{align*}
  h_{12} &= 1 + 0.9h_{12} \\
  \Rightarrow \quad h_{12} &= 10
\end{align*}
\]
Solution, continued

\[
\begin{align*}
    h_{12} &= 1 + 0.9h_{12} &\Rightarrow h_{12} &= 10 \\
    h_{13} &= 1 + 0.9h_{13} + 0.1h_{23}
\end{align*}
\]
Solution, continued

\[
\begin{align*}
    h_{12} &= 1 + 0.9h_{12} \quad \Rightarrow \quad h_{12} = 10 \\
    h_{13} &= 1 + 0.9h_{13} + 0.1h_{23} \\
    h_{21} &= 1 + 0.9h_{21} + 0.1h_{31}
\end{align*}
\]
Expected hitting time

Solution, continued

\[
\begin{align*}
h_{12} &= 1 + 0.9h_{12} \quad \Rightarrow \quad h_{12} = 10 \\
h_{13} &= 1 + 0.9h_{13} + 0.1h_{23} \\
h_{21} &= 1 + 0.9h_{21} + 0.1h_{31} \\
h_{23} &= 1 + 0.9h_{23} \quad \Rightarrow \quad h_{23} = 10
\end{align*}
\]
Solution, continued

\[
\begin{align*}
  h_{12} &= 1 + 0.9h_{12} \\
  h_{13} &= 1 + 0.9h_{13} + 0.1h_{23} \\
  h_{21} &= 1 + 0.9h_{21} + 0.1h_{31} \\
  h_{23} &= 1 + 0.9h_{23} \\
  h_{31} &= 1 + h_{31}
\end{align*}
\]

\[\Rightarrow h_{12} = 10\]

\[\Rightarrow h_{23} = 10\]

\[\Rightarrow h_{31} = \infty\]
Expected hitting time

Solution, continued

\[
\begin{align*}
  h_{12} &= 1 + 0.9h_{12} & \Rightarrow & & h_{12} = 10 \\
  h_{13} &= 1 + 0.9h_{13} + 0.1h_{23} \\
  h_{21} &= 1 + 0.9h_{21} + 0.1h_{31} \\
  h_{23} &= 1 + 0.9h_{23} & \Rightarrow & & h_{23} = 10 \\
  h_{31} &= 1 + h_{31} & \Rightarrow & & h_{31} = \infty \\
  h_{32} &= 1 + h_{32} & \Rightarrow & & h_{32} = \infty
\end{align*}
\]
Expected hitting time

Solution, continued

\[
\left\{
\begin{align*}
h_{12} &= 1 + 0.9h_{12} &\Rightarrow h_{12} &= 10 \\
h_{13} &= 1 + 0.9h_{13} + 0.1h_{23} \\
h_{21} &= 1 + 0.9h_{21} + 0.1h_{31} \\
h_{23} &= 1 + 0.9h_{23} &\Rightarrow h_{23} &= 10 \\
h_{31} &= 1 + h_{31} &\Rightarrow h_{31} &= \infty \\
h_{32} &= 1 + h_{32} &\Rightarrow h_{32} &= \infty
\end{align*}
\right.
\]
Expected hitting time

Solution, continued

\[
\begin{align*}
    h_{12} &= 1 + 0.9 h_{12} & \Rightarrow & & h_{12} &= 10 \\
    h_{13} &= 1 + 0.9 h_{13} + 0.1 h_{23} \\
    h_{21} &= 1 + 0.9 h_{21} + 0.1 h_{31} \\
    h_{23} &= 1 + 0.9 h_{23} & \Rightarrow & & h_{23} &= 10 \\
    h_{31} &= 1 + h_{31} & \Rightarrow & & h_{31} &= \infty \\
    h_{32} &= 1 + h_{32} & \Rightarrow & & h_{32} &= \infty \\

\end{align*}
\]
Expected hitting time

Solution, continued

\[
\begin{align*}
  h_{12} & = 1 + 0.9h_{12} \quad \Rightarrow \quad h_{12} = 10 \\
  h_{13} & = 1 + 0.9h_{13} + 0.1h_{23} \\
  h_{21} & = 1 + 0.9h_{21} + 0.1h_{31} \\
  h_{23} & = 1 + 0.9h_{23} \quad \Rightarrow \quad h_{23} = 10 \\
  h_{31} & = 1 + h_{31} \quad \Rightarrow \quad h_{31} = \infty \\
  h_{32} & = 1 + h_{32} \quad \Rightarrow \quad h_{32} = \infty
\end{align*}
\]

\[
\begin{align*}
  h_{12} & = 1 + 0.9h_{12} \\
  h_{13} & = 1 + 0.9h_{13} + 0.1 \times 10 \quad \Rightarrow \quad h_{13} = 20
\end{align*}
\]
Expected hitting time

