Markov Chains

Stochastics

Illés Horváth

2021/10/26

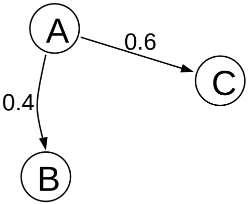
1. Definition, introduction
2. Structural properties
3. Long-term behaviour
4. Reducible Markov chains
5. Periodic Markov chains
6. Stationary vector interpretations
7. Problems

Randomly changing system

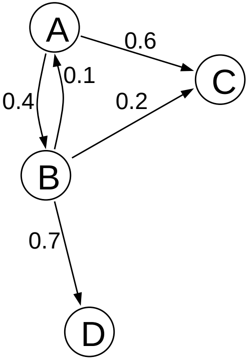
A system starts from state A.

Randomly changing system

It can change to state B or C randomly (with prob. 0.4 and 0.6).

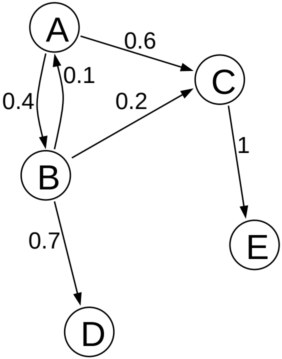


From state B, it can go to A, C or D randomly.



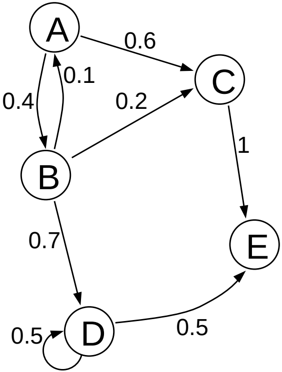
Randomly changing system

From C, it can only go to E.

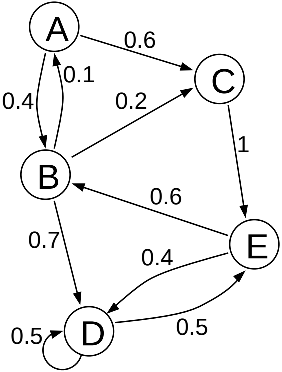


Randomly changing system

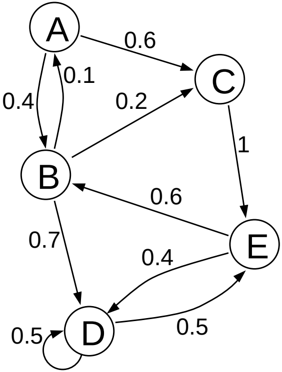
From D, it may go to E or back to D.



This is a Markov chain.



A possible realization



A possible realization: A, B, D, D, E, D, E, B, C, E, B, A, C . . .

Markov chain

A Markov chain is a system that can randomly change its state.

The probabilities with which it chooses the next state only depend on the current state, not what has happened before. This is known as the *Markov property*.

Markov chain

A Markov chain is a system that can randomly change its state.

The probabilities with which it chooses the next state only depend on the current state, not what has happened before. This is known as the *Markov property*.

The Markov property is a sort of memoryless property: the future behaviour of the Markov chain only depends on the current state, not the past.

A Markov chain is a system that can randomly change its state.

The probabilities with which it chooses the next state only depend on the current state, not what has happened before. This is known as the *Markov property*.

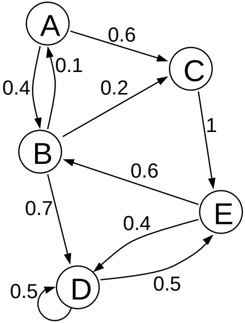
The Markov property is a sort of memoryless property: the future behaviour of the Markov chain only depends on the current state, not the past.

To define a Markov chain, we need the following information:

- list of states (e.g. A, B, C, D, E),
- initial state, and
- either a directed graph with probabilities on the edges (as before), or a *transition probability matrix* P.

Transition probability matrix

Element (i, j) of the transition probability matrix P is the probability that the Markov chain will go (transition) to state j next, assuming it is currently in state i.



	A	B	C	D	E
A	0	0.4	0.6	0	0
B	0.1	0	0.2	0.7	0
C	0	0	0	0	1
D	0	0	0	0.5	0.5
E	0	0.6	0	0.4	0

Initial state vector

The initial state can also be given in vector form, e.g. if the initial state is A, then the initial vector is

$$v_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Initial state vector

The initial state can also be given in vector form, e.g. if the initial state is A, then the initial vector is

$$v_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

With this notation, it is also possible to have a random initial vector, e.g.

$$v_0 = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \end{pmatrix}.$$

means that the Markov chain starts from state A with probability $1/2$ or state C with probability $1/2$.

Initial state vector

The initial state can also be given in vector form, e.g. if the initial state is A, then the initial vector is

$$v_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

With this notation, it is also possible to have a random initial vector, e.g.

$$v_0 = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \end{pmatrix}.$$

means that the Markov chain starts from state A with probability 1/2 or state C with probability 1/2.

The elements of v_0 are nonnegative and their sum is 1.

Stochastic matrix

Properties of the transition probability matrix P:

- Its elements are nonnegative:

$$p_{ij} \geq 0 \, \forall i, j.$$

- Each row of P has a sum equal to 1:

$$\sum_i p_{ij} = 1.$$

Stochastic matrix

Properties of the transition probability matrix P:

- Its elements are nonnegative:

$$p_{ij} \geq 0 \, \forall i, j.$$

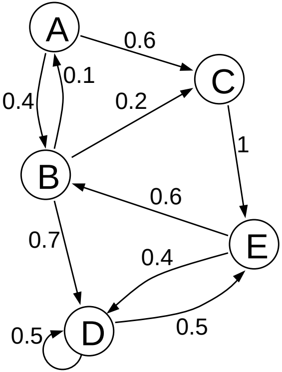
- Each row of P has a sum equal to 1:

$$\sum_{i} p_{ij} = 1.$$

Any square matrix satisfying the above properties is a valid probability transition matrix. A square matrix satisfying the above properties is also known as a *stochastic matrix*.

State probabilities

Assume the Markov chain starts from state A. Where can it be after 1 step? Where can it be after 2 steps?



State probabilities

After 1 step, the state is random: it can be either B with probability 0.4 or C with probability 0.6, so the state probability vector after 1 step is

$$v_1 = \begin{pmatrix} 0 & 0.4 & 0.6 & 0 & 0 \end{pmatrix}.$$

State probabilities

After 1 step, the state is random: it can be either B with probability 0.4 or C with probability 0.6, so the state probability vector after 1 step is

$$v_1 = \begin{pmatrix} 0 & 0.4 & 0.6 & 0 & 0 \end{pmatrix}.$$

In 2 steps, the Markov chain can go:

- A$\to$B$\to$A with probability $0.4 \cdot 0.1$, or
- A$\to$B$\to$C with probability $0.4 \cdot 0.2$, or
- A$\to$B$\to$D with probability $0.4 \cdot 0.7$, or
- A$\to$C$\to$E with probability $0.6 \cdot 1$.

So

$$v_2 = \begin{pmatrix} 0.04 & 0 & 0.08 & 0.28 & 0.60 \end{pmatrix}.$$

How can we compute v_n, the state probability vector in general?

How can we compute v_n, the state probability vector in general?

Lemma

$$v_n = v_0 \cdot P^n$$

State probabilities

How can we compute v_n, the state probability vector in general?

Lemma

$$v_n = v_0 \cdot P^n$$

Proof (sketch). Apply induction on n; the vector-matrix product $v_{n-1} \cdot P$ computes the probability of being in each state after n steps by applying total probability according to the state after $n-1$ steps.

State probabilities

For example, from the initial vector

$$v_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and transition probability matrix

$$P = \begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0.6 & 0 & 0.4 & 0 \end{bmatrix}$$

we have

$$v_1 = v_0 \cdot P = \begin{pmatrix} 0 & 0.4 & 0.6 & 0 & 0 \end{pmatrix}.$$

and

$$v_2 = v_1 \cdot P = v_0 \cdot P^2 = \begin{pmatrix} 0.04 & 0 & 0.08 & 0.28 & 0.60 \end{pmatrix}.$$

State probabilities

$$\begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0.6 & 0 & 0.4 & 0 \end{bmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0.4 & 0.6 & 0 & 0 \end{pmatrix}$$

State probabilities

$$\begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0.6 & 0 & 0.4 & 0 \end{bmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0.4 & 0.6 & 0 & 0 \end{pmatrix}$$

$$\begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0.6 & 0 & 0.4 & 0 \end{bmatrix}$$

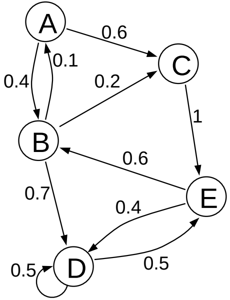
$$\begin{pmatrix} 0 & 0.4 & 0.6 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0.04 & 0 & 0.08 & 0.28 & 0.60 \end{pmatrix}$$

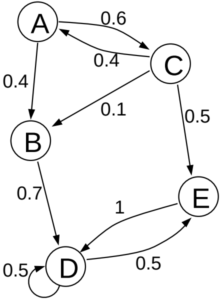
Communicating classes

If for two states, the Markov chain can get from the first state to the second (in a finite number of steps with positive probability), and can also get from the second state to the first state (in a finite number of steps with positive probability), then the two states belong to the same *communicating class*.

