Markov Chains I

Károly Simon

This course is based on the book:
Essentials of Stochastic processes
by R. Durrett

Department of Stochastics
Budapest University of Technology and Economics
www.math.bme.hu/~simonk

Autumn 2025, BME

1. We collect a lot of natural examples, which can be studied by the theory of Markov chains.
2. We introduce the most important notions and most important theorems.
3. Compute the stationary distributions.
4. Recurrence properties of Markov chains.
5. We study the death and birth processes as a special case of reversible Markov chains.
6. Exist distributions for absorbing Markov chains.
7. Branching processes.

Gambler's ruin

Gambler's ruin (cont.)

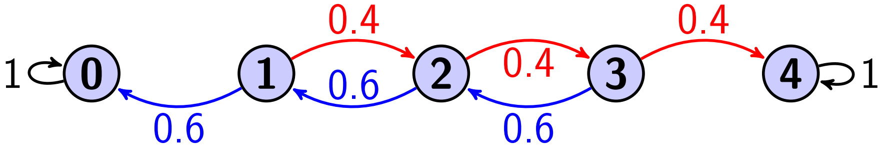
Example 1.1

We start with a gambling game, in which in every turn:

- we win \$1 with probability $p = 0.4$,
- we lose \$1 with probability $1 - p = 0.6$.

The game stops if we reach a fixed amount of $N = \$4$ or if we lose all our money.

We start at $\$X_0$, where $X_0 \in \{1, 2, 3\}$.
Let X_n be the amount of money we have after n turns.
In this case



X_n has the "Markov property". That is:
if we know X_n, any other information about the past is irrelevant for predicting the next state of X_{n+1}. Thus:

$$(1) \quad \mathbb{P}\left(X_{n+1}=j \middle| X_n=i, X_{n-1}=b_{n-1},\ldots,X_0=b_0\right) = \mathbb{P}\left(X_{n+1}=j \middle| X_n=i\right).$$

Homogeneous discrete-time Markov chain

Definition 1.2

Let S be a finite or a countably infinite (we call it countable) set. We say that X_n is a (time) homogeneous discrete-time Markov chain on state space S, with transition matrix $\mathbf{P} = p(i,j)$, if for any n, and any $i, j, b_{n-1}, \dots, b_0 \in S$:

$$\mathbb{P}\left(X_{n+1} = j \middle| X_n = i, X_{n-1} = b_{n-1}, \dots, X_0 = b_0\right) = p(i,j) \tag{2}$$

We consider only time homogeneous Markov chains and some times we abbreviate them MC.

Initial distribution

A Markov chain is determined by its initial distribution and its transition matrix. The initial distribution $\alpha = (\alpha_i)_{i \in S}$, $(\alpha_i \geq 0, \sum_{i \in S} \alpha_i = 1)$ is the distribution of the state from which a Markov chain starts. When we insist that the Markov chain starts from a given $i \in S$ (in this case $\alpha_i = 1$ and $\alpha_j = 0$ for $j \in S,\ j \neq i$) then all probabilities and expectations are denoted by

$$\mathbb{P}_i\left(\cdot\right), \mathbb{E}_i\left[\cdot\right].$$

In some cases, we write $\mathbb{P}_\alpha\left(\cdot\right), \quad \mathbb{E}_\alpha\left[\cdot\right]$ or we specify the initial distribution α in words, and then we write simply $\mathbb{P}\left(\cdot\right), \mathbb{E}\left[\cdot\right]$.

In the Gambler's ruin example, if $N = 4$ then the transition matrix $\mathbf{P}$ is a 5×5 matrix

	0	**1**	**2**	**3**	**4**
0	1	0	0	0	0
1	0.6	0	0.4	0	0
2	0	0.6	0	0.4	0
3	0	0	0.6	0	0.4
4	0	0	0	0	1

Here and many places later, the bold green numbers like $0, \dots, 4$ are the elements of the state space. So, they are NOT part of the matrix. They are the indices. The matrix above is a 5×5 matrix. For example: $p(0,0) = 1$ and $p(3,4) = 0.4$.

A simulation with Mathematica

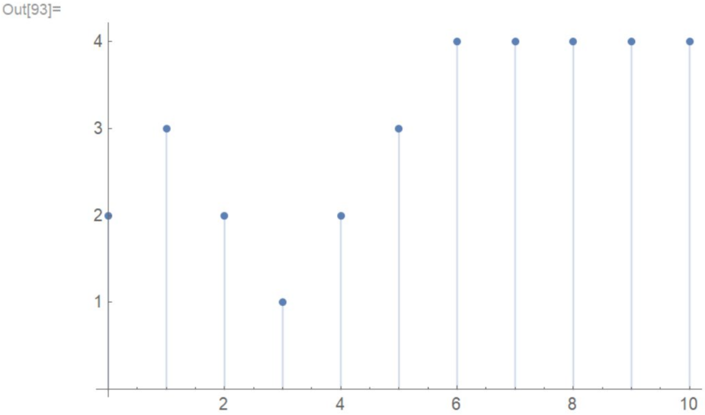


Figure: Gamblar's ruin simulation

Andrey Markov, 1856 – 1922

Ehrenfest chain

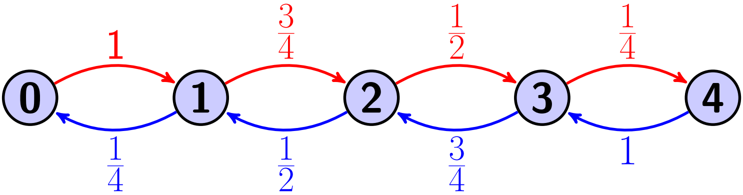
Example 1.3

We have two urns (left and right urn), in which there are a total of N balls. We pick a random ball and take it into the other urn. Let X_n be the number of balls in the left urn after the n^{th} draw. X_n has the Markov-property, because

$$p(i, i+1) = \frac{N-i}{N}, \; p(i, i-1) = \frac{i}{N} \text{ if } 0 \leq i \leq N$$

and $p(i, j) = 0$ otherwise.

$N = 4$, the corresponding graph and transition matrix:



	0	**1**	**2**	**3**	**4**
0	0	1	0	0	0
1	1/4	0	3/4	0	0
2	0	2/4	0	2/4	0
3	0	0	3/4	0	1/4
4	0	0	0	1	0

A simulation with Mathematica

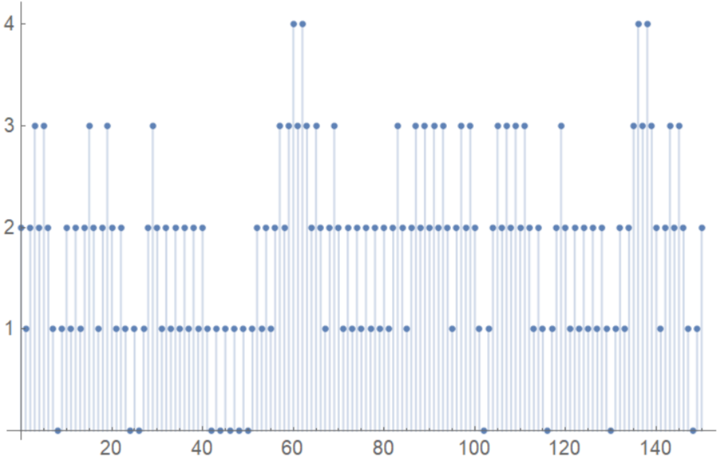


Figure: A simulation for Ehrenfest chain simulation

Another simulation with Mathematica

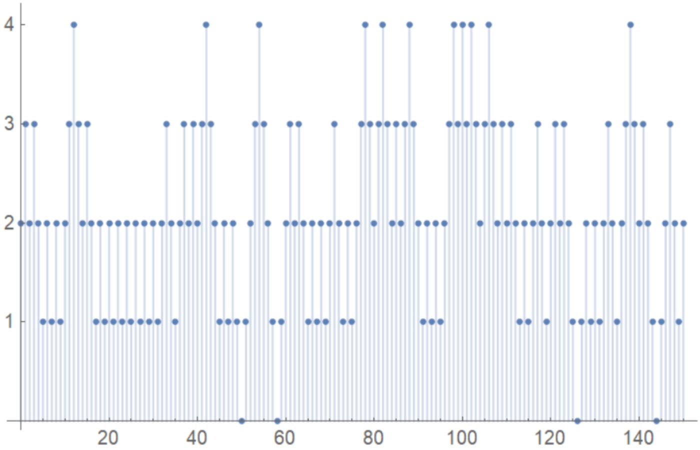


Figure: Another simulation for Ehrenfest chain simulation

The Mathematica code for the previous two simulations

$$\mathcal{P} = \texttt{DiscreteMarkovProcess}\left[\{0, 0, 1, 0, 0\}, \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & \frac{2}{4} & 0 & \frac{2}{4} & 0 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}\right]$$

```
data = RandomFunction[𝒫, {0, 150}]
```

```
TemporalData[ Time: 0 to 150
              Data points: 151      Paths: 1 ]
```

```
ListPlot[data - 1, Filling → Axis, Ticks → {Automatic, {0, 1, 2, 3, 4}}]
```

Compare the previous two chains I.

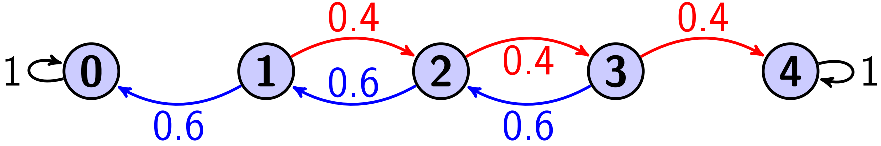


Figure: Gambler's ruin chain:

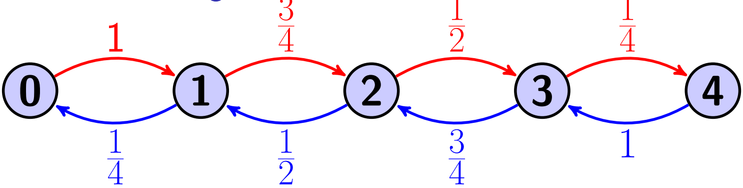


Figure: Ehrenfest chain

Compare the previous two chains II.