Solution, continued

\[
\begin{align*}
    h_{12} &= 1 + 0.9h_{12} & \Rightarrow & & h_{12} = 10 \\
    h_{13} &= 1 + 0.9h_{13} + 0.1h_{23} \\
    h_{21} &= 1 + 0.9h_{21} + 0.1h_{31} \\
    h_{23} &= 1 + 0.9h_{23} & \Rightarrow & & h_{23} = 10 \\
    h_{31} &= 1 + h_{31} & \Rightarrow & & h_{31} = \infty \\
    h_{32} &= 1 + h_{32} & \Rightarrow & & h_{32} = \infty
\end{align*}
\]

\[
\begin{align*}
    h_{12} &= 1 + 0.9h_{12} \\
    h_{13} &= 1 + 0.9h_{13} + 0.1 \times 10 & \Rightarrow & & h_{13} = 20 \\
    h_{21} &= 1 + 0.9h_{21} + 0.1 \times \infty & \Rightarrow & & h_{21} = \infty
\end{align*}
\]
Expected hitting time

Solution, continued

\[
\begin{align*}
    h_{12} &= 1 + 0.9 h_{12} \\
    h_{13} &= 1 + 0.9 h_{13} + 0.1 h_{23} \\
    h_{21} &= 1 + 0.9 h_{21} + 0.1 h_{31} \\
    h_{23} &= 1 + 0.9 h_{23} \\
    h_{31} &= 1 + h_{31} \\
    h_{32} &= 1 + h_{32}
\end{align*}
\]

\[\Rightarrow \quad h_{12} = 10 \]

\[\Rightarrow \quad h_{23} = 10 \]

\[\Rightarrow \quad h_{31} = \infty \]

\[\Rightarrow \quad h_{32} = \infty \]

\[\Rightarrow \quad h_{12} = 1 + 0.9 h_{12} \\
    h_{13} = 1 + 0.9 h_{13} + 0.1 \times 10 \Rightarrow h_{13} = 20 \\
    h_{21} = 1 + 0.9 h_{21} + 0.1 \times \infty \Rightarrow h_{21} = \infty \\
    h_{23} = 1 + 0.9 h_{23}
\]
Solution, continued

\[
\begin{align*}
  h_{12} &= 1 + 0.9 h_{12} \quad \Rightarrow \quad h_{12} = 10 \\
  h_{13} &= 1 + 0.9 h_{13} + 0.1 h_{23} \\
  h_{21} &= 1 + 0.9 h_{21} + 0.1 h_{31} \\
  h_{23} &= 1 + 0.9 h_{23} \quad \Rightarrow \quad h_{23} = 10 \\
  h_{31} &= 1 + h_{31} \quad \Rightarrow \quad h_{31} = \infty \\
  h_{32} &= 1 + h_{32} \quad \Rightarrow \quad h_{32} = \infty
\end{align*}
\]

\[
\Rightarrow \quad \begin{align*}
  h_{12} &= 1 + 0.9 h_{12} \\
  h_{13} &= 1 + 0.9 h_{13} + 0.1 \times 10 \quad \Rightarrow \quad h_{13} = 20 \\
  h_{21} &= 1 + 0.9 h_{21} + 0.1 \times \infty \quad \Rightarrow \quad h_{21} = \infty \\
  h_{23} &= 1 + 0.9 h_{23} \\
  h_{31} &= 1 + h_{31}
\end{align*}
\]
Expected hitting time

Solution, continued

\[
\begin{align*}
  h_{12} &= 1 + 0.9 h_{12} \quad \Rightarrow \quad h_{12} = 10 \\
  h_{13} &= 1 + 0.9 h_{13} + 0.1 h_{23} \\
  h_{21} &= 1 + 0.9 h_{21} + 0.1 h_{31} \\
  h_{23} &= 1 + 0.9 h_{23} \quad \Rightarrow \quad h_{23} = 10 \\
  h_{31} &= 1 + h_{31} \quad \Rightarrow \quad h_{31} = \infty \\
  h_{32} &= 1 + h_{32} \quad \Rightarrow \quad h_{32} = \infty
\end{align*}
\]

\[
\begin{align*}
  h_{12} &= 1 + 0.9 h_{12} \\
  h_{13} &= 1 + 0.9 h_{13} + 0.1 \times 10 \quad \Rightarrow \quad h_{13} = 20 \\
  h_{21} &= 1 + 0.9 h_{21} + 0.1 \times \infty \quad \Rightarrow \quad h_{21} = \infty \\
  h_{23} &= 1 + 0.9 h_{23} \\
  h_{31} &= 1 + h_{31} \\
  h_{32} &= 1 + h_{32}
\end{align*}
\]
The **mean first passage time**: 

\[ \mu_{ij} = \text{Def } E[T_{ij}] \]
The mean first passage time:

\[ \mu_{ij} = \text{Def } E[T_{ij}] \]