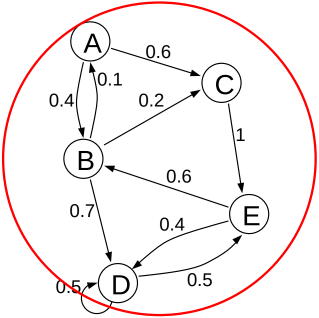
Communicating classes

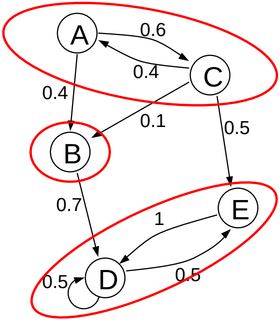
If for two states, the Markov chain can get from the first state to the second (in a finite number of steps with positive probability), and can also get from the second state to the first state (in a finite number of steps with positive probability), then the two states belong to the same *communicating class*.





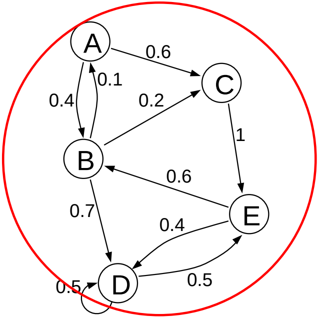
If for two states, the Markov chain can get from the first state to the second (in a finite number of steps with positive probability), and can also get from the second state to the first state (in a finite number of steps with positive probability), then the two states belong to the same *communicating class*.

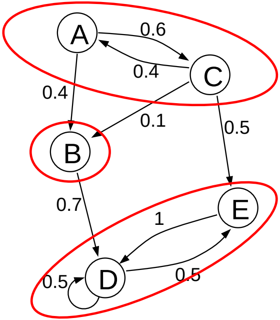




Irreducible and reducible Markov chains

If a Markov chain has a single communicating class, it is called *irreducible*. If it has multiple communicating classes, it is called *reducible*.

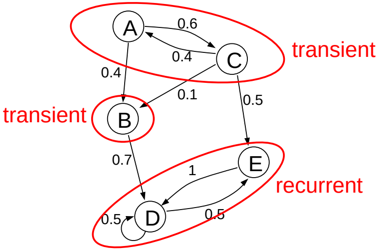




Recurrent and transient classes

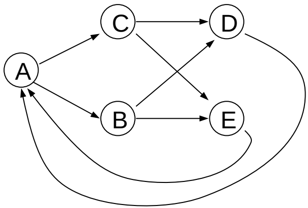
Communicating classes with no transitions leading out of the class are called *recurrent classes*, and states inside recurrent classes are *recurrent states*.

Communicating classes with at least one transition leading out of the class are called *transient classes*, and states inside transient classes are *transient states*.

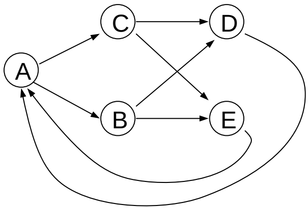


Periodicity

Is this Markov chain irreducible?



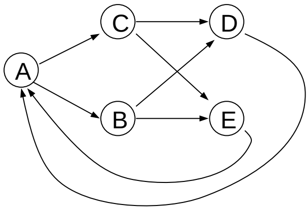
Is this Markov chain irreducible?



Yes, it is. We can get from any state to any other state.

Periodicity

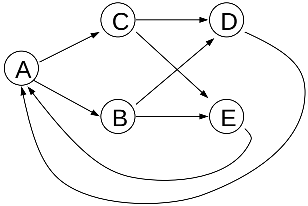
Is this Markov chain irreducible?



Yes, it is. We can get from any state to any other state.

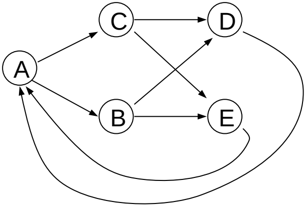
How many steps does it take to get from state A back to state A?

Periodicity



We can get from state A back to state A in 3, 6, 9, . . . steps.

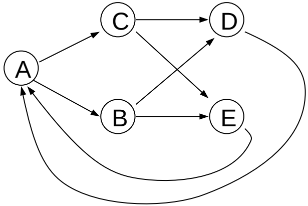
Periodicity



We can get from state A back to state A in 3, 6, 9,. . . steps.

From state B we can get back to state B only in 3, 6, 9,. . . steps, too.

Periodicity

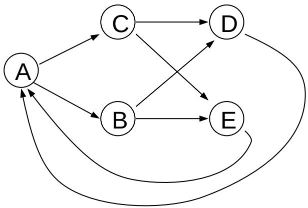


We can get from state A back to state A in 3, 6, 9,. . . steps.

From state B we can get back to state B only in 3, 6, 9,. . . steps, too.

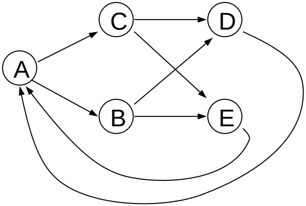
And this holds for any state.

Periodicity



If in an irreducible Markov chain we can only get from state A back to state A in a number of steps divisible by some $d > 1$, then we say that the Markov chain is *periodic with period* d. The value d is the same for all other states too.

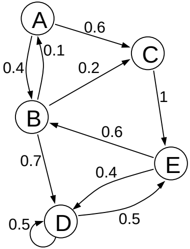
Periodicity



If in an irreducible Markov chain we can only get from state A back to state A in a number of steps divisible by some $d > 1$, then we say that the Markov chain is *periodic with period d*. The value d is the same for all other states too.

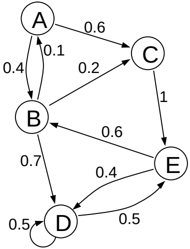
The above Markov chain is periodic with period 3.

Periodicity



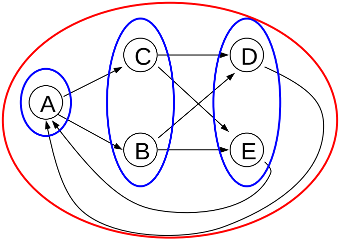
In this Markov chain, we can get from state D back to state D 1 step, or 2 steps, or 3, etc. This is an *aperiodic* Markov chain.

Periodicity



In this Markov chain, we can get from state D back to state D 1 step, or 2 steps, or 3, etc. This is an *aperiodic* Markov chain.

(We can also think of it as a Markov chain with $d = 1$.)



In a periodic Markov chain, we can divide the states into d *periodicity classes*. These are different from communicating classes! The above Markov chain is irreducible, it has one communicating class (containing every state) and 3 periodicity classes.

Example

In London, each rainy day is followed by a rainy day with probability 70% and by a sunny day with probability 30%. A sunny day is followed by a rainy day with probability 50% and by a sunny day with probability 50%. Assuming it is raining today, what is the probability that it will be raining 2 days from now? And 3 days from now?

Example

In London, each rainy day is followed by a rainy day with probability 70% and by a sunny day with probability 30%. A sunny day is followed by a rainy day with probability 50% and by a sunny day with probability 50%. Assuming it is raining today, what is the probability that it will be raining 2 days from now? And 3 days from now?

(This is not a perfect model as weather typically has some long term behaviour, but we will work with this now.)

Example

In London, each rainy day is followed by a rainy day with probability 70% and by a sunny day with probability 30%. A sunny day is followed by a rainy day with probability 50% and by a sunny day with probability 50%. Assuming it is raining today, what is the probability that it will be raining 2 days from now? And 3 days from now?

(This is not a perfect model as weather typically has some long term behaviour, but we will work with this now.)

There are 2 states: rainy and sunny. The probability transition matrix is

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix}$$

Example

In London, each rainy day is followed by a rainy day with probability 70% and by a sunny day with probability 30%. A sunny day is followed by a rainy day with probability 50% and by a sunny day with probability 50%. Assuming it is raining today, what is the probability that it will be raining 2 days from now? And 3 days from now?

(This is not a perfect model as weather typically has some long term behaviour, but we will work with this now.)

There are 2 states: rainy and sunny. The probability transition matrix is

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix}$$

Today it is raining, so

$$v_0 = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Then

$$v_1 = v_0 \cdot P = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix}.$$

Then

$$v_1 = v_0 \cdot P = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix}.$$

$$v_2 = v_1 \cdot P = \begin{pmatrix} 0.64 & 0.36 \end{pmatrix}.$$

Example

Then

$$v_1 = v_0 \cdot P = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix}.$$

$$v_2 = v_1 \cdot P = \begin{pmatrix} 0.64 & 0.36 \end{pmatrix}.$$

$$v_3 = v_1 \cdot P = \begin{pmatrix} 0.628 & 0.372 \end{pmatrix}.$$

Example

Then

$$v_1 = v_0 \cdot P = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix}.$$

$$v_2 = v_1 \cdot P = \begin{pmatrix} 0.64 & 0.36 \end{pmatrix}.$$

$$v_3 = v_1 \cdot P = \begin{pmatrix} 0.628 & 0.372 \end{pmatrix}.$$

They seem to converge. We aim to identify the limit next.