First, we consider the Gambler's ruin case. Let us say we start from state 2. In the gambler' ruin case with probability 0.16 we reach state 4 in two steps, and with probability 0.36 we reach state 0 and then we stay there forever. Therefore the states 0 and 4 are absorbing states. That is the probability that starting from 2 we ever return to 2 at least one more time is less than $p := 0.48 = 1 - (0.16 + 0.36)$. Then after the first return, everything starts as before independently. So, the probability that we return to 2 at least twice is less than p^2, and similarly, the probability that we return to 2 at least n times is less than p^n.

Compare the previous two chains III.

So, the probability that we return to 2 infinitely many times is $\lim_{n\to\infty} p^n = 0$. That is starting from 2, we visit 2 only finitely many times almost surely. We call those states where we return only finitely many times almost surely, transient states. Since the same reasoning applies for states $1, 3$ we can see that in the Gambler's ruin example, states $1, 2, 3$ are transient. The states where we return infinitely many times almost surely are called recurrent. Every state is either transient or recurrent.

Compare the previous two chains IV.

We spend only finite time at each transient states. So, if the state space S is finite, then we spend finite time altogether at all transient states together. This implies that
for a finite state MC we always have recurrent states.
Clearly the absorbing states $\{0, 4\}$ are always recurrent states. The following interesting questions will be answered later. To answer the first of the following two problems we need to learn about the so-called exit distributions and to answer the second one we need to study the so-called exit times .

Compare the previous two chains V.

Problem 1.4

Starting from 2 *what is the probability that the gambler eventually wins? That is she gets to* 4*?*

We answer this on slide 59 in File MC 2, see also slide 48.

Problem 1.5

Starting from 2*, what is the expected number of steps until the gambler gets to either* 0 *(ruin) or to* 4 *(success)?*

We answer this question on slide 93 in the following File.

Compare the previous two chains VI.

Now we turn to the Ehrenfest chain:

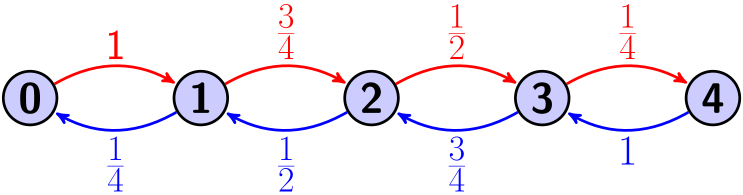


Figure: Ehrenfest chain

We consider the case again when we start from state 2. Then with 1/2-1/2 probability, we jump to either state 1 or 3.

Compare the previous two chains VII.

The probability that we do not return to 2 in any of the next $2n$ steps is $(1/4)^n$. So, the probability that we actually never return to state 2 is $\lim_{n\to\infty}\left(\frac{1}{4}\right)^n = 0$. So we return to 2 almost surely. But when we are at 2 then the whole argument repeats. So we obtain that we return to 2 infinitely many times almost surely. This means that 2 is a recurrent state. With a very similar argument, one can show that the same holds for all the other states. This means that all of the states are recurrent. In this case, we can reach from every state to every state with positive probability (after some steps). In such a situation we say that the MC is irreducible.

Compare the previous two chains VII.

Here we can ask the following question:

Problem 1.6

What is the expected number of steps so that starting from $i \in \{0, \dots, 4\}$ we get back to i for the first time?

The answer is the reciprocal of the i-th component of the so-called stationary distribution which is a probability vector $\pi = (\pi_i)_{i \in S}$, $\pi_i \geq 0$, $\sum_{i \in S} \pi_i = 1$ satisfying:

$$\pi^T \cdot P = \pi^T \tag{3}$$

Question: When do we have stationary distributions?

Theorem 1.7

Assume that card$(S) = k$, finite, and we also assume that the probability transition matrix P is irreducible. Then there is a unique stationary distribution with all components positive. That is, there exists a unique $\pi = (\pi_i)_{i \in S}$ such that $\sum_{i \in S} \pi_i = 1$ and $\pi_i > 0$ for all $i \in S$.

This follows from Theorem 7.4 (Perron-Frobenius Theorem). However, since the proof is very nice I present it here.

Proof of Theorem 1.7, Slide I

Let I be the $k \times k$ identity matrix. Since the rows of $P - I$ add to 0 , the rank of the matrix $P - I$ is less than or equal to $k - 1$, and there is a vector $v \neq 0$ such that $v^T \cdot P = v^T$. Consider the MC with probability transition matrix

$$Q = \frac{1}{2}(I + P).$$

This stays put with probability $\frac{1}{2}$ and takes a step according to P with also probability $\frac{1}{2}$ (so-called lazy chain). Let $R := Q^{k-1}$.

(4) $\quad v^T \cdot P = v^T \implies v^T \cdot Q = v^T$ and $v^T \cdot R = v^T$.

Now we prove that all elements of R are positive.

Proof of Theorem 1.7 Slide, II

The irreducibility of P implies the existence of a path, in the associated directed graph, from x to y, for every $x \neq y$, $x, y \in S$. Consider the shortest path between x and y in the MC corresponding to P. It does not go through twice on any element of S. So, the length of the shortest path is at most $k - 1$. Since we can stay put in the Markov chain corresponding to Q (the lazy chain) at any states for any time, we get that there is a path between any $x \neq y$ in R of length $k - 1$. The same clearly holds also for $x = y$. This implies that for $R = (r(x, y))_{x,y \in S}$,

$$r(x, y) > 0, \quad x, y \in S. \tag{5}$$

Proof of Theorem 1.7 Slide, II

We defined the vector v by $v^T \cdot P = v^T, v \neq 0$. Clearly all (non-zero) constant multiple of v satisfies this equation. Hence, we may assume that at least one of the components of v is positive. Now we prove that

$$v(x) > 0, \qquad \forall x \in S. \tag{6}$$

Proof of Theorem 1.7 Slide, III

Assume that there are both negative and positive components of v. Then by $v^T = v^T \cdot R$ and $r(x,y) > 0$ for all $x, y \in S$ we get
(7)

$$|v(y)| = \left|\sum_x v(x) r(x,y)\right| < \sum_x |v(x)|\, r(x,y), \quad \forall y \in S.$$

Now we use that R is a stochastic matrix, so all row sums are equal to 1. That is $\sum_{y \in S} r(x,y) = 1$, for all $x \in S$. Hence, by inequality (7):

$$\sum_y |v(y)| < \sum_x |v(x)|\,,$$

which is a contradiction. This completes the proof of (6).

Proof of Theorem 1.7 Slide, IV

Now we prove the uniqueness of the stationary distribution. This follows from the fact that

$$\text{rank}(P - I) = k - 1. \tag{8}$$

Namely, $\text{rank}(P - I) \leq k - 1$ has already been verified. If $\text{rank}(P - I) < k - 1$ then there exists a vector $w \neq 0$, $w \perp v$ with $w^T \cdot P = w^T$. We get, as above, that all the coordinates of w have the same sign let say all coordinates are positive. But then $w \perp v$ cannot happen. This completes the proof of the uniqueness.

Mathematica code for the stationary distribution

In this special case we use Mathematica we get
$\pi = \left(\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16}\right)$.

In[119]:= $p = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & \frac{2}{4} & 0 & \frac{2}{4} & 0 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$

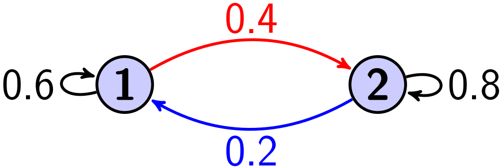
```
In[120]:= invmatrep =
  Inverse[ReplacePart[p - IdentityMatrix[Length[p[[1]]]], {i_, Length[p[[1]]]} :> 1]]

In[121]:= invmatrep[[Length[p[[1]]]]]
```

Out[121]= $\left\{\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16}\right\}$

Weather chain

Let X_n be the weather on day n on a given island, with

(9) $$X_n := \begin{cases} 1, & \text{if day } n \text{ is rainy;} \\ 2, & \text{if day } n \text{ is sunny} \end{cases}$$



	1	2
1	0.6	0.4
2	0.2	0.8

Question: What is the long-run fraction of sunny days?

π for the Weather chain

For weather chain: $\mathbf{P} = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$. We are looking for a random vector $\boldsymbol{\pi} = (\pi_1, \pi_2)$ for which:

$$(\pi_1, \pi_2) \cdot \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix} = (\pi_1, \pi_2).$$

The solution is $\boldsymbol{\pi} = (\frac{1}{3}, \frac{2}{3})$. This follows from the general result about the stationary distribution of two-states MC:

Stationary state for general two states MC

Lemma 2.1

A two-state MC's transition matrix can be written in the following way:

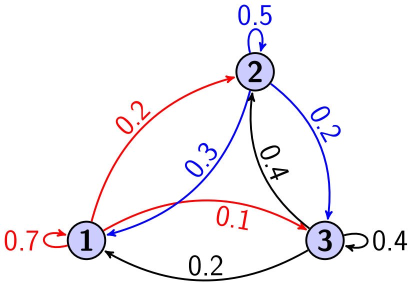
$$\mathbf{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

Then the stationary distribution is $\boldsymbol{\pi} = \left(\frac{b}{a+b}, \frac{a}{a+b}\right)$.

The proof is trivial.

Social mobility chain

Let X_n be a family's social class in the n^{th} generation, if
lower class:1 middle class:2 upper class:3



	1	2	3
1	0.7	0.2	0.1
2	0.3	0.5	0.2
3	0.2	0.4	0.4

Question: Do the fractions of people in the three classes stabilize after a long time?

For the social mobility chain

For the social mobility chain $\mathbf{P} = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}$ the equation of $\boldsymbol{\pi}^T \cdot \mathbf{P} = \boldsymbol{\pi}^T$ is

$$\begin{array}{ccccccc} 0.7\pi_1 & + & 0.3\pi_2 & + & 0.2\pi_3 & = & \pi_1 \\ 0.2\pi_1 & + & 0.5\pi_2 & + & 0.4\pi_3 & = & \pi_2 \\ 0.1\pi_1 & + & 0.2\pi_2 & + & 0.4\pi_3 & = & \pi_3 \end{array}$$

The 3^{rd} equation gives us no more information than we have already known. So, we can throw it away, and we

For the social mobility chain (cont.)

replace it with the condition that the sum of the components of $\boldsymbol{\pi}$ equals to 1. We obtain after this replacement:

$$
\begin{array}{rcrcrcr}
0.7\pi_1 & + & 0.3\pi_2 & + & 0.2\pi_3 & = & \pi_1 \\
0.2\pi_1 & + & 0.5\pi_2 & + & 0.4\pi_3 & = & \pi_2 \\
\pi_1 & + & \pi_2 & + & \pi_3 & = & 1
\end{array} \tag{10}
$$

For the social mobility chain (cont.)