Lemma

\[ \mu_{ij} = 1 + \sum_{\text{all } k} p_{ik} h_{kj}, \]

where \( h_{ij} \) is the mean hitting time.
The mean first passage time:

$$\mu_{ij} = \text{Def } E[T_{ij}]$$

Lemma

$$\mu_{ij} = 1 + \sum_{\text{all } k} p_{ik} h_{kj},$$

where $h_{ij}$ is the mean hitting time.

The mean recurrence time:

$$\mu_i = \text{Def } \mu_{ii}$$
Mean recurrence time = 1 / empirical frequency

If the process has left state $i$, how many steps will it take, on average, to return to state $i$?

How many times, on the average, will the process visit state $i$?

**Theorem**

The mean recurrence time is the reciprocal of the empirical frequency:

$$\mu_{ii} = \frac{1}{F_{ii}}$$

The theorem is valid for all irreducible chains and remains valid in case $F_{ii} = 0$ and $1/0$ is interpreted as $+\infty$. 

Gerard Vreeswijk  
MAIR: Markov models for multi-agent learning
If the process has left *i*, *how many steps will it take*, on average, to return to *i*?
1. If the process has left $i$, how many steps will it take, on average, to return to $i$?

2. How many times, on the average, will the process visit state $i$?
If the process has left \( i \), **how many steps will it take**, on average, to return to \( i \)?

**How many times**, on the average, will the process visit state \( i \)?

**Theorem**

The mean recurrence time is the reciprocal of the empirical frequency:

\[
\mu_{ii} = \frac{1}{F_{ii}}
\]
Mean recurrence time $= 1 / \text{empirical frequency}$

1. If the process has left $i$, how many steps will it take, on average, to return to $i$?

2. How many times, on the average, will the process visit state $i$?

**Theorem**

The mean recurrence time is the reciprocal of the empirical frequency:

$$\mu_{ii} = \frac{1}{F_{ii}}$$

The theorem is valid for all irreducible chains and remains valid in case $F_{ii} = 0$ and $1/0$ is interpreted as $+\infty$. 
Because $\mu_{ii} = \frac{1}{F_{ii}}$, immediately
$\mu_i = (10.12, 10.27, 10.12, 10.22, 10.22, 10.06)$
Example

![Diagram of a graph with nodes labeled e, f, c, a, b, d and edges with probabilities 0.3, 0.2, 0.7, 0.8.]

\[ F = \bar{P} = \begin{pmatrix}
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
\end{pmatrix}. \]
Example

Because $\mu_{ii} = 1/F_{ii}$, immediately

$$\mu_i = \left( \frac{1}{0.12}, \frac{1}{0.27}, \frac{1}{0.12}, \frac{1}{0.22}, \frac{1}{0.22}, \frac{1}{0.06} \right)$$

$$F = \bar{P} = \begin{pmatrix} 0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \end{pmatrix}.$$
Example

\[ F = \bar{\mathbf{P}} = \begin{pmatrix}
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0.12 & 0.27 & 0.12 & 0.22 & 0.22 & 0.06
\end{pmatrix}. \]

Because \( \mu_{ii} = 1/F_{ii} \), immediately

\[ \mu_i = \left( \frac{1}{0.12}, \frac{1}{0.27}, \frac{1}{0.12}, \frac{1}{0.22}, \frac{1}{0.22}, \frac{1}{0.06} \right) = (8.5, 3.7, 8.5, 4.7, 4.7, 15.5). \]
Problem

If in the bonus-malus problem, suppose \( p_0 = 0.8, \ p_1 = 0.1, \ p_2 = 0.05, \) and \( p_{\geq 3} = 0.05. \)
If in the bonus-malus problem, suppose $p_0 = 0.8$, $p_1 = 0.1$, $p_2 = 0.05$, and $p_{\geq 3} = 0.05$. Then

$$\lim_{n \to \infty} P^n \text{ exists and equals}$$

$$\begin{pmatrix} 0.6337 & 0.1584 & 0.1188 & 0.0891 \\ 0.6337 & 0.1584 & 0.1188 & 0.0891 \\ 0.6337 & 0.1584 & 0.1188 & 0.0891 \\ 0.6337 & 0.1584 & 0.1188 & 0.0891 \end{pmatrix}$$
Problem