Stationary vector

$v_{\mathrm{st}} = \begin{pmatrix} x_1 & x_2 & \dots & x_k \end{pmatrix}$ is a *stationary vector* or *stationary distribution* for P if

$$v_{\mathrm{st}} \cdot P = v_{\mathrm{st}},$$

$x_i \geq 0$ $(i = 1, \dots, k)$ and

$$x_1 + \dots + x_k = 1.$$

Stationary vector

$v_{\mathrm{st}} = \begin{pmatrix} x_1 & x_2 & \dots & x_k \end{pmatrix}$ is a *stationary vector* or *stationary distribution* for P if

$$v_{\mathrm{st}} \cdot P = v_{\mathrm{st}},$$

$x_i \geq 0$ $(i = 1, \dots, k)$ and

$$x_1 + \dots + x_k = 1.$$

If we start the Markov chain from $v_0 = v_{\mathrm{st}}$, then

$$v_1 = v_{\mathrm{st}} \cdot P = v_{\mathrm{st}}$$

and so on, so

$$v_n = v_{\mathrm{st}} \quad \forall n.$$

A Markov chain started from a stationary vector is called a *stationary Markov chain*.

Long-term behaviour

Theorem (Perron–Frobenius)

(a) *There is always at least one stationary vector for any (finite state) Markov chain.*

(b) *If the Markov chain is irreducible, then v_{st} is unique and its elements are strictly positive.*

(c) *If the Markov chain is irreducible and aperiodic, then v_{st} is unique, its elements are strictly positive and*

$$\lim_{n\to\infty} v_n = v_{st}$$

for any v_0 initial vector.

Theorem (Perron–Frobenius)

(a) *There is always at least one stationary vector for any (finite state) Markov chain.*

(b) *If the Markov chain is irreducible, then v_{st} is unique and its elements are strictly positive.*

(c) *If the Markov chain is irreducible and aperiodic, then v_{st} is unique, its elements are strictly positive and*

$$\lim_{n \to \infty} v_n = v_{st}$$

for any v_0 initial vector.

No proof, but the stationary vector is the left-eigenvector corresponding to the dominant eigenvalue of P which is 1.

Example

If

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix},$$

then for $v_{\mathrm{st}} = \begin{pmatrix} x_1 & x_2 \end{pmatrix}$ the definition of the stationary vector gives

$$\begin{aligned} 0.7x_1 + 0.5x_2 &= x_1 \\ 0.3x_1 + 0.5x_2 &= x_2 \\ x_1 + x_2 &= 1, \end{aligned}$$

Example

If

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix},$$

then for $v_{\mathrm{st}} = \begin{pmatrix} x_1 & x_2 \end{pmatrix}$ the definition of the stationary vector gives

$$\begin{aligned} 0.7x_1 + 0.5x_2 &= x_1 \\ 0.3x_1 + 0.5x_2 &= x_2 \\ x_1 + x_2 &= 1, \end{aligned}$$

whose solution is $x_1 = 5/8 = 0.625$ and $x_2 = 3/8 = 0.375$, so

$$v_{\mathrm{st}} = \begin{pmatrix} 0.625 & 0.375 \end{pmatrix}.$$

Computing the stationary distribution

In general, the stationary distribution can be computed by solving the linear system of equations

$$v_{\mathrm{st}} \cdot P = v_{\mathrm{st}}\,,$$
$$x_1 + \cdots + x_k = 1.$$

In general, the stationary distribution can be computed by solving the linear system of equations

$$v_{\mathrm{st}} \cdot P = v_{\mathrm{st}},$$
$$x_1 + \cdots + x_k = 1.$$

The general method to do this is Gaussian elimination.

Computing the stationary distribution

In general, the stationary distribution can be computed by solving the linear system of equations

$$v_{\mathrm{st}} \cdot P = v_{\mathrm{st}},$$
$$x_1 + \cdots + x_k = 1.$$

The general method to do this is Gaussian elimination.

In certain special cases, the solution may be easier. We will address this during problem solving.

So if the Markov chain is irreducible and aperiodic,

$$\lim_{n\to\infty} v_n = v_{\mathrm{st}}.$$

So if the Markov chain is irreducible and aperiodic,

$$\lim_{n\to\infty} v_n = v_{\mathrm{st}}.$$

The convergence is fast (exponentially fast – no proof); in general, already for $n \geq 10$,

$$v_n \approx v_{\mathrm{st}}$$

is a good approximation.

So if the Markov chain is irreducible and aperiodic,

$$\lim_{n\to\infty} v_n = v_{\mathrm{st}}.$$

The convergence is fast (exponentially fast – no proof); in general, already for $n \geq 10$,

$$v_n \approx v_{\mathrm{st}}$$

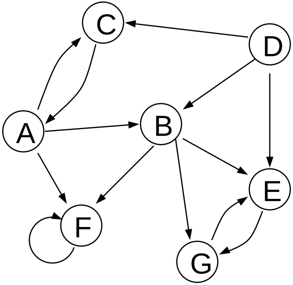
is a good approximation.

Also, part (c) of the Perron–Frobenius theorem is equivalent to

$$\lim_{n\to\infty} P^n = \begin{bmatrix} v_{\mathrm{st}} \\ \hline v_{\mathrm{st}} \\ \hline \vdots \\ \hline v_{\mathrm{st}} \end{bmatrix}.$$

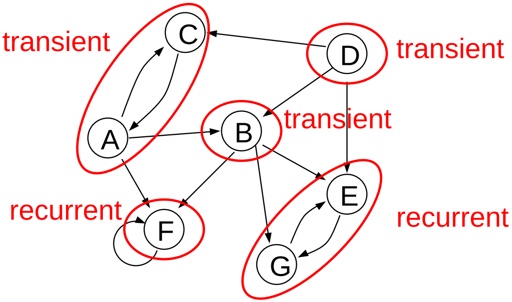
Reducible Markov chains

A reducible Markov chain will eventually end up in one of the recurrent classes.



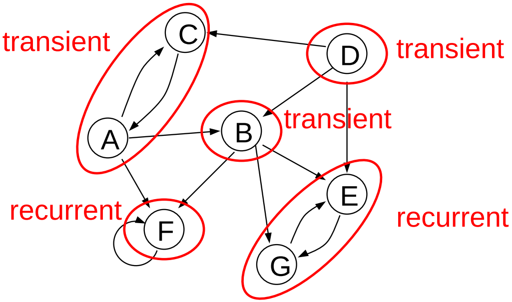
Reducible Markov chains

A reducible Markov chain will eventually end up in one of the recurrent classes.



Reducible Markov chains

Each recurrent class can be regarded as a smaller irreducible Markov chain.



Reducible Markov chains

For a reducible Markov chain, there is a stationary vector for each of the recurrent classes. For each class, the stationary vector is unique.

Reducible Markov chains

For a reducible Markov chain, there is a stationary vector for each of the recurrent classes. For each class, the stationary vector is unique.

Stationary vectors of the entire Markov chain are convex linear combinations of the stationary vectors of the classes.

Reducible Markov chains

For a reducible Markov chain, there is a stationary vector for each of the recurrent classes. For each class, the stationary vector is unique.

Stationary vectors of the entire Markov chain are convex linear combinations of the stationary vectors of the classes.

The probability of any transient state is 0 in any of the stationary vectors.

Reducible Markov chains

For a reducible Markov chain, there is a stationary vector for each of the recurrent classes. For each class, the stationary vector is unique.

Stationary vectors of the entire Markov chain are convex linear combinations of the stationary vectors of the classes.

The probability of any transient state is 0 in any of the stationary vectors.

Even for a reducible Markov chain, v_n will converge to one of the stationary vectors – but it can be random, which one.

Reducible Markov chains

For a reducible Markov chain, there is a stationary vector for each of the recurrent classes. For each class, the stationary vector is unique.

Stationary vectors of the entire Markov chain are convex linear combinations of the stationary vectors of the classes.

The probability of any transient state is 0 in any of the stationary vectors.

Even for a reducible Markov chain, v_n will converge to one of the stationary vectors – but it can be random, which one.

We will focus mostly on irreducible Markov chains.

Stationary vector interpretations

For an irreducible, aperiodic Markov chain,

$$v_n \to v_{\mathrm{st}}$$

for any v_0. This means that in the long run, the Markov chain "forgets" the initial state and will be very close to stationary. As a consequence, in many scenarios where a Markov chain has been running for a long time, the initial state is irrelevant, as the Markov chain will be stationary anyway.