After straightforward algebraic manipulations we get:

$$\begin{array}{rcrcrcl} -0.3\pi_1 & + & 0.3\pi_2 & + & 0.2\pi_3 & = & 0 \\ 0.2\pi_1 & + & -0.5\pi_2 & + & 0.4\pi_3 & = & 0 \\ \pi_1 & + & \pi_2 & + & \pi_3 & = & 1 \end{array} \tag{11}$$

$$\boldsymbol{\pi}^T \cdot \mathbf{A} = (0, 0, 1),$$

where $\boldsymbol{\pi}^T$ is a row vector and

$$\mathbf{A} := \begin{bmatrix} -0.3 & 0.2 & 1 \\ 0.3 & -0.5 & 1 \\ 0.2 & 0.4 & 1 \end{bmatrix}$$

For the social mobility chain (cont.)

So

$$\boldsymbol{\pi}^T = (0, 0, 1) \cdot A^{-1} \tag{12}$$

Steps of computing vector π:

1. Start with the transition matrix **P**,
2. subtract 1 from its diagonal elements,
3. replace the last column with the vector whose all elements are equal to 1.
4. The matrix that we obtained is called A.

For the social mobility chain (cont.)

5. By formula (12): The last row of matrix A^{-1} is $\boldsymbol{\pi}$.

In the case of the social mobility chain:

$$A^{-1} = \begin{pmatrix} -\frac{90}{47} & \frac{20}{47} & \frac{70}{47} \\ -\frac{10}{47} & -\frac{50}{47} & \frac{60}{47} \\ \frac{22}{47} & \frac{16}{47} & \frac{9}{47} \end{pmatrix}.$$

And from it: $\boldsymbol{\pi} = \left(\frac{22}{47}, \frac{16}{47}, \frac{9}{47}\right)$.

Multistep transition probabilities

Let $p^m(i,j)$ be the probability that the Markov chain with transition matrix $\mathbf{P} = p(i,j)$, starting from state i is in state j after m steps.

$$p^m(i,j) \overset{\text{in general}}{\neq} \underbrace{p(i,j) \cdots p(i,j)}_{m} \tag{13}$$

Multistep transition probabilities (cont.)

We would like to compute the m-step transition matrix with $\mathbf{P}$.
First observe that

$$p^{m+n}(i,j) = \sum_k p^m(i,k) \cdot p^n(k,j). \tag{14}$$

This is called the Chapman-Kolmogorov equation.
The proof is obvious from the following Figure:

Multistep transition probabilities (cont.)

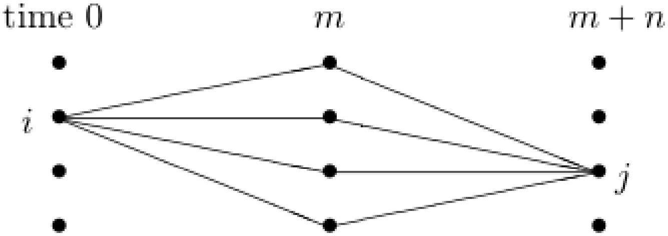


Figure: The Figure is from [2]

Multistep transition probabilities (cont.)

Theorem 3.1

The m-step transition probability $\mathbb{P}\left(X_{n+m} = j | X_n = i\right)$ *is the* (i, j)*-th element of the m-th power of the transition matrix.*

Multistep transition probabilities (cont.)

In the Gambler's ruin example, where the transition matrix was:

P	**0**	**1**	**2**	**3**	**4**
0	1	0	0	0	0
1	0.6	0	0.4	0	0
2	0	0.6	0	0.4	0
3	0	0	0.6	0	0.4
4	0	0	0	0	1

Multistep transition probabilities (cont.)

The $\lim_{n\to\infty} \mathbf{P}^n$ limit also exists, and we will see that it equals to:

$\lim_{n\to\infty} \mathbf{P}^n$	**0**	**1**	**2**	**3**	**4**
0	1	0	0	0	0
1	57/65	0	0	0	8/65
2	45/65	0	0	0	20/65
3	27/65	0	0	0	38/65
4	0	0	0	0	1

In the Ehrenfest chain example, where the transition matrix was:

	0	**1**	**2**	**3**	**4**
0	0	1	0	0	0
1	1/4	0	3/4	0	0
2	0	2/4	0	2/4	0
3	0	0	3/4	0	1/4
4	0	0	0	1	0

The $\lim_{n \to \infty} \mathbf{P}^n$ limit does NOT exist.

- A square matrix $\mathbf{P}$ is a stochastic matrix if all elements are non negative and all the row-sums are equal to 1.
- For a stochastic matrix $\mathbf{P}$ we obtain the corresponding adjacency matrix A_P by replacing all non-zero elements of P by 1. So, if

$$P = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.2 & 0.1 & 0.7 \\ 0.7 & 0.3 & 0 \end{pmatrix} \text{ then } A_P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

- We are given a Markov Chain (MC) X_n with (finite or countably infinite) state space S and transition matrix $\mathbf{P} = (p(i,j))_{i,j \in S}$ (which is always a stochastic matrix).
- We write

$$\mathbb{P}_x(A) := \mathbb{P}\left(A | X_0 = x\right).$$

$\mathbb{E}_x$ notates the expected value for the probability $\mathbb{P}_x$.
The time of the first visit to y:

$$T_y := \min\left\{n \geq 1 : X_n = y\right\}$$

So, even if we start from y, $T_y \neq 0$.

- Let $i, j \in S$, where S is the state space. We say that i and j communicate if there exists an n and an m such that $p^n(i,j) > 0$ and $p^m(j,i) > 0$.
- Observe that "communicates with" is an equivalence relation. The classes of the corresponding partition of S are called communication classes or simply classes.
- If there is only one communication class (everybody communicates with everybody) then we say that the Markov Chain (MC) is irreducible.

- Consider the MC with $S := \{1, 2, 3, 4\}$ and
$$\mathbf{P} := \begin{pmatrix} 0 & 0.4 & 0 & 0.6 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.3 & 0 & 0.7 \\ 0.1 & 0 & 0.9 & 0 \end{pmatrix}.$$ Then
$$\mathbf{P}^2 = \begin{pmatrix} 0.26 & 0. & 0.74 & 0. \\ 0. & 0.35 & 0. & 0.65 \\ 0.22 & 0. & 0.78 & 0. \\ 0. & 0.31 & 0. & 0.69 \end{pmatrix}$$ This chain is irreducible because for every $i, j \in S$ either $p(i,j) > 0$ or $p^2(i,j) > 0$ (here $p^2(i,j)$ is the (i,j)-th element of $\mathbf{P}^2$.

- The corresponding adjacency matrices for every n are:

$$A_{P^{2n-1}} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, A_{P^{2n}} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

- For the chain above the greatest common divisor (gcd):

$$\gcd\{n : p^n(i,i) > 0\} = 2 \text{ for } \forall i \in S. \qquad (15)$$

Then we say that the period of every state is 2. In general, the period of state i is

$$d_i := \gcd\{n : p^n(i,i) > 0\}.$$

We will see that in a communication class all elements have the same period. So, for an irreducible MC all elements have the same period. If this period is equal to 1 then we say that the irreducible chain is aperiodic.

- We say that a state $i \in S$ is transient if the MC returns to i finitely many times almost surely.

- We say that a state $i \in S$ is recurrent if the MC returns to i infinitely many times almost surely. Every state is either recurrent or transient.
- If an element of a communication class is recurrent then all other elements of this class are also recurrent. These classes are the recurrent classes, while the other classes are the transient classes.
- If a communication class is closed (no arrow goes out of the class) then it is recurrent class. The non-closed communication classes are the transient class.

- Let $i \in S$ be a recurrent state. We say that i is positive recurrent if the expected time of the first return to i (starting from i) is finite.
- Let $i \in S$ be a recurrent state. We say that i is null recurrent if the expected time of the first return to i (starting from i) is infinite.
- A state $i \in S$ is ergodic if i aperiodic and positive recurrent.
- A Marov chain is ergodic if all of it states are ergodic. In particular, a Markov chain is ergodic if there is an N_0 such that for every $m \geq N_0$ for every $i, j \in S$ the state j can be reached from i in m steps.

- A state $i \in S$ is absorbing if $p_{ii} = 1$ (we cannot go anywhere from this state, it is a trap).
- A Markov Chain is absorbing if every state can reach an absorbing state.
- Stationary distribution π is a probability measure on S ($\pi(i) \geq 0$ and $\sum_{i \in S} \pi(i) = 1$) which satisfies:

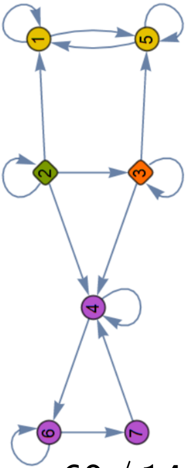
$$\pi^T \cdot \mathbf{P} = \pi^T \tag{16}$$

Convention: every vector is a column vector. When I need a row vector, I write transpose of the vector as above.

An example of irreducible classes

Example 4.1

	1	2	3	4	5	6	7
1	0.7	0	0	0	0.3	0	0
2	0.1	0.2	0.3	0.4	0	0	0
3	0	0	0.5	0.3	0.2	0	0
4	0	0	0	0.5	0	0.5	0
5	0.6	0	0	0	0.4	0	0
6	0	0	0	0	0	0.2	0.8
7	0	0	0	1	0	0	0

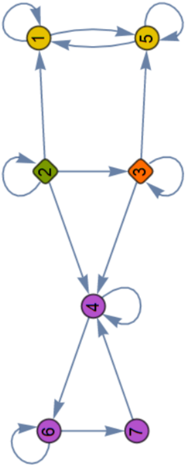


Let us create a graph whose vertices are the elements of state space $S = \{1, \dots, 7\}$ and it has directed edge (i, j) if $p(i, j) > 0$. $A \subset S$ is **closed** if it is impossible to get out. So

$$i \in A \text{ and } j \notin A \text{ then } p(i, j) = 0.$$

In the example above: sets $\{1, 5\}$ and $\{4, 6, 7\}$ are closed, so is their union, and even $\{1, 5, 4, 6, 7, 3\}$ and S itself are closed too.
$B \subset S$ is **irreducible** if any two of its elements communicate with one another: $\forall i, j \in B, \quad i \rightsquigarrow j$.