If in the bonus-malus problem, suppose $p_0 = 0.8$, $p_1 = 0.1$, $p_2 = 0.05$, and $p_{\geq 3} = 0.05$. Then

$$\lim_{n \to \infty} P^n$$
exists and equals

$$\begin{pmatrix}
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
0.6337 & 0.1584 & 0.1188 & 0.0891
\end{pmatrix}$$

1. **Determine the empirical frequencies.**
If in the bonus-malus problem, suppose $p_0 = 0.8$, $p_1 = 0.1$, $p_2 = 0.05$, and $p_{\geq 3} = 0.05$. Then

$$\lim_{n \to \infty} P^n$$
exists and equals

$$\begin{pmatrix}
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
0.6337 & 0.1584 & 0.1188 & 0.0891 \\
\end{pmatrix}$$

1. Determine the empirical frequencies.
2. Determine the mean recurrence times.

Hint: Since
Determine the empirical frequencies.

\[
F = P^* = \lim_{n \to \infty} P^n.
\]

The empirical frequencies can be read off from \( P^* \):

\[
(F_1, F_2, F_3, F_4) = (0.6337, 0.1584, 0.1188, 0.0891).
\]

Determine the mean recurrence times.

\[
\mu_{ii} = \frac{1}{F_{ii}}.
\]

So

\[
(\mu_1, \mu_2, \mu_3, \mu_4) = \left( \frac{1}{0.6337}, \frac{1}{0.1584}, \frac{1}{0.1188}, \frac{1}{0.0891} \right) \approx (1.58, 6.31, 8.42, 11.22).
\]
Determine the empirical frequencies.

**Solution:** The process mixes so

\[ F = P^* = \lim_{n \to \infty} P^n. \]
Determine the empirical frequencies.

**Solution:** The process mixes so

\[ F = P^* = \lim_{n \to \infty} P^n. \]

The empirical frequencies can be read off from \( P^* \):

\[ (F_1, F_2, F_3, F_4) \]
Determine the empirical frequencies. 

**Solution:** The process mixes so

\[ F = P^* = \lim_{n \to \infty} P^n. \]

The empirical frequencies can be read off from \( P^* \):

\[ (F_1, F_2, F_3, F_4) = (0.6337, 0.1584, 0.1188, 0.0891). \]
Solution bonus-malus problem

1. Determine the empirical frequencies.

   **Solution:** The process mixes so

   \[ F = P^* = \lim_{n \to \infty} P^n. \]

   The empirical frequencies can be read off from \( P^* \):

   \[ (F_1, F_2, F_3, F_4) = (0.6337, 0.1584, 0.1188, 0.0891). \]

2. Determine the mean recurrence times.
Determine the empirical frequencies.

**Solution:** The process mixes so

\[ F = P^* = \lim_{n \to \infty} P^n. \]

The empirical frequencies can be read off from \( P^* \):

\[ (F_1, F_2, F_3, F_4) = (0.6337, 0.1584, 0.1188, 0.0891). \]

Determine the mean recurrence times.

**Solution:** This chain is irreducible. Hence, \( \mu_{ii} = 1/F_{ii} \).
Determine the empirical frequencies.

**Solution:** The process mixes so

\[ F = P^* = \lim_{n \to \infty} P^n. \]

The empirical frequencies can be read off from \( P^* \):

\[
(F_1, F_2, F_3, F_4) = (0.6337, 0.1584, 0.1188, 0.0891).
\]

Determine the mean recurrence times.

**Solution:** This chain is irreducible. Hence, \( \mu_{ii} = 1/F_{ii} \). So

\[
(\mu_1, \mu_2, \mu_3, \mu_4) = (1/0.6337, 1/0.1584, 1/0.1188, 1/0.0891)
\]
1. Determine the empirical frequencies.

**Solution:** The process mixes so

\[ F = P^* = \lim_{n \to \infty} P^n. \]

The empirical frequencies can be read off from \( P^* \):

\[ (F_1, F_2, F_3, F_4) = (0.6337, 0.1584, 0.1188, 0.0891). \]

2. Determine the mean recurrence times.

**Solution:** This chain is irreducible. Hence, \( \mu_{ii} = 1/F_{ii} \). So

\[ (\mu_1, \mu_2, \mu_3, \mu_4) = (1/0.6337, 1/0.1584, 1/0.1188, 1/0.0891) \]
\[ \approx (1.58, 6.31, 8.42, 11.22). \]