Stationary vector interpretations

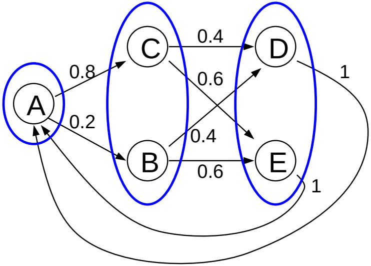
For an irreducible, aperiodic Markov chain,

$$v_n \to v_{\mathrm{st}}$$

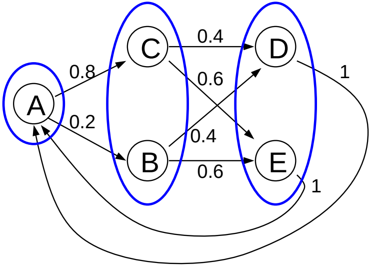
for any v_0. This means that in the long run, the Markov chain "forgets" the initial state and will be very close to stationary. As a consequence, in many scenarios where a Markov chain has been running for a long time, the initial state is irrelevant, as the Markov chain will be stationary anyway.

What about irreducible, periodic Markov chains?

Periodic Markov chains

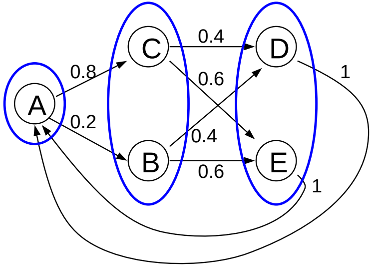


Periodic Markov chains



A possible realization: A, C, D, A, C, E, A, B, E, A, C, E, A, . . .

Periodic Markov chains



A possible realization: A, C, D, A, C, E, A, B, E, A, C, E, A, . . .

(Every third state is A, followed by either B or C, followed by either D or E.)

Periodic Markov chains

For an irreducible, periodic Markov chain, v_{st} is unique. For the above Markov chain,

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{3} & \frac{1}{15} & \frac{4}{15} & \frac{14}{75} & \frac{11}{75} \end{pmatrix}.$$

For an irreducible, periodic Markov chain, v_{st} is unique. For the above Markov chain,

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{3} & \frac{1}{15} & \frac{4}{15} & \frac{14}{75} & \frac{11}{75} \end{pmatrix}.$$

If the period is d, v_{st} assigns $\frac{1}{d}$ total weight to each of the periodicity classes:

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{3} & \underbrace{\frac{1}{15} \quad \frac{4}{15}}_{\frac{1}{3}} & \underbrace{\frac{14}{75} \quad \frac{11}{75}}_{\frac{1}{3}} \end{pmatrix}.$$

Periodic Markov chains

However, a periodic Markov chain does not “forget” the initial state.

Periodic Markov chains

However, a periodic Markov chain does not "forget" the initial state.

If the periodicity classes are denoted by $1, \ldots, d$, then if the Markov chain started from class 1, then after nd steps, it can only be in class 1 again. After $nd + 1$ steps, it can only be in class 2, and so on.

Periodic Markov chains

However, a periodic Markov chain does not "forget" the initial state.

If the periodicity classes are denoted by $1, \ldots, d$, then if the Markov chain started from class 1, then after nd steps, it can only be in class 1 again. After $nd + 1$ steps, it can only be in class 2, and so on.

The following is true: for large n, v_n is approximately equal to v_{st} *conditioned on being in the periodicity class where the Markov chain can be in after n steps.*

Periodic Markov chains

However, a periodic Markov chain does not "forget" the initial state.

If the periodicity classes are denoted by $1, \ldots, d$, then if the Markov chain started from class 1, then after nd steps, it can only be in class 1 again. After $nd + 1$ steps, it can only be in class 2, and so on.

The following is true: for large n, v_n is approximately equal to v_{st} *conditioned on being in the periodicity class where the Markov chain can be in after n steps.*

This conditional distribution can be computed by replacing all elements of $d \times v_{\mathrm{st}}$ by zeros except for one periodicity class.

Periodic Markov chains

For the previous periodic Markov chain example with period 3, we have

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{3} & \frac{1}{15} & \frac{4}{15} & \frac{14}{75} & \frac{11}{75} \end{pmatrix}.$$

If the Markov chain started from state 1, then after $3n$ steps, it can only be in class 1, and so

$$v_{3n} \approx 3 \cdot \begin{pmatrix} \frac{1}{3} & \cancelto{0}{\frac{1}{15}} & \cancelto{0}{\frac{4}{15}} & \cancelto{0}{\frac{14}{75}} & \cancelto{0}{\frac{11}{75}} \end{pmatrix} = (1\ 0\ 0\ 0\ 0).$$

After $3n+1$ steps, it can only be in class 2, so

$$v_{3n+1} \approx 3 \cdot \begin{pmatrix} \cancelto{0}{\frac{1}{3}} & \frac{1}{15} & \frac{4}{15} & \cancelto{0}{\frac{14}{75}} & \cancelto{0}{\frac{11}{75}} \end{pmatrix} = (0\ \tfrac{1}{5}\ \tfrac{4}{5}\ 0\ 0).$$

Periodic Markov chains

And after $3n+2$ steps, it can only be in class 3, so

$$v_{3n+2} \approx 3 \cdot \begin{pmatrix} \cancelto{0}{\frac{1}{3}} & \cancelto{0}{\frac{1}{15}} & \cancelto{0}{\frac{4}{15}} & \frac{14}{75} & \frac{11}{75} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \frac{14}{25} & \frac{11}{25} \end{pmatrix}.$$

Periodic Markov chains

And after $3n+2$ steps, it can only be in class 3, so

$$v_{3n+2} \approx 3 \cdot \left(\overset{0}{\cancel{\frac{1}{3}}} \;\; \overset{0}{\cancel{\frac{1}{15}}} \;\; \overset{0}{\cancel{\frac{4}{15}}} \;\; \frac{14}{75} \;\; \frac{11}{75} \right) = \left(0 \; 0 \; 0 \; \frac{14}{25} \; \frac{11}{25} \right).$$

If the Markov chain started from class 2, then after $3n$ steps, it can only be in class 2 again, so in this case,

$$v_{3n} \approx 3 \cdot \left(\overset{0}{\cancel{\frac{1}{3}}} \;\; \frac{1}{15} \;\; \frac{4}{15} \;\; \overset{0}{\cancel{\frac{14}{75}}} \;\; \overset{0}{\cancel{\frac{11}{75}}} \right) = \left(0 \; \frac{1}{5} \; \frac{4}{5} \; 0 \; 0 \right).$$

Similarly,

$$v_{3n+1} \approx 3 \cdot \left(\overset{0}{\cancel{\frac{1}{3}}} \;\; \overset{0}{\cancel{\frac{1}{15}}} \;\; \overset{0}{\cancel{\frac{4}{15}}} \;\; \frac{14}{75} \;\; \frac{11}{75} \right) = \left(0 \; 0 \; 0 \; \frac{14}{25} \; \frac{11}{25} \right),$$

etc.; the approximations are shifted according to the initial class.

Stationary vector interpretations

What else does the stationary vector tell? Use the notation

$$v_{\mathrm{st}} = (\, x_1 \; x_2 \; \dots \; x_k \,).$$

Stationary vector interpretations

What else does the stationary vector tell? Use the notation

$$v_{\mathrm{st}} = (\, x_1 \; x_2 \; \ldots \; x_k \,).$$

Lemma

For an irreducible Markov chain, the ratio of time spent at state i in the long run is x_i.

Stationary vector interpretations

What else does the stationary vector tell? Use the notation

$$v_{\mathrm{st}} = (\, x_1 \; x_2 \; \ldots \; x_k \,).$$

Lemma

For an irreducible Markov chain, the ratio of time spent at state i in the long run is x_i.

Lemma

For an irreducible Markov chain, the average number of steps taken between two visits to state i is $\frac{1}{x_i}$.

Both lemmas are valid for both periodic and aperiodic Markov chains.

Stationary vector interpretations

For the London weather example, let us denote the states as 1: rainy, 2: sunny. Then a possible realization is

$$1, 1, 2, 1, 1, 1, 2, 2, 1, 2, 1, 1, 2, 1, 1, 2, \ldots$$

Stationary vector interpretations

For the London weather example, let us denote the states as 1: rainy, 2: sunny. Then a possible realization is

$$1, 1, 2, 1, 1, 1, 2, 2, 1, 2, 1, 1, 2, 1, 1, 2, \ldots$$

The first lemma states that the ratio of 1's, that is, the ratio of rainy days in the long run will be $x_1 = 0.625 = 62.5\%$ of the total days, while the ratio of sunny days will be $x_2 = 0.375 = 37.5\%$.

Stationary vector interpretations

For the London weather example, let us denote the states as 1: rainy, 2: sunny. Then a possible realization is

$$1, 1, 2, 1, 1, 1, 2, 2, 1, 2, 1, 1, 2, 1, 1, 2, \ldots$$

The first lemma states that the ratio of 1's, that is, the ratio of rainy days in the long run will be $x_1 = 0.625 = 62.5\%$ of the total days, while the ratio of sunny days will be $x_2 = 0.375 = 37.5\%$.

The second lemma states that the average number of steps between consecutive 1's, that is, the average number of days between consecutive rainy days, is $\frac{1}{x_1} = \frac{1}{0.625} = 1.6$, and the average number of steps between consecutive 2's (or the average number of days between consecutive sunny days) is $\frac{1}{x_2} = \frac{1}{0.375} \approx 2.667$.

Ergodic theorem

The next theorem is essentially the law of large numbers for irreducible Markov chains. Let the states of the MC be $\{1, \ldots, k\}$.