So, in the graph we can get from every element of B to any other through directed edges; and the irreducible and closed sets are: $\{1, 5\}$ and $\{4, 6, 7\}$. That is the irreducible classes are: $\{1, 5\}$ and $\{4, 6, 7\}$.



Definitions I

Here we follow Senata's book [8, Section 1.2]. For a $k \geq 1$ we use the shorthand notation

$$[k] := \{1, \dots, k\}.$$

We consider here only square matrices with non-negative elements. If we replace all positive elements of such a matrix to get its adjacency matrix. That is the adjacency matrix is a $0-1$ matrix. Let $A = (a_{i,j})_{i,j=1}^{n}$ be an $n \times n$ adjacency matrix. Then $a_{i,j} \in \{0, 1\}$.
We say that $i, i_1, \dots, i_{k-1}, j$ is a chain of length of k between i and j if

$$a_{i,i_1} \cdot a_{i_1,i_2} \cdots a_{i_{k-1}j} = 1.$$

Definitions II

We can associate a directed graph $\mathcal{G}_A = (E, V)$ with the adjacency matrix A such that

1. the set of vertices $V = [n]$ and
2. the set of edges E is defined as follows: there is directed edge between vertices i, j if and only if $a_{i,j} = 1$.

In this way $i, i_1, \ldots, i_{k-1}, j$ is a chain of length of k between i and j if and only if $i, i_1, \ldots, i_{k-1}, j$ is a chain of length of k in the directed graph $G_{\mathcal{A}}$.

Definitions III

Definition 5.1

We write

1. $i \rightarrow j$ if there is a path between i and j. Then i and j communicate. If $i \not\rightarrow j$ then i and j does not communicate.
2. $i \leftrightarrow j$ if $i \rightarrow j$ and $j \rightarrow i$.
3. i is transient if $\exists j$ such that $i \rightarrow j$ but $j \not\rightarrow i$
4. recurrent states are does which are NOT transient.
5. For a $C \subset [n]$ we say that
    1. C is irreducible if $i \leftrightarrow j$ for all $i, j \in [n]$.
    2. C is closed if $\forall i, j \in [n]$ $i \in C, j \notin C$ implies that $i \not\rightarrow j$.

If i is a recurrent state and $i \leftrightarrow j$ then j is also a recurrent state.

1. The recurrent states form classes in which everybody communicates with everybody and a member of such a clas does not communicate to anyone out of the clas. These classes are the recurrent self-communication classes.
2. Those transient states which communicate with some other states can be divided into transient classes such that any two member of such a class communicate. These are the transient communication classes.
3. There can be transient states that do not communicate with any one. They together form a class let us call it inessential class.

Path-diagram I

The path diagram for the incidens matrix $A=(a_{i,j})_{i,j=1}^{n}$:

1. Start with index 1. This is the first stage , and determine all j for which $a_{1,j}=1$. These j's form the second stage .
2. Starting from all such j repeat the previous procedure to form stage 3 and so on.
3. Stop when an index appears second time.
4. The diagram terminates when every index which appears in the diagram has been repeated.
5. If some indices were left over start with any of them and draw a similar diagram regarding the indices of the previous diagrams as "occoured in a previous stage".

Path-diagram III

Now we follow all of these on an axample:

	1	2	3	4	5	6	7	8	9
1	1	1	0	0	0	0	0	0	0
2	1	1	1	0	0	0	1	0	0
3	0	0	0	0	0	0	1	0	0
4	0	0	0	1	0	0	0	0	1
5	0	0	0	0	1	0	0	0	0
6	0	0	1	0	0	1	0	0	0
7	0	0	1	0	0	0	0	0	0
8	0	1	0	0	0	1	0	1	0
9	0	0	0	1	0	0	0	0	1

Diagram 1

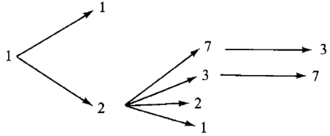
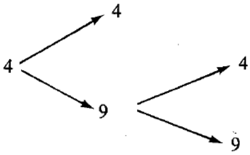


Diagram 2

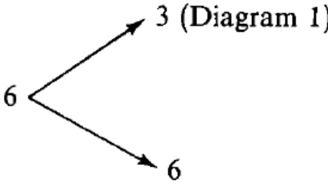


Path-diagram IV

	1	2	3	4	5	6	7	8	9
1	1	1	0	0	0	0	0	0	0
2	1	1	1	0	0	0	1	0	0
3	0	0	0	0	0	0	1	0	0
4	0	0	0	1	0	0	0	0	1
5	0	0	0	0	1	0	0	0	0
6	0	0	1	0	0	1	0	0	0
7	0	0	1	0	0	0	0	0	0
8	0	1	0	0	0	1	0	1	0
9	0	0	0	1	0	0	0	0	1

Diagram 3

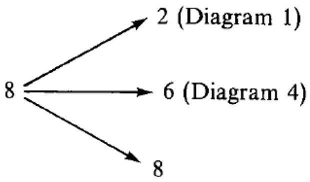
Diagram 4



Path-diagram V

	1	2	3	4	5	6	7	8	9
1	1	1	0	0	0	0	0	0	0
2	1	1	1	0	0	0	1	0	0
3	0	0	0	0	0	0	1	0	0
4	0	0	0	1	0	0	0	0	1
5	0	0	0	0	1	0	0	0	0
6	0	0	1	0	0	1	0	0	0
7	0	0	1	0	0	0	0	0	0
8	0	1	0	0	0	1	0	1	0
9	0	0	0	1	0	0	0	0	1

Diagram 5



Recurrent and tarnsient self-communication classes

1. Diagram 1 $\Longrightarrow$ $\{3,7\}$ recurrent class, $\{1,2\}$ transient class.
2. Diagram 2 $\Longrightarrow$ $\{4,9\}$ recurrent class,
3. Diagram 3 $\Longrightarrow$ $\{5\}$ recurrent class,
4. Diagram 4 $\Longrightarrow$ $\{6\}$ transient class,
5. Diagram 5 $\Longrightarrow$ $\{8\}$ transient class,

The recurrent self-communication classes:
$\{5\}, \quad \{4,9\} \quad \{3,7\}.$
The transient self-communication classes:
$\{1,2\}, \quad \{6\}, \quad \{8\}.$

Canonical form I

The so-called canonical form of the matrix on the left-hand side is the matrix on the righ-hand side.

	1	2	3	4	5	6	7	8	9
1	1	1	0	0	0	0	0	0	0
2	1	1	1	0	0	0	1	0	0
3	0	0	0	0	0	0	1	0	0
4	0	0	0	1	0	0	0	0	1
5	0	0	0	0	1	0	0	0	0
6	0	0	1	0	0	1	0	0	0
7	0	0	1	0	0	0	0	0	0
8	0	1	0	0	0	1	0	1	0
9	0	0	0	1	0	0	0	0	1

	5	4	9	3	7	1	2	6	8
5	1	0	0	0	0	0	0	0	0
4	0	1	1	0	0	0	0	0	0
9	0	1	1	0	0	0	0	0	0
3	0	0	0	0	1	0	0	0	0
7	0	0	0	1	0	0	0	0	0
1	0	0	0	0	0	1	1	0	0
2	0	0	0	1	1	1	1	0	0
6	0	0	0	1	0	0	0	1	0
8	0	0	0	0	0	0	1	1	1

.

Canonical form II

Assume that a matrix T has canonical form:

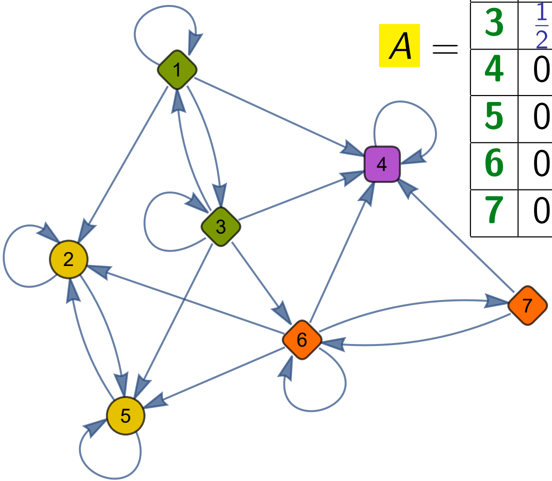
$$T = \left[\begin{array}{ccccc|c} T_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & T_2 & & & & \vdots \\ & 0 & & & & \\ & \vdots & & & & \\ 0 & 0 & \dots & T_z & & 0 \\ \hline & R & & & & Q \end{array}\right] \qquad Q = \begin{bmatrix} Q_1 & & & \\ & Q_2 & & \\ & & \ddots & 0 \\ S & & & Q_w \end{bmatrix}$$

Canonical form III

Then the k-th power T^k of T is of the form:

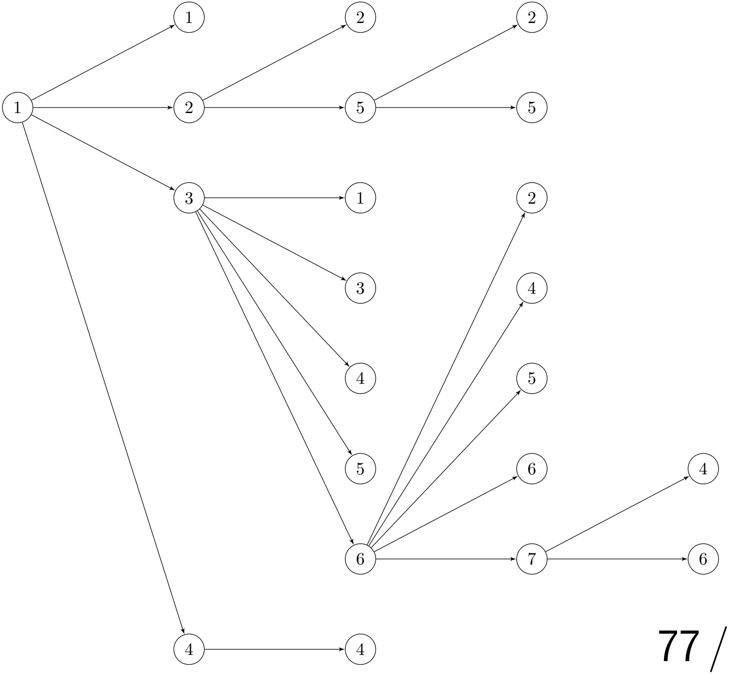
$$
T^k = \left[\begin{array}{cccc|c}
T_1^k & & & & \\
 & T_2^k & & 0 & 0 \\
0 & & \ddots & & \\
 & & & T_z^k & \\
\hline
 & R_k & & & Q^k
\end{array}\right], \quad
Q^k = \begin{bmatrix}
Q_1^k & & & \\
 & Q_2^k & & \\
 & & \ddots & 0 \\
S_k & & & Q_w^k
\end{bmatrix}
$$

Example 5.2



$A =$

	1	2	3	4	5	6	7
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0
2	0	$\frac{5}{6}$	0	0	$\frac{1}{6}$	0	0
3	$\frac{1}{2}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0
4	0	0	0	1	0	0	0
5	0	$\frac{1}{3}$	0	0	$\frac{2}{3}$	0	0
6	0	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
7	0	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0


1
1
2
3
4
2
5
1
3
4
5
6
4
2
5
2
4
5
6
7
4
6

Recurrent Classes: $\{2,5\}$ and $\{4\}$

Transient Classes: $\{1,3\}$ and $\{6,7\}$

$T =$

	4	**2**	**5**	**6**	**7**	**3**	**1**
4	1	0	0	0	0	0	0
2	0	$\frac{5}{6}$	$\frac{1}{6}$	0	0	0	0
5	0	$\frac{1}{3}$	$\frac{2}{3}$	0	0	0	0
6	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	0	0
7	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	0	0
3	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$

$$= \left(\begin{array}{c||c} U & \mathbf{0}_{3,4} \\ \hline\hline V & W \end{array}\right)$$

That is

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \; V = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{8} & 0 & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix}, \; W = \begin{pmatrix} \frac{1}{2} & \frac{1}{8} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{8} & 0 & \frac{1}{8} & \frac{1}{2} \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} \end{pmatrix},$$

and $\mathbf{0}_{3,4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Let $I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, and

$$\mathbf{0}_{4,4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\pi := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 2 & 6 & 1 & 3 & 4 & 5 \end{pmatrix}, \pi^{-1} := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 2 & 5 & 6 & 7 & 3 & 1 \end{pmatrix}$$

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \Pi^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In the matrix Π in the i-th row the 1 is at the position $\pi(i)$. With this notation:

$$T(i,j) = A(\pi^{-1}(i), \pi^{-1}(j)), \quad A(i,j) = T(\pi(i), \pi(j)). \tag{17}$$

By the definition of matrix products we get

$$T = \Pi^{-1} \cdot A \cdot \Pi. \tag{18}$$

We know that

$$T^n = \left(\begin{array}{c||c} U^n & \mathbf{0}_{3,4} \\ \hline\hline S_n & W^n \end{array}\right). \tag{19}$$

Moreover,

1. Using that the matrix W corresponds to the transient states we get that $\lim_{n\to\infty} W^n = \mathbf{0}$.
2. We learned that $\lim_{n\to\infty} U^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} =: B$.

So $T_\infty := \lim_{n\to\infty} T^n = \left(\begin{array}{c||c} B & \mathbf{0}_{3,4} \\ \hline\hline X & \mathbf{0}_{4,4} \end{array}\right)$, where $X = \lim_{n\to\infty} S_n$.

Using that

$$\left(\begin{array}{c||c} U & \mathbf{0} \\ \hline\hline V & W \end{array}\right) \cdot \left(\begin{array}{c||c} B & \mathbf{0}_{3,4} \\ \hline\hline X & \mathbf{0}_{4,4} \end{array}\right) = T \cdot T_\infty = T_\infty = \left(\begin{array}{c||c} B & \mathbf{0}_{3,4} \\ \hline\hline X & \mathbf{0}_{4,4} \end{array}\right)$$

We get that

$$V \cdot B + W \cdot X = X = I \cdot X. \tag{20}$$

On slide 79 we defined the matrices W, I, V, B we can compute:

$$X = (W - I)^{-1} \cdot (-V \cdot B) = \begin{pmatrix} \frac{3}{7} & \frac{8}{21} & \frac{4}{21} \\ \frac{5}{7} & \frac{4}{21} & \frac{2}{21} \\ \frac{58}{119} & \frac{122}{357} & \frac{61}{357} \\ \frac{59}{119} & \frac{40}{119} & \frac{20}{119} \end{pmatrix}. \tag{21}$$

Hence,

$$T_\infty = \left(\begin{array}{c||c} B & \mathbf{0}_{3,4} \\ \hline\hline X & \mathbf{0}_{4,4} \end{array}\right) = \left(\begin{array}{ccc||cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \hline\hline \frac{3}{7} & \frac{8}{21} & \frac{4}{21} & 0 & 0 & 0 & 0 \\ \frac{5}{7} & \frac{4}{21} & \frac{2}{21} & 0 & 0 & 0 & 0 \\ \frac{58}{119} & \frac{122}{357} & \frac{61}{357} & 0 & 0 & 0 & 0 \\ \frac{59}{119} & \frac{40}{119} & \frac{20}{119} & 0 & 0 & 0 & 0 \end{array}\right).$$

Finally we get for $A_\infty := \lim_{n\to\infty} A^n$ that

$$A_\infty = \Pi \cdot T_\infty \cdot \Pi^{-1} = \begin{pmatrix} 0 & \frac{40}{119} & 0 & \frac{59}{119} & \frac{20}{119} & 0 & 0 \\ 0 & \frac{2}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{122}{357} & 0 & \frac{58}{119} & \frac{61}{357} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{8}{21} & 0 & \frac{3}{7} & \frac{4}{21} & 0 & 0 \\ 0 & \frac{4}{21} & 0 & \frac{5}{7} & \frac{2}{21} & 0 & 0 \end{pmatrix},$$

where the permutation matrices Π and Π^{-1} were defined on slide 80. This implies for example that starting from 5 after very many steps, the probability that we are at 2 is apprIximately $\frac{2}{3}$ and that we are at 5 is approximately $\frac{1}{3}$.

On the following slides we state the limit theorems.

One of the important consequence of the following theorems is that under some not restrictive conditions, the same thing happens as on slide 48. That is $\lim_{n\to\infty} P^n$ exists and equal to a matrix whose all rows are equal to π.

Limit Theorems (Preparation)

- Given a Markov Chain (X_n) on a
- state space S (finite or countably infinite)
- transition matrix $\mathbf{P} = (p(i,j))_{i,j \in S}$.
- $p^m(i,j)$: the probability that starting from i we will be in j after m steps.

Definition 6.1 (Abbreviations used below)

- $\mathcal{I}$: irreducible,
- $\mathcal{A}$: aperiodic,
- $\mathcal{R}$: all states are recurrent,
- $\mathcal{S}$: $\exists \pi$ stationary distribution.

The Limit theorems below hold for countable state spaces. This means that the state space S is either countably infinite of finite.

Theorem 6.2 (Convergence Theorem)

$\mathcal{I}$ and $\mathcal{A}$ and $\mathcal{S}$ implies that

(a) *The MC is positive recurrent,*

(b) $\lim_{n\to\infty} p^n(i,j) = \boldsymbol{\pi}(j), \forall i,j$

(c) $\forall j, \boldsymbol{\pi}(j) > 0.$

(d) *The stationary distribution is unique.*

Theorem 6.3 (Asymptotic frequency)

$\mathcal{I}$ *and* $\mathcal{R}$ $\Longrightarrow \lim_{n\to\infty} \frac{\#\{k\leq n: X_k = j\}}{n} = \frac{1}{\mathbb{E}_j[T_j]}, \ \forall j \in S,$
where $\mathbb{E}_j\left[T_j\right]$ *is the expected time of the first return to* j*, starting from* j*.*

Theorem 6.4 (π is unique)

$\mathcal{I}$ *and* $\mathcal{S}$ $\Longrightarrow \pi(j) = \frac{1}{\mathbb{E}_j[T_j]}, \ \forall j \in S.$
In particular, π *is unique.*

Theorem 6.5

Let $f : S \to \mathbb{R}$, *s.t.* $\sum_{i \in S} |f(i)| \cdot \pi(i) < \infty$. *Then*

$$(22) \quad \mathcal{I} \text{ and } \mathcal{S} \Longrightarrow \lim_{n\to\infty} \frac{1}{n} \cdot \sum_{m=1}^{n} f(X_m) = \sum_{i \in S} f(i) \cdot \pi(i).$$

The Limit theorems below hold for finite state spaces.

Theorem 6.6 (Finite state space I)

$\#S < \infty$ and $\mathcal{I}$ and $\mathcal{A}$ then

(a) *π exists and unique,*

(b) *$\pi_i > 0$ for all $i \in S$.*

(c) *For every initial distribution $\boldsymbol{\alpha}$ on S we have $\lim_{n\to\infty} \boldsymbol{\alpha}^T \cdot P^n = \pi^T$*

The proof is [7, p. 19]. If $\#S < \infty$ then the assumptions of the theorem are equivalent to P is primitive: ($\exists k$ s.t. $P^k > 0$ that is all elements of P^k are positive.)

Theorem 6.7 (Finite state space II)

$\#S < \infty$ and $\mathcal{I}$ then

- *(a) π exists and unique,*
- *(b) $\pi_i > 0$ for all $i \in S$.*
- *(c) But it is not necessarily true that for every initial distribution $\boldsymbol{\alpha}$ on S we have $\lim_{n \to \infty} \boldsymbol{\alpha}^T \cdot P^n = \pi^T$*

We give a proof only in the very special case when $\text{card}(S) = 2$ and all elements of the transition probability matrix is positive. That is

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

where $0 < p, q < 1$. This matrix has eigenvalues 1 and $1-p-q$. We diagonalize $P = Q \cdot D \cdot Q^{-1}$, where:

$$Q = \begin{bmatrix} 1 & -p \\ 1 & q \end{bmatrix}, \quad Q^{-1} = \begin{bmatrix} q/(p+q) & p/(p+q) \\ -1/(p+q) & 1/(p+q) \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1-p-q \end{bmatrix}.$$

The columns of Q are right eigenvectors of P and the rows of Q^{-1} are left eigenvectors. The eigenvectors are unique up to a multiplicative constant.

The constant is chosen such that the left eigenvector for eigenvalue 1 is a probability vector:

$$\pi = (q/(p+q), p/(p+q))$$

is the unique invariant probability distribution for P. In the following computation the key fact is the observation that

$$1 > |1 - p - q| \, . \tag{23}$$

$$P^n = \left(QDQ^{-1}\right)^n = QD^nQ^{-1} = Q\begin{bmatrix} 1 & 0 \\ 0 & (1-p-q)^n \end{bmatrix}Q^{-1}$$
$$= \begin{bmatrix} [q + p(1-p-q)^n]/(p+q) & [p - p(1-p-q)^n]/(p+q) \\ [q - q(1-p-q)^n]/(p+q) & [p + q(1-p-q)^n]/(p+q) \end{bmatrix}$$

Using that $|1-p-q| < 1$, we get

$$\lim_{n\to\infty} \mathbf{P}^n = \begin{bmatrix} q/(p+q) & p/(p+q) \\ q/(p+q) & p/(p+q) \end{bmatrix} = \begin{bmatrix} \pi \\ \pi \end{bmatrix}.$$

Notation

Let $A = (a_{ij})$ be matrix of $N \times N$. We are assuming from now on that A is nonnegative.
Hence $a_{ij} \geq 0$.

$a_{ij}^{(m)}$ denotes element (i, j) of matrix A^m.

Notation (cont.)

Definition 7.1 (Adjacency matrix of directed graphs)

Let $G = (V, E)$ be a directed graph. We denote the set of vertices by V and the set of edges by E.
The adjacency matrix of graph G (the matrix of its vertices): $A_G = (a_{ij})$

$$a_{ij} = \begin{cases} 1, & \text{if } (i,j) \in E; \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

Notation (cont.)

It is easy to see that

$$a_{ij}^{(m)} = \# \{\text{paths with length } m \text{ from } i \text{ to } j\} . \tag{25}$$

On the other hand, for every nonnegative $N \times N$ matrix A there exists a directed graph G_A in which $V(G) := \{1, \dots, N\}$ and

$$(i, j) \in E(G) \text{ if and only if } a_{ij} > 0.$$

Notation (cont.)

Definition 7.2 (irreducible matrices)

Matrix A is irreducible, if $\forall (i, j),\ \exists m = m(i, j)$, so that $a_{ij}^{(m)} > 0$

It's obvious that A is irreducible if and only if G_A is strongly connected, so there is a path in each direction between each pair of vertices of the graph.

Notation (cont.)

Definition 7.3 (Primitive matrices)

We say that a nonnegative matrix A is primitive, if $\exists M : \forall i, j,\ a_{ij}^{(M)} > 0$

- If a matrix is irreducible and aperiodic then this matrix is primitive (see [7, p. 19]).
- It is easy to see that if a nonnegative matrix is irreducible and at least one of its diagonal elements is nonzero, then it is primitive.

Notation (cont.)

The proof of the following Perron-Frobenius Theorem is available for example in the Appendix of the book Karlin-Taylor [5].

Perron-Frobenius Theorem I

Theorem 7.4

Let A be a $N \times N$ nonnegative matrix. Then

(i) *A has eigenvalue $\lambda \in \mathbb{R}_0^+$ (so called as Perron-Frobenius eigenvalue) such that no other eigenvalues of A are greater than λ in absolute value.*

(ii) $\min_i \sum_{j=1}^{N} a_{ij} \le \lambda \le \max_i \sum_{j=1}^{N} a_{ij}$.

(iii) *We can choose the left and right eigenvectors $\mathbf{u}$ and $\mathbf{v}$ of λ so that all of their components are nonnegative.*

$$\mathbf{u}^T \cdot A = \lambda \mathbf{u}^T, \; A \cdot \mathbf{v} = \lambda \cdot \mathbf{v}.$$

Perron-Frobenius Theorem II

From now on we normalize **u** and **v** so that

$$\sum_{i=1}^{N} u_i = 1 \quad \text{and} \quad \sum_{i=1}^{N} u_i v_i = 1 . \tag{26}$$

If we additionally assume that A is irreducible, then:

(iv) λ is eigenvalue with multiplicity 1 and all elements of **u** and **v** are strictly positive.

(v) λ is the only eigenvalue for which there exists an eigenvector with only nonnegative elements.

Perron-Frobenius Theorem III

And if we assume that A is primitive, then:

(vi) $\forall i, j$:

$$\lim_{n\to\infty} \lambda^{-n} a_{ij}^{(n)} = u_j v_i, \tag{27}$$

where $\mathbf{u}, \mathbf{v}$ are the left and right eigenvectors with positive components corresponding to λ which satisfy condition (26).

Application for Markov chains

In our case the matrix A is te transition matrix P which is a stochastic matrix. Then all row sums are equal to 1.This implies that

- $\lambda = 1$ according to (ii) on slide 103 and
- $\mathbf{v} = (1, \ldots, 1)$.
- $\mathbf{u}^T \cdot P = \mathbf{u}^T$ by (iii) on slide 103 and by (26). That is the stationary distribution $\pi = \mathbf{u}$.

Application for Markov chains (cont.)

Then (vi) on slide 105 reads like: $\forall i, j \in S$

$$\lim_{n\to\infty} p_{i,j}^n = u_j = \pi_j, \tag{28}$$

here $p_{i,j}^n$ was defined on slide 43. So, Theorem 6.6 is a corollary of the Peron-Frobenius Therem.
Moreover, let Π be an $|S| \times |S|$ matrix, (where $|S|$ is the cardinality of S) such that all **rows** of Π are equal to π. Then

$$\lim_{n\to\infty} P^n = \Pi. \tag{29}$$

Observe that (28) is the same as (29) in terms of components. Speed of convergence:see [6, Theorem 4.9].

$\#S < \infty$, irreducible with period d

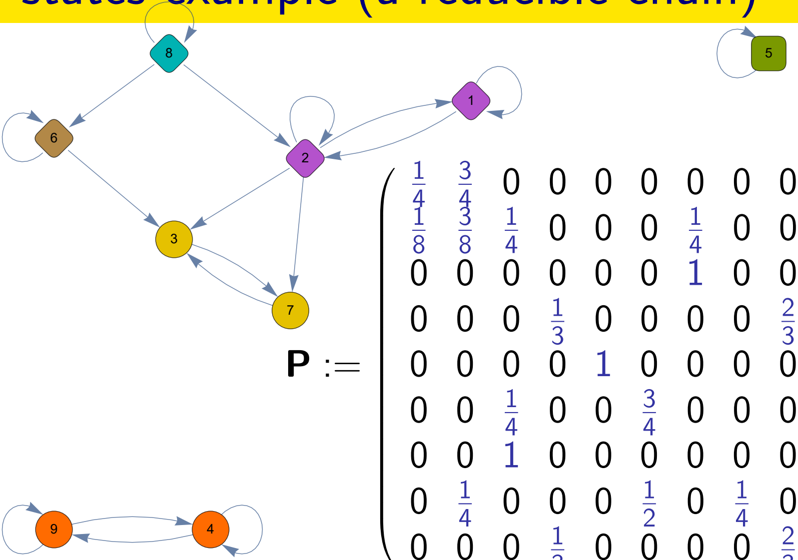
Theorem 7.5

Assume that $\#S < \infty$, P is irreducible, periodic with period $d > 1$. Then P has d eigenvalues with absolute value 1, each of them is simple. In particular 1 is a simple eigenvalue that is there is a unique invariant probability vector π corresponding to the eigenvalue 1. Let $\boldsymbol{\alpha}$ be a probability distribution on S. That is $\boldsymbol{\alpha} = (\alpha_i)_{i \in S}$ with $\sum_{i \in S} \alpha_i = 1$ and $\alpha_i \geq 0$. Then

$$\lim_{n \to \infty} \frac{1}{d} \left(\boldsymbol{\alpha}^T \cdot P^{n+1} + \cdots + \boldsymbol{\alpha}^T \cdot P^{n+d} \right) = \pi. \tag{30}$$

This Theorem is a corollary of Theorem 6.2.

A 9-states example (a reducible chain)



$$\mathbf{P} := \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{2}{3} \end{pmatrix}$$

Continuation

This shows that $\{1, 2\}$, $\{6\}$ and $\{8\}$ are transient classes. This implies that their measure by the stationary distribution must be zero. On each of the recurrent classes we have different stationary distributions which have nothing to do with each other. On the class $\{3, 7\}$, $\{4, 9\}$ and $\{5\}$ the stationary distributions in this order are: $\widehat{\pi} := \left(0, 0, \frac{1}{2}, 0, 0, 0, \frac{1}{2}, 0, 0\right)$.

$\widetilde{\pi} := \left(0, 0, 0, \frac{1}{3}, 0, 0, 0, 0, \frac{2}{3}\right)$,

$\overline{\pi} := (0, 0, 0, 0, 1, 0, 0, 0, 0)$.

Let $\pi := \alpha_1\widehat{\pi} + \alpha_2\widetilde{\pi} + \alpha_3\overline{\pi}$, where $\alpha_i \geq 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Then π is one of the uncountably many stationary distributions of the chain.

Continuation

We obtained $\tilde{\pi}$ on the previous slide by the Mathematica

```
In[78]:= PDF[StationaryDistribution[
  DiscreteMarkovProcess[4, {{1/4, 3/4, 0, 0, 0, 0, 0, 0, 0}, {1/8, 3/8, 1/4, 0, 0, 0, 1/4, 0, 0},
   {0, 0, 0, 0, 0, 0, 1, 0, 0}, {0, 0, 0, 1/3, 0, 0, 0, 0, 2/3}, {0, 0, 0, 0, 1, 0, 0, 0, 0},
   {0, 0, 1/4, 0, 0, 3/4, 0, 0, 0}, {0, 0, 1, 0, 0, 0, 0, 0, 0}, {0, 1/4, 0, 0, 0, 1/2, 0, 1/4, 0},
   {0, 0, 0, 1/3, 0, 0, 0, 0, 2/3}}]], 9]

Out[78]= 2/3
```

Explanation: The very first number in the code is 4. It says that we are in the recurrence class that contains 4. The very last number is 9. This gives the measure of state 9 for that stationary distribution which is supported by the recurrence class that contains 4.

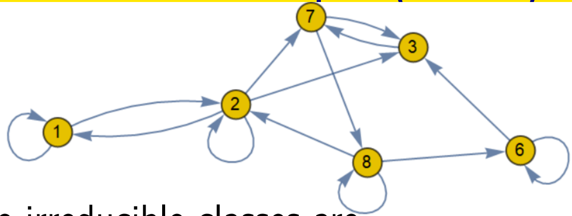
Another 9 states example

Example 7.6

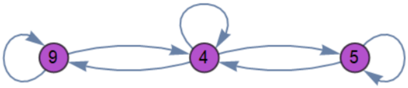
Find the all of the stationary distributions for the Markov chain given by P, where P is:

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

Another 9 states example (cont.)



That is the irreducible classes are $\{1, 2, 3, 6, 7, 8\}$ (above) and, $\{4, 5, 9\}$ (below). Then we run the Mathematica code on the next slide. The only thing missing from this code is the definition of the matrix p which should be defined first as P.



Another 9 states example (cont.)

```
In[43]:= pbig = p[[{1, 2, 3, 6, 7, 8}, {1, 2, 3, 6, 7, 8}]]

In[44]:= psmall = p[[{4, 5, 9}, {4, 5, 9}]]

In[47]:= invmatrep =
          Inverse[ReplacePart[pbig - IdentityMatrix[Length[pbig[[1]]]],
            {i_, Length[pbig[[1]]]} :> 1]]
         invmatrep[[Length[pbig[[1]]]]]
```

Out[48]=

$$\left\{\frac{2}{31}, \frac{4}{31}, \frac{7}{31}, \frac{4}{31}, \frac{8}{31}, \frac{6}{31}\right\}$$

```
In[49]:= invmatrep =
          Inverse[ReplacePart[psmall - IdentityMatrix[Length[psmall[[1]]]],
            {i_, Length[psmall[[1]]]} :> 1]]
         invmatrep[[Length[psmall[[1]]]]]
```

Out[50]=

$$\left\{\frac{3}{7}, \frac{2}{7}, \frac{2}{7}\right\}$$

Another 9 states example (cont.)

That is let

$$\pi^{(1)} := \left(\frac{2}{31}, \frac{4}{31}, \frac{7}{31}, 0, 0, \frac{4}{31}, \frac{8}{31}, \frac{6}{31}, 0\right),$$

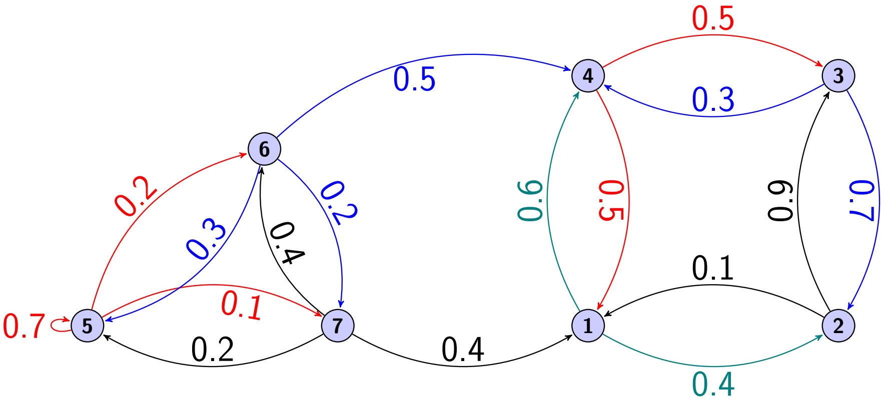
$$\pi^{(2)} := \left(0, 0, 0, \frac{3}{7}, \frac{2}{7}, 0, 0, 0, \frac{2}{7}, \right).$$

Then for every $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ with $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$ the vector

$$\pi = \alpha_1 \cdot \pi^{(1)} + \alpha_2 \cdot \pi^{(2)} \tag{31}$$

is a stationary distribution and all stationary distributions π can be presented of the form as in (31) for suitable α_1, α_2.

Example 7.7 (Triangle-square chain)



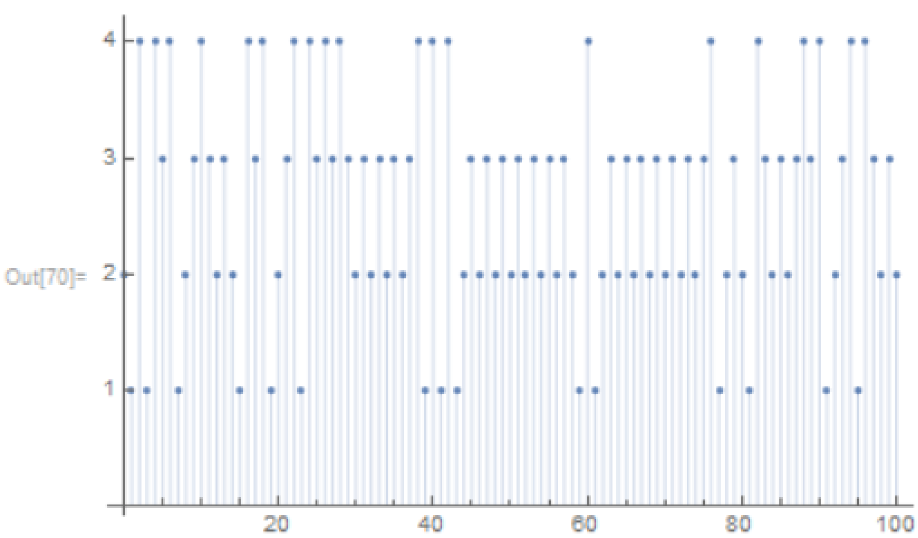
Transient states: 1, 2, 3, Recurrent states: 4, 5, 6, 7.

Example 7.7 (cont.)

It is enough to focus on the right hand side square shaped part. That is the subgraph of vertices $\{1, 2, 3, 4\}$. The transition matrix is:

$$\mathbf{P} = \begin{pmatrix} 0 & 0.4 & 0 & 0.6 \\ 0.1 & 0 & 0.9 & 0 \\ 0 & 0.7 & 0 & 0.3 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix}$$

Example 7.7 (cont.)



Stationary distribution with Mathematica

In[74]:=
```
Clear[p]
```

In[75]:=

$$p = \begin{pmatrix} 0 & \frac{4}{10} & 0 & \frac{6}{10} \\ \frac{1}{10} & 0 & \frac{9}{10} & 0 \\ 0 & \frac{7}{10} & 0 & \frac{3}{10} \\ \frac{5}{10} & 0 & \frac{5}{10} & 0 \end{pmatrix}$$

Out[75]= $\left\{\left\{0, \frac{2}{5}, 0, \frac{3}{5}\right\}, \left\{\frac{1}{10}, 0, \frac{9}{10}, 0\right\}, \left\{0, \frac{7}{10}, 0, \frac{3}{10}\right\}, \left\{\frac{1}{2}, 0, \frac{1}{2}, 0\right\}\right\}$

In[76]:=
```
invmatrep =
 Inverse[ReplacePart[p - IdentityMatrix[Length[p[[1]]]], {i_, Length[p[[1]]]} :> 1]]
```

Out[76]= $\left\{\left\{-\frac{61}{88}, -\frac{25}{176}, \frac{17}{88}, \frac{113}{176}\right\}, \left\{\frac{5}{11}, -\frac{25}{22}, -\frac{5}{11}, \frac{25}{22}\right\}, \left\{\frac{39}{88}, -\frac{85}{176}, -\frac{83}{88}, \frac{173}{176}\right\}, \left\{\frac{1}{8}, \frac{5}{16}, \frac{3}{8}, \frac{3}{16}\right\}\right\}$

In[77]:=
```
invmatrep[[Length[p[[1]]]]]
```

Out[77]= $\left\{\frac{1}{8}, \frac{5}{16}, \frac{3}{8}, \frac{3}{16}\right\}$

$P = S \cdot D \cdot S^{-1}$

$$\mathrm{S} = \begin{pmatrix} -1 & 1 & -\frac{3\sqrt{3}}{5} & \frac{3\sqrt{3}}{5} \\ 1 & 1 & -\frac{3}{5} & -\frac{3}{5} \\ -1 & 1 & \frac{\sqrt{3}}{5} & -\frac{\sqrt{3}}{5} \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad \mathrm{D} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{3}}{5} & 0 \\ 0 & 0 & 0 & \frac{\sqrt{3}}{5} \end{pmatrix}$$

Let $D_{odd} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; D_{even} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

$$\lim_{n\to\infty} D^{2n+1} = D_{odd}, \; \lim_{n\to\infty} D^{2n} = D_{even}$$

$\lim_{n\to\infty} P^{2n+1} = S \cdot D_{odd} \cdot S^{-1}$, and
$\lim_{n\to\infty} P^{2n} = S \cdot D_{even} \cdot S^{-1}$.

Cont.

This yields that

$$\lim_{n\to\infty} P^{2n+1} = \begin{pmatrix} 0 & \frac{5}{8} & 0 & \frac{3}{8} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{5}{8} & 0 & \frac{3}{8} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \end{pmatrix} \quad \lim_{n\to\infty} P^{2n} = \begin{pmatrix} \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{5}{8} & 0 & \frac{3}{8} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{5}{8} & 0 & \frac{3}{8} \end{pmatrix}$$

and $\lim_{n\to\infty} \frac{1}{2}\left(P^{2n+1} + P^{2n+1}\right) = \begin{pmatrix} \frac{1}{8} & \frac{5}{16} & \frac{3}{8} & \frac{3}{16} \\ \frac{1}{8} & \frac{5}{16} & \frac{3}{8} & \frac{3}{16} \\ \frac{1}{8} & \frac{5}{16} & \frac{3}{8} & \frac{3}{16} \\ \frac{1}{8} & \frac{5}{16} & \frac{3}{8} & \frac{3}{16} \end{pmatrix}$. Note that all row vectors are the same and identical to π.

Inventory chain Durrett, Example 1.6

s, S **storage strategy**:

- Given $s < S$
- Let X_n be the amount of stock on hand at the end of day n.

Strategy:

- If $X_n \leq s$ we fill up the stock during the night so that the stock at the beginning of day $n+1$ is S.
- If $X_n > s$ we do not do anything.

Inventory chain Durrett, Example 1.6 (cont.)

Let D_{n+1} be the demand of this item on day $n+1$.

Using the $x^+ := \max\{x, 0\}$ notation:

$$X_{n+1} = \begin{cases} (X_n - D_{n+1})^+, & \text{if } X_n > s; \\ (S - D_{n+1})^+, & \text{if } X_n \leq s. \end{cases}$$

Inventory chain Durrett, Example 1.6 (cont.)

In an example with $s = 1, S = 5$ and
$\mathbb{P}(D_{n+1} = 0) = 0.3, \ \mathbb{P}(D_{n+1} = 1) = 0.4$
$\mathbb{P}(D_{n+1} = 2) = 0.2, \ \mathbb{P}(D_{n+1} = 3) = 0.1$

	0	**1**	**2**	**3**	**4**	**5**
0	0	0	0.1	0.2	0.4	0.3
1	0	0	0.1	0.2	0.4	0.3
2	0.3	0.4	0.3	0	0	0
3	0.1	0.2	0.4	0.3	0	0
4	0	0.1	0.2	0.4	0.3	0
5	0	0	0.1	0.2	0.4	0.3

Inventory chain Durrett, Example 1.6 (cont.)

For $s = 1$ and $S = 5$ the stationary distribution is:

$$\pi = \left\{ \frac{177}{1948}, \frac{379}{2435}, \frac{225}{974}, \frac{105}{487}, \frac{98}{487}, \frac{1029}{9740} \right\}$$

Assume that the profit of every single item is \$12, but the daily storage fee is \$2.

Question:

- What is the long-term profit on this item for the previous choice of s, S?

Inventory chain Durrett, Example 1.6 (cont.)

- How should we choose values of s, S to maximize the profit?

Repair chain

A machine has 3 critical components which can go wrong, but the machine operates until all of them stops working. If at least two components are broken, they get repaired for the next day. We assume that on a single day maximum 1 component can go wrong, and the probability of component 1, 2 and 3 failing is (in order) 0.01, 0.02 and 0.04.

If we are to model this process with a Markov chain, it is recommended to use state space of broken parts: $\{0, 1, 2, 3, 12, 13, 23\}$. The transition matrix is:

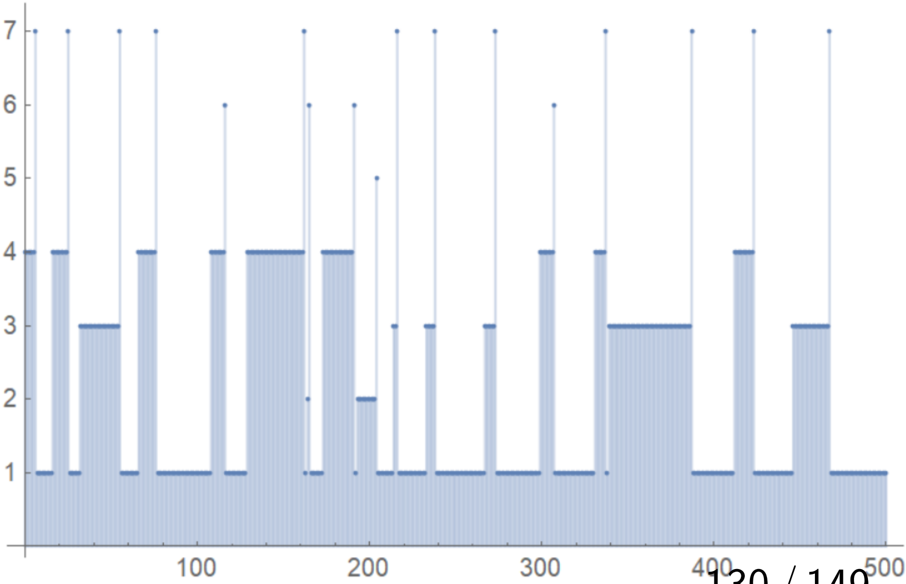
Repair chain (cont.)

	0	**1**	**2**	**3**	**12**	**13**	**23**
0	0.93	0.01	0.02	0.04	0	0	0
1	0	0.94	0	0	0.02	0.04	0
2	0	0	0.95	0	0.01	0	0.04
3	0	0	0	0.97	0	0.01	0.02
12	1	0	0	0	0	0	0
13	1	0	0	0	0	0	0
13	1	0	0	0	0	0	0

Repair chain (cont.)

Question: How many components are used of type $1, 2$ and 3 in 1000 days?

Repair chain (cont.)



Repair c

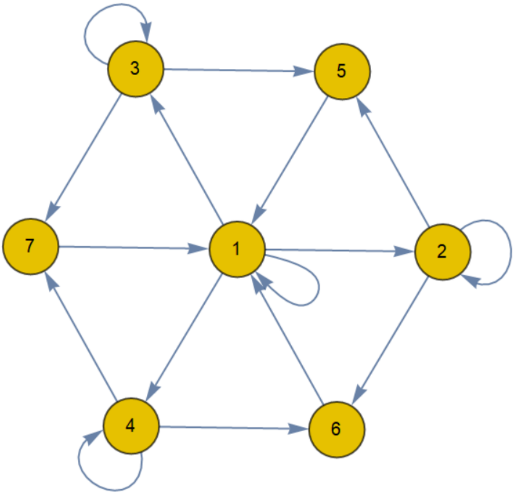


Figure: Prepared with Wolfram mathematica

Repair chain (cont.)

Stationary distribution:

$$\pi = (0.336, 0.056, 0.134, 0.448, 0.002, 0.006, 0.014)$$

Wright-Fisher model

Example 7.8

A (fixed size) generation consists of $2N$ genes with type either a or A. If there are $j \in \{0, \ldots, 2N\}$ a-type gene in the parent population, then the next generaton's building will be determined with $2N$ independent binomial trials, with probabilities $p_j = \frac{j}{2N}$, $q_j = 1 - \frac{j}{2N}$. So, if X_n is the number of a-type genes in the n^{th} generation, then the appropriate Markov-chain is:

$\mathbb{P}(X_{n+1} = k | X_n = j) = p(j, k) = \binom{2N}{k} p_j^k q_j^{2N-k}$.

The transition matrix for the Wright-Fisher model when $2N = 6$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{15625}{46656} & \frac{3125}{7776} & \frac{3125}{15552} & \frac{625}{11664} & \frac{125}{15552} & \frac{5}{7776} & \frac{1}{46656} \\ \frac{64}{729} & \frac{64}{243} & \frac{80}{243} & \frac{160}{729} & \frac{20}{243} & \frac{4}{243} & \frac{1}{729} \\ \frac{1}{64} & \frac{3}{32} & \frac{15}{64} & \frac{5}{16} & \frac{15}{64} & \frac{3}{32} & \frac{1}{64} \\ \frac{1}{729} & \frac{4}{243} & \frac{20}{243} & \frac{160}{729} & \frac{80}{243} & \frac{64}{243} & \frac{64}{729} \\ \frac{1}{46656} & \frac{5}{7776} & \frac{125}{15552} & \frac{625}{11664} & \frac{3125}{15552} & \frac{3125}{7776} & \frac{15625}{46656} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In the Wright-Fisher model above we have **absorbing states** when $x = 0$ and $x = 2N$. This means that if the process ever reaches one of these states, it remains there forever.

We modify the model so that there will be no absorbing state:

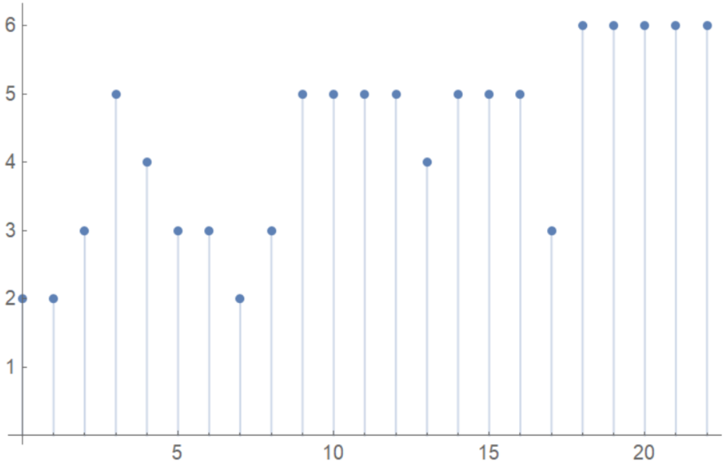


Figure: Simulation for the Wright-Fisher model, $2N = 6$, starting from 2

Wright-Fisher model with mutations

Example 7.9

In this model every gene can mutate before creating the new generation. An a can mutate into A with probability α_1 and the reverse side has probability α_2.
In this case the transition matrix is the same, but now, for the mutation, the probabilities are modified.

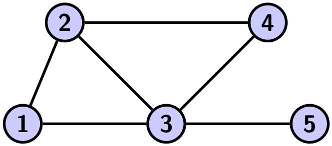
$$p_j = \frac{j}{2N}(1-\alpha_1) + \left(1 - \frac{j}{2N}\right)\alpha_2,$$

and

$$q_j = \frac{j}{2N}\alpha_1 + \left(1 - \frac{j}{2N}\right)(1-\alpha_2).$$

Simple RW on simple graphs

Example 7.10



	1	**2**	**3**	**4**	**5**
1	0	1/2	1/2	0	0
2	1/3	0	1/3	1/3	0
3	1/4	1/4	0	1/4	1/4
4	0	1/2	1/2	0	0
5	0	0	1	0	0

Simple graph and the transition matrix of the corresponding simple random walk (