Ergodic theorem

The next theorem is essentially the law of large numbers for irreducible Markov chains. Let the states of the MC be $\{1, \ldots, k\}$.

Theorem (Ergodic theorem)

Denote the realization of a Markov chain by the sequence of states

$$X_1, X_2, \ldots$$

If the Markov chain is irreducible, then for any function f given on the states,

$$\lim_{n\to\infty} \frac{f(X_1) + \cdots + f(X_n)}{n} = \mathbb{E}_{st}(f),$$

where

$$\mathbb{E}_{st}(f) = x_1 f(1) + \cdots + x_k f(k),$$

where

$$v_{st} = \begin{pmatrix} x_1 & \ldots & x_k \end{pmatrix}.$$

Ergodic theorem

An ice cream seller in London is selling on average 800 pounds of ice cream per day on a sunny day, but only 120 pounds on average on a rainy day. What is his long term average daily income?

Ergodic theorem

An ice cream seller in London is selling on average 800 pounds of ice cream per day on a sunny day, but only 120 pounds on average on a rainy day. What is his long term average daily income?

Denoting the states of the Markov chain with 1: rainy and 2 sunny, define the function f to be

$$f(1) = 120, \qquad f(2) = 800.$$

An ice cream seller in London is selling on average 800 pounds of ice cream per day on a sunny day, but only 120 pounds on average on a rainy day. What is his long term average daily income?

Denoting the states of the Markov chain with 1: rainy and 2 sunny, define the function f to be

$$f(1) = 120, \qquad f(2) = 800.$$

Then the ergodic theorem states that the long term average daily income is

$$\mathbb{E}_{\mathrm{st}}(f) = x_1 f(1) + x_2 f(2) = 0.625 \cdot 120 + 0.375 \cdot 800 = 375$$

(pounds per day).

Central limit theorem for Markov chains

The ergodic theorem is essentially the law of large numbers for Markov chains.

Central limit theorem for Markov chains

The ergodic theorem is essentially the law of large numbers for Markov chains.

So does the central limit theorem hold, too?

Central limit theorem for Markov chains

The ergodic theorem is essentially the law of large numbers for Markov chains.

So does the central limit theorem hold, too?

Actually it does! That is,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{f(X_1)+\cdots+f(X_n)-n\cdot\mathbb{E}_{\mathrm{st}}(f)}{\sigma(f)\sqrt{n}} < x\right) = \Phi(x) \quad \forall x \in \mathbb{R}.$$

Central limit theorem for Markov chains

The ergodic theorem is essentially the law of large numbers for Markov chains.

So does the central limit theorem hold, too?

Actually it does! That is,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{f(X_1) + \cdots + f(X_n) - n \cdot \mathbb{E}_{\mathrm{st}}(f)}{\sigma(f)\sqrt{n}} < x \right) = \Phi(x) \quad \forall x \in \mathbb{R}.$$

However, there is one difficulty: the value of $\sigma(f)$ is difficult to compute.

The ergodic theorem is essentially the law of large numbers for Markov chains.

So does the central limit theorem hold, too?

Actually it does! That is,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{f(X_1)+\cdots+f(X_n)-n\cdot\mathbb{E}_{\mathrm{st}}(f)}{\sigma(f)\sqrt{n}}<x\right)=\Phi(x) \quad \forall x\in\mathbb{R}.$$

However, there is one difficulty: the value of $\sigma(f)$ is difficult to compute.

Overall, the Markov chain CLT is not very practical and we are not going to use it.

Summary

- There is always at least one v_{st}.
- For an irreducible Markov chain, $v_{\mathrm{st}} = (x_1 \ \ldots \ x_k)$ is unique and $x_i > 0 \, \forall i$.
- A reducible Markov chain will eventually end up in one the recurrent classes, which can be viewed as a smaller Markov chain itself.
- For an irreducible, aperiodic Markov chain, $v_n \to v_{\mathrm{st}}$ rapidly.
- For an irreducible, periodic Markov chain, v_n changes periodically, and can be approximated by v_{st} conditioned on the corresponding periodicity class.
- For irreducible Markov chains, the long term average ratio of time spent in state i is x_i, and the average number of steps spent between returns to state i is $\frac{1}{x_i}$.
- Ergodic theorem: the long term average of functions converges to $\mathbb{E}_{\mathrm{st}}(f)$, the stationary expected value of the function.

Infinite state space

For Markov chains in an infinite state space, a wide variety of behaviour is possible.

- There are Markov chains with a single v_{st} and $v_n \to v_{\mathrm{st}}$ for any v_0; essentially, all theorems for finite state space Markov chains are valid. These are sometimes called stable Markov chains.

For Markov chains in an infinite state space, a wide variety of behaviour is possible.

- There are Markov chains with a single v_{st} and $v_n \to v_{\mathrm{st}}$ for any v_0; essentially, all theorems for finite state space Markov chains are valid. These are sometimes called stable Markov chains.
- There are Markov chains with no v_{st} (all states are transient).

For Markov chains in an infinite state space, a wide variety of behaviour is possible.

- There are Markov chains with a single v_{st} and $v_n \to v_{\mathrm{st}}$ for any v_0; essentially, all theorems for finite state space Markov chains are valid. These are sometimes called stable Markov chains.
- There are Markov chains with no v_{st} (all states are transient).
- There are Markov chains with infinitely many v_{st}'s where v_n might be divergent for some choices of v_0.

Infinite state space

For Markov chains in an infinite state space, a wide variety of behaviour is possible.

- There are Markov chains with a single v_{st} and $v_n \to v_{\mathrm{st}}$ for any v_0; essentially, all theorems for finite state space Markov chains are valid. These are sometimes called stable Markov chains.
- There are Markov chains with no v_{st} (all states are transient).
- There are Markov chains with infinitely many v_{st}'s where v_n might be divergent for some choices of v_0.
- Apart from transient and recurrent states, there may be null recurrent states, at which the Markov chain returns infinitely many times, but the expected number of steps between two returns is infinite.

For Markov chains in an infinite state space, a wide variety of behaviour is possible.

- There are Markov chains with a single v_{st} and $v_n \to v_{\mathrm{st}}$ for any v_0; essentially, all theorems for finite state space Markov chains are valid. These are sometimes called stable Markov chains.
- There are Markov chains with no v_{st} (all states are transient).
- There are Markov chains with infinitely many v_{st}'s where v_n might be divergent for some choices of v_0.
- Apart from transient and recurrent states, there may be null recurrent states, at which the Markov chain returns infinitely many times, but the expected number of steps between two returns is infinite.

In general, spectral properties of the generator may help decide which of the above cases holds.

For Markov chains in an infinite state space, a wide variety of behaviour is possible.

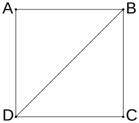
- There are Markov chains with a single v_{st} and $v_n \to v_{\mathrm{st}}$ for any v_0; essentially, all theorems for finite state space Markov chains are valid. These are sometimes called stable Markov chains.
- There are Markov chains with no v_{st} (all states are transient).
- There are Markov chains with infinitely many v_{st}'s where v_n might be divergent for some choices of v_0.
- Apart from transient and recurrent states, there may be null recurrent states, at which the Markov chain returns infinitely many times, but the expected number of steps between two returns is infinite.

In general, spectral properties of the generator may help decide which of the above cases holds.

We will discuss a few infinite state space Markov chains - for continuous time Markov chains.

Problem 1

A drunk man is walking around in a small town. The map of the town is the following:



Whenever the man arrives at any of the corners (A, B, C or D), he will choose his next destination randomly from among the streets available, except the street where he just arrived from.
Is the sequence of corners he visits a Markov chain? If not, propose a Markov chain that describes the situation.

Problem 1

Solution.
If we only know that right now he is in corner B, it does not fully describe where he can go next: if he arrived from A to B in the previous step, then he can go to C or D next, but if he arrived from C, he can go to A or B next.

Solution.
If we only know that right now he is in corner B, it does not fully describe where he can go next: if he arrived from A to B in the previous step, then he can go to C or D next, but if he arrived from C, he can go to A or B next.

If we include the previous corner in the state too, then it will be a Markov chain. The states are AB, AD, BA, BC, BD, CB, CD, DA, DB, DC. For example, AB means that right now he is in B, coming from A in the previous step.

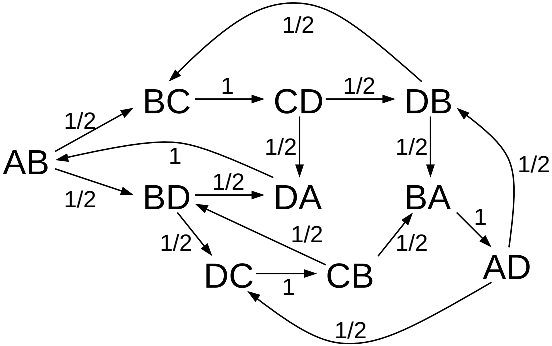
Problem 1

Solution.
If we only know that right now he is in corner B, it does not fully describe where he can go next: if he arrived from A to B in the previous step, then he can go to C or D next, but if he arrived from C, he can go to A or B next.

If we include the previous corner in the state too, then it will be a Markov chain. The states are AB, AD, BA, BC, BD, CB, CD, DA, DB, DC. For example, AB means that right now he is in B, coming from A in the previous step. From state AB, the Markov chain can transition to either BD or BC with probability $1/2$ each.

Problem 1

The graph of the entire Markov chain:



Problem 2

Electric Ltd. takes two types of contract jobs: A and B. A type A job lasts for one month and their income from it is 1.4 million HUF, while a type B job lasts for 2 months and their income is 2.7 million HUF. At the beginning of each month, they are open to new contract offers unless they are in the middle of a type B job.
At the beginning of each month, they will receive a contract offer for a type B job with 60% probability, while they will receive a contract offer for a type A job with 50% probability (independently from type B offers). If they receive both types of offers, they accept a type A offer.

Problem 2

(a) Model the monthly activity of Electric Ltd. with a Markov-chain. What are the states? What is the transition matrix? Is the Markov chain irreducible? Is it aperiodic?

(b) Calculate the stationary distribution. Based on the stationary distribution, calculate the long-term average monthly income.

(c) What is the average amount of time between consecutive idle months?

(d) They are reconsidering their policy to accept a type A offer when both are available. What is their long-term average monthly income in case they prefer a type B offer when both A and B are available?

Problem 2

To determine the states, let's look at the possible activities of Electric Ltd. They may be doing a type A job, or a type B job, or they can be idle, which will be denoted by 0.

Problem 2

To determine the states, let's look at the possible activities of Electric Ltd. They may be doing a type A job, or a type B job, or they can be idle, which will be denoted by 0.

A possible realization is:

$$A, B, B, 0, B, B, A, 0, A, A, B, B, B, B, A, B, B, \ldots$$

Problem 2

To determine the states, let's look at the possible activities of Electric Ltd. They may be doing a type A job, or a type B job, or they can be idle, which will be denoted by 0.

A possible realization is:

$$A,\ B,\ B,\ 0,\ B,\ B,\ A,\ 0,\ A,\ A,\ B,\ B,\ B,\ B,\ A,\ B,\ B, \ldots$$

Is this a Markov chain?

Problem 2

To determine the states, let's look at the possible activities of Electric Ltd. They may be doing a type A job, or a type B job, or they can be idle, which will be denoted by 0.

A possible realization is:

$$A, B, B, 0, B, B, A, 0, A, A, B, B, B, B, A, B, B, \ldots$$

Is this a Markov chain?

No, it is not! The issue is that if right now they are doing a type B job, what can happen next month depends on whether it's the first half or second half of a type B job: from the first half of a type B job, they will always go to the second half, while after the second half of a type B job, in the next month they can be idle, or do a type A job, or start another type B job.

Problem 2

A possible solution is to consider the first and second month of a type B job as separate states B1 and B2. Then the previous realization looks like this:

A, B1, B2, 0, B1, B2, A, 0, A, A, B1, B2, B1, B2, A, B1, B2,. . .

Problem 2

A possible solution is to consider the first and second month of a type B job as separate states B1 and B2. Then the previous realization looks like this:

A, B1, B2, 0, B1, B2, A, 0, A, A, B1, B2, B1, B2, A, B1, B2,. . .

This is now a Markov chain on the states A, B1, B2, 0.

Problem 2

A possible solution is to consider the first and second month of a type B job as separate states B1 and B2. Then the previous realization looks like this:

A, B1, B2, 0, B1, B2, A, 0, A, A, B1, B2, B1, B2, A, B1, B2,...

This is now a Markov chain on the states A, B1, B2, 0. Let's calculate the transition probabilities from state A. They will finish the type A job at the end of the month, so for the next month, they will be open to contract offers.

Problem 2

A possible solution is to consider the first and second month of a type B job as separate states B1 and B2. Then the previous realization looks like this:

A, B1, B2, 0, B1, B2, A, 0, A, A, B1, B2, B1, B2, A, B1, B2,. . .

This is now a Markov chain on the states A, B1, B2, 0. Let's calculate the transition probabilities from state A. They will finish the type A job at the end of the month, so for the next month, they will be open to contract offers.

- If they receive an offer for a type A job, they will take it; this has probability 0.5.

Problem 2

A possible solution is to consider the first and second month of a type B job as separate states B1 and B2. Then the previous realization looks like this:

A, B1, B2, 0, B1, B2, A, 0, A, A, B1, B2, B1, B2, A, B1, B2,...

This is now a Markov chain on the states A, B1, B2, 0. Let's calculate the transition probabilities from state A. They will finish the type A job at the end of the month, so for the next month, they will be open to contract offers.

- If they receive an offer for a type A job, they will take it; this has probability 0.5.
- They will start a type B job if they receive a type B offer *and* do not receive a type A offer; this has probability $0.6 \cdot (1 - 0.5) = 0.3$.

Problem 2

A possible solution is to consider the first and second month of a type B job as separate states B1 and B2. Then the previous realization looks like this:

A, B1, B2, 0, B1, B2, A, 0, A, A, B1, B2, B1, B2, A, B1, B2,. . .

This is now a Markov chain on the states A, B1, B2, 0. Let's calculate the transition probabilities from state A. They will finish the type A job at the end of the month, so for the next month, they will be open to contract offers.

- If they receive an offer for a type A job, they will take it; this has probability 0.5.
- They will start a type B job if they receive a type B offer *and* do not receive a type A offer; this has probability $0.6 \cdot (1 - 0.5) = 0.3$.
- They will be idle if they do not receive either a type A or a type B offer; this has probability $(1 - 0.5)(1 - 0.6) = 0.2$.

Problem 2

The probability transition matrix is

	A	$B1$	$B2$	0
A	0.5	0.3	0	0.2
$B1$	0	0	1	0
$B2$	0.5	0.3	0	0.2
0	0.5	0.3	0	0.2

Problem 2

The probability transition matrix is

	A	$B1$	$B2$	0
A	0.5	0.3	0	0.2
$B1$	0	0	1	0
$B2$	0.5	0.3	0	0.2
0	0.5	0.3	0	0.2

The Markov chain may get from any state to any state, so it is irreducible.

Problem 2

The probability transition matrix is

	A	$B1$	$B2$	0
A	0.5	0.3	0	0.2
$B1$	0	0	1	0
$B2$	0.5	0.3	0	0.2
0	0.5	0.3	0	0.2

The Markov chain may get from any state to any state, so it is irreducible.

Quick rule of thumb to help decide periodicity: if the matrix P has at least one strictly positive element in the diagonal (which corresponds to a loop), then the Markov chain must be aperiodic.

The converse is not true.

Problem 2

The stationary distribution $v_{\mathrm{st}} = \left(x_1 \; x_2 \; x_3 \; x_4 \right)$ can be computed from

$$\begin{aligned} v_{\mathrm{st}} \cdot P &= v_{\mathrm{st}}, \\ x_1 + x_2 + x_3 + x_4 &= 1, \end{aligned}$$

from which

$$\begin{aligned} 0.5x_1 + 0.5x_2 + 0.5x_4 &= x_1, \\ 0.3x_1 + 0.3x_2 + 0.3x_4 &= x_2, \\ x_2 &= x_3, \\ 0.2x_1 + 0.2x_2 + 0.2x_4 &= x_4, \\ x_1 + x_2 + x_3 + x_4 &= 1. \end{aligned}$$

Problem 2

From the first four equations, the ratios of x_1, x_2, x_3, x_4 relative to each other can be expressed directly as

$$x_1 : x_2 : x_3 : x_4 = 5 : 3 : 3 : 2,$$

Problem 2

From the first four equations, the ratios of x_1, x_2, x_3, x_4 relative to each other can be expressed directly as

$$x_1 : x_2 : x_3 : x_4 = 5 : 3 : 3 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \left(\frac{5}{13} \; \frac{3}{13} \; \frac{3}{13} \; \frac{2}{13} \right).$$

From the first four equations, the ratios of x_1, x_2, x_3, x_4 relative to each other can be expressed directly as

$$x_1 : x_2 : x_3 : x_4 = 5 : 3 : 3 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{5}{13} & \frac{3}{13} & \frac{3}{13} & \frac{2}{13} \end{pmatrix}.$$

Their long-term average monthly income can be computed from the ergodic theorem. Note that to apply the ergodic theorem, we need to assign a value to each state. For A, their income is 1.4 (million HUF), for state 0, their income is 0, and for states B1 and B2, we need to split the 2.7.

Problem 2

From the first four equations, the ratios of x_1, x_2, x_3, x_4 relative to each other can be expressed directly as

$$x_1 : x_2 : x_3 : x_4 = 5 : 3 : 3 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{5}{13} & \frac{3}{13} & \frac{3}{13} & \frac{2}{13} \end{pmatrix}.$$

Their long-term average monthly income can be computed from the ergodic theorem. Note that to apply the ergodic theorem, we need to assign a value to each state. For A, their income is 1.4 (million HUF), for state 0, their income is 0, and for states B1 and B2, we need to split the 2.7.
If we split it evenly, then their long-term average monthly income is

$$\frac{5}{13} \cdot 1.4 + \frac{3}{13} \cdot 1.35 + \frac{3}{13} \cdot 1.35 + \frac{2}{13} \cdot 0 \approx 1.16.$$

Problem 2

If we split the income for a type B job in some other way, for example $2.7 + 0$, then their long-term average monthly income is

$$\frac{5}{13} \cdot 1.4 + \frac{3}{13} \cdot 2.7 + \frac{3}{13} \cdot 0 + \frac{2}{13} \cdot 0 \approx 1.16.$$

Problem 2

If we split the income for a type B job in some other way, for example $2.7 + 0$, then their long-term average monthly income is

$$\frac{5}{13} \cdot 1.4 + \frac{3}{13} \cdot 2.7 + \frac{3}{13} \cdot 0 + \frac{2}{13} \cdot 0 \approx 1.16.$$

In fact, the long-term average monthly income will be the same for any split due to $x_2 = x_3$.

If we split the income for a type B job in some other way, for example $2.7+0$, then their long-term average monthly income is

$$\frac{5}{13}\cdot 1.4+\frac{3}{13}\cdot 2.7+\frac{3}{13}\cdot 0+\frac{2}{13}\cdot 0\approx 1.16.$$

In fact, the long-term average monthly income will be the same for any split due to $x_2=x_3$.

Since the stationary probability of an idle month is $x_4=\frac{2}{13}$, the average amount of time between two consecutive idle months is

$$\frac{1}{2/13}=6.5$$

months.

Problem 2

If they decide to prefer type B jobs instead of type A jobs, then the transition probabilities are different.

- They will start a type A job if they receive a type A offer *and* do not receive a type B offer; this has probability $0.5 \cdot (1 - 0.6) = 0.2$.

Problem 2

If they decide to prefer type B jobs instead of type A jobs, then the transition probabilities are different.

- They will start a type A job if they receive a type A offer *and* do not receive a type B offer; this has probability $0.5 \cdot (1 - 0.6) = 0.2$.
- If they receive an offer for a type B job, they will take it; this has probability 0.6.

Problem 2

If they decide to prefer type B jobs instead of type A jobs, then the transition probabilities are different.

- They will start a type A job if they receive a type A offer *and* do not receive a type B offer; this has probability $0.5 \cdot (1 - 0.6) = 0.2$.
- If they receive an offer for a type B job, they will take it; this has probability 0.6.
- They will be idle if they do not receive either a type A or a type B offer; this has probability $(1 - 0.5)(1 - 0.6) = 0.2$.

Problem 2

If they decide to prefer type B jobs instead of type A jobs, then the transition probabilities are different.

- They will start a type A job if they receive a type A offer *and* do not receive a type B offer; this has probability $0.5 \cdot (1 - 0.6) = 0.2$.
- If they receive an offer for a type B job, they will take it; this has probability 0.6.
- They will be idle if they do not receive either a type A or a type B offer; this has probability $(1 - 0.5)(1 - 0.6) = 0.2$.

Accordingly,

$$P = \begin{bmatrix} 0.2 & 0.6 & 0 & 0.2 \\ 0 & 0 & 1 & 0 \\ 0.2 & 0.6 & 0 & 0.2 \\ 0.2 & 0.6 & 0 & 0.2 \end{bmatrix}.$$

Problem 2

The stationary distribution is different as well:

$$x_1 : x_2 : x_3 : x_4 = 2 : 6 : 6 : 2,$$

Problem 2

The stationary distribution is different as well:

$$x_1 : x_2 : x_3 : x_4 = 2 : 6 : 6 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \left(\frac{2}{16} \; \frac{6}{16} \; \frac{6}{16} \; \frac{2}{16} \right),$$

Problem 2

The stationary distribution is different as well:

$$x_1 : x_2 : x_3 : x_4 = 2 : 6 : 6 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \left(\frac{2}{16} \; \frac{6}{16} \; \frac{6}{16} \; \frac{2}{16} \right),$$

and according to the ergodic theorem, their long-term average monthly income is

$$\frac{2}{16} \cdot 1.4 + \frac{6}{16} \cdot 1.35 + \frac{6}{16} \cdot 1.35 + \frac{2}{16} \cdot 0 \approx 1.19.$$

Problem 2

The stationary distribution is different as well:

$$x_1 : x_2 : x_3 : x_4 = 2 : 6 : 6 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \left(\frac{2}{16} \; \frac{6}{16} \; \frac{6}{16} \; \frac{2}{16} \right),$$

and according to the ergodic theorem, their long-term average monthly income is

$$\frac{2}{16} \cdot 1.4 + \frac{6}{16} \cdot 1.35 + \frac{6}{16} \cdot 1.35 + \frac{2}{16} \cdot 0 \approx 1.19.$$

This is actually higher than when they prefer type A jobs, despite the fact that a type A job offers more income per month than a type B job.

Problem 2

The stationary distribution is different as well:

$$x_1 : x_2 : x_3 : x_4 = 2 : 6 : 6 : 2,$$

and accordingly,

$$v_{\mathrm{st}} = \left(\frac{2}{16} \; \frac{6}{16} \; \frac{6}{16} \; \frac{2}{16} \right),$$

and according to the ergodic theorem, their long-term average monthly income is

$$\frac{2}{16} \cdot 1.4 + \frac{6}{16} \cdot 1.35 + \frac{6}{16} \cdot 1.35 + \frac{2}{16} \cdot 0 \approx 1.19.$$

This is actually higher than when they prefer type A jobs, despite the fact that a type A job offers more income per month than a type B job. The reason is that the probability of idle months is now significantly lower.

Problem 4

A football association has 3 leagues. Pegleg FC starts from league 3. If they are currently in league 3, they get promoted with probability $2/3$ for the next season. From league 2, they get promoted with probability $1/2$ for the next season and get relegated with probability $1/6$ (otherwise, they remain in the current league). From league 1, they get relegated with probability $1/2$.

a) Calculate the stationary distribution.
b) What is the probability that 10 years from now, they will play in league 1?
c) What is the probability that 10 years from now, they will get relegated at the end of the season?
d) What is the long term ratio of years they spend in league 2?
e) Calculate the average number of years that pass between 2 consecutive appearances in league 3.

Problem 4

Solution. The states are 1, 2 and 3: state 1 corresponds to the lowest division, so division 3, state 2 is league 2 and state 3 is league 1. Then

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

Problem 4

Solution. The states are 1, 2 and 3: state 1 corresponds to the lowest division, so division 3, state 2 is league 2 and state 3 is league 1. Then

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

and the stationary vector $v_{\mathrm{st}} = (\, x_1 \; x_2 \; x_3 \,)$ can be computed from

$$\begin{aligned} \frac{1}{3}x_1 + \frac{1}{6}x_2 &= x_1 \\ \frac{2}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3 &= x_2 \\ \frac{1}{2}x_2 + \frac{1}{2}x_3 &= x_3 \\ x_1 + x_2 + x_3 &= 1 \end{aligned}$$

Problem 4

From the third equation, $x_2 = x_3$, and from the first equation, $x_2 = 4x_1$, so we have $x_1 : x_2 : x_3 = 1 : 4 : 4$ and

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \end{pmatrix}.$$

Problem 4

From the third equation, $x_2 = x_3$, and from the first equation, $x_2 = 4x_1$, so we have $x_1 : x_2 : x_3 = 1 : 4 : 4$ and

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \end{pmatrix}.$$

This is an irreducible and aperiodic Markov chain, so $v_n \approx v_{\mathrm{st}}$ for large n. 10 years is a long time, so the probability that they will play in league 1 can be approximated by $x_3 = 4/9$.

Problem 4

From the third equation, $x_2 = x_3$, and from the first equation, $x_2 = 4x_1$, so we have $x_1 : x_2 : x_3 = 1 : 4 : 4$ and

$$v_{\mathrm{st}} = \begin{pmatrix} \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \end{pmatrix}.$$

This is an irreducible and aperiodic Markov chain, so $v_n \approx v_{\mathrm{st}}$ for large n. 10 years is a long time, so the probability that they will play in league 1 can be approximated by $x_3 = 4/9$.

Let A denote the event that they get relegated at the end of the season 10 seasons from now. Its probability depends on the league they are in 10 years from now, so let B_3, B_2 and B_1 denote the events that they are in league 1, 2 or 3 respectively 10 years from now. Then according to total probability:

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \mathbb{P}(A|B_3)\mathbb{P}(B_3) \\ &\approx 0 \cdot \frac{1}{9} + \frac{1}{6} \cdot \frac{4}{9} + \frac{1}{2} \cdot \frac{4}{9} \approx 0.296.\end{aligned}$$

Problem 4

The long term average ratio of the time they spend in league 2 is $x_2 = \frac{4}{9}$.

Problem 4

The long term average ratio of the time they spend in league 2 is $x_2 = \frac{4}{9}$.

The average number of years that pass between 2 consecutive appearances in league 3 is $\frac{1}{x_1} = 9$.

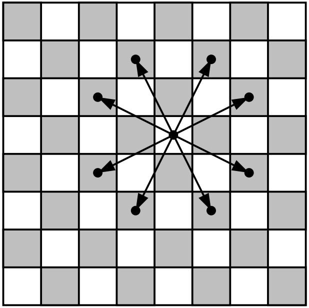
Problem 6

A knight is moving around the squares of the chessboard randomly; the next step is taken uniformly among all possible steps from the current square.

a. Argue that the position of the knight is a Markov chain.
b. Is the Markov chain irreducible or not? Is it aperiodic or periodic?
c. Calculate the stationary distribution.
d. Compute the conditional probability that the knight will be on A1 after 1000 steps, assuming it is on A1 now.
e. Compute the conditional probability that the knight will be on A2 after 1000 steps, assuming it is on A1 now.
f. Compute the conditional probability that the knight will be on A1 after 1001 steps, assuming it is on A1 now.
g. Compute the conditional probability that the knight will be on A2 after 1001 steps, assuming it is on A1 now.

Problem 6

Solution. The possible moves of the knight:



From every square, it can move to 8 other squares, except near the border of the chessboard, where there are fewer target squares.

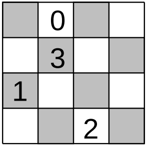
Problem 6

Since it selects randomly in each step, depending only on the current square, this is a Markov chain. It has 64 states (the squares of the chessboard).

Problem 6

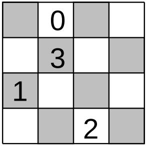
Since it selects randomly in each step, depending only on the current square, this is a Markov chain. It has 64 states (the squares of the chessboard).

The Markov chain is irreducible since the knight can get from anywhere to anywhere; this can be seen e.g. using the following 3-step combination that will move the knight to an adjacent square. It can be repeated to move the knight around the board.



Since it selects randomly in each step, depending only on the current square, this is a Markov chain. It has 64 states (the squares of the chessboard).

The Markov chain is irreducible since the knight can get from anywhere to anywhere; this can be seen e.g. using the following 3-step combination that will move the knight to an adjacent square. It can be repeated to move the knight around the board.



The Markov chain is periodic with period 2 since the knight always moves from a light square to a dark square and vice versa.

Problem 6

The transition probability matrix is a 64×64 matrix that has element $P_{i,j} = \frac{1}{d_i}$ if d_i denotes the number of possible target squares from square i. The values of d_i for each square is

2	3	4	4	4	4	3	2
3	4	6	6	6	6	4	3
4	6	8	8	8	8	6	4
4	6	8	8	8	8	6	4
4	6	8	8	8	8	6	4
4	6	8	8	8	8	6	4
3	4	6	6	6	6	4	3
2	3	4	4	4	4	3	2

Problem 6

These actually help with computing v_{st}; for the vector v whose elements are d_i for each state, we have

$$(v \cdot P)_i = \sum_j v_j P_{ji} = \sum_{j:j \text{ neighbor of } i} d_j \cdot \frac{1}{d_j} = d_i = v_i,$$

which means v satisfies $v \cdot P = v$, and thus $v_{\mathrm{st}} = \frac{1}{336} v$ since the sum of all d_i's is 336.

Problem 6

These actually help with computing v_{st}; for the vector v whose elements are d_i for each state, we have

$$(v \cdot P)_i = \sum_j v_j P_{ji} = \sum_{j: j \text{ neighbor of } i} d_j \cdot \frac{1}{d_j} = d_i = v_i,$$

which means v satisfies $v \cdot P = v$, and thus $v_{\mathrm{st}} = \frac{1}{336} v$ since the sum of all d_i's is 336.

Accordingly, the probability

$$\mathbb{P}(\text{A1 after 1000 steps}|\text{A1 initially})$$

can be computed the following way: first we check whether it is possible to be on A1 again (due to periodicity).

Problem 6

These actually help with computing v_{st}; for the vector v whose elements are d_i for each state, we have

$$(v \cdot P)_i = \sum_j v_j P_{ji} = \sum_{j:j \text{ neighbor of } i} d_j \cdot \frac{1}{d_j} = d_i = v_i,$$

which means v satisfies $v \cdot P = v$, and thus $v_{\mathrm{st}} = \frac{1}{336} v$ since the sum of all d_i's is 336.

Accordingly, the probability

$$\mathbb{P}(\text{A1 after 1000 steps}|\text{A1 initially})$$

can be computed the following way: first we check whether it is possible to be on A1 again (due to periodicity). It is possible because 1000 is even. Then the probability is approximately the period (2) multiplied by the stationary probability of the square A1, which is $\frac{4}{336}$.

Problem 6

For the probability

$$\mathbb{P}(\text{A2 after 1000 steps}|\text{A1 initially}),$$

we once again first check whether it is possible to be on A2 after 1000 steps, starting from A1.

For the probability

$$\mathbb{P}(\text{A2 after 1000 steps}|\text{A1 initially}),$$

we once again first check whether it is possible to be on A2 after 1000 steps, starting from A1. This is not possible since 1000 is even, A1 is white and A2 is black. So

$$\mathbb{P}(\text{A2 after 1000 steps}|\text{A1 initially}) = 0.$$

Problem 7

John has liability insurance for his car. The insurance company puts drivers into 4 categories: 1, 2, 3, 4. If a driver does not cause any accidents for an entire year, he moves up by 1 category (if he was in category 4, he stays there). If a driver causes a major accident, next year he goes into category 1. If a driver causes a minor accident, but no major accidents during a year, next year he moves down by 1 category (if he was in category 1, he stays there).
John causes a major accident during a year with probability $1/12$, and the probability that he causes a minor accident but no major accidents during a year is $1/4$.

Problem 7

(a) Model this process with a Markov chain. What are the states? Calculate the transition matrix. Is the Markov chain irreducible? Is it aperiodic?

(b) What is the conditional probability that John will be in category 2 two years from now, assuming that now he is in category 4?

(c) What is the probability that he will be in category 2 ten years from now?

(d) In the long run, how often does he move from category 3 to category 4 on average?

(e) For each category, the annual cost is respectively 120000, 72000, 54000, 36000 HUF. What is the long-term average annual cost paid by John?

Problem 7

Solution.

(a) States are 1, 2, 3, 4 according to the categories.

Solution.

(a) States are 1, 2, 3, 4 according to the categories.

$$P = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 1/12 & 1/4 & 0 & 2/3 \\ 1/12 & 0 & 1/4 & 2/3 \end{bmatrix}.$$

Problem 7

Solution.

(a) States are 1, 2, 3, 4 according to the categories.

$$P = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 1/12 & 1/4 & 0 & 2/3 \\ 1/12 & 0 & 1/4 & 2/3 \end{bmatrix}.$$

The Markov chain is irreducible and aperiodic (rule of thumb: if there is a positive element in the diagonal of P, then it must be aperiodic).

(b) If he is in category 4 now, then

$$v_0 = (\,0\ 0\ 0\ 1\,),$$

and

$$v_1 = v_0 P = \left(\frac{1}{12}\ 0\ \frac{1}{4}\ \frac{2}{3} \right)$$

and

$$v_2 = v_1 P = \left(\frac{5}{48}\ \frac{17}{144}\ \frac{1}{6}\ \frac{11}{18} \right),$$

so the probability that he will be in category 2 is $\frac{17}{144} \approx 0.118$.

Problem 7

(c) 10 years is a long time, so $v_{10} \approx v_{\mathrm{st}}$.

$v_{\mathrm{st}} = (x_1\ x_2\ x_3\ x_4)$ can be computed from

$$v_{\mathrm{st}} = v_{\mathrm{st}} P,$$

$$x_1 + x_2 + x_3 + x_4 = 1.$$

(c) 10 years is a long time, so $v_{10} \approx v_{\mathrm{st}}$.

$v_{\mathrm{st}} = (x_1\ x_2\ x_3\ x_4)$ can be computed from

$$v_{\mathrm{st}} = v_{\mathrm{st}} P,$$

$$x_1 + x_2 + x_3 + x_4 = 1.$$

The result is

$$v_{\mathrm{st}} = \left(\frac{1}{6}\ \frac{1}{6}\ \frac{2}{9}\ \frac{4}{9} \right),$$

so the probability that he will be in category 2 ten years from now is $\frac{1}{6}$.

Problem 7

(d) In the long run, the probability that he is in category 3 in a given year is $x_3 = \frac{2}{9}$, and

$$\mathbb{P}(\text{he moves from 3 to 4}) =$$
$$\mathbb{P}(\text{he moves from 3 to 4}|\text{he is in 3})\mathbb{P}(\text{he is in 3}) =$$
$$\frac{2}{3} \cdot \frac{2}{9} = \frac{4}{27}.$$

Problem 7

(d) In the long run, the probability that he is in category 3 in a given year is $x_3 = \frac{2}{9}$, and

$$\mathbb{P}(\text{he moves from 3 to 4}) =$$
$$\mathbb{P}(\text{he moves from 3 to 4}|\text{he is in 3})\mathbb{P}(\text{he is in 3}) =$$
$$\frac{2}{3} \cdot \frac{2}{9} = \frac{4}{27}.$$

(e) Due to the ergodic theorem, the long term average fee he pays is

$$120000 \cdot \frac{1}{6} + 72000 \cdot \frac{1}{6} + 54000 \cdot \frac{2}{9} + 36000 \cdot \frac{4}{9} = 60000$$

HUF per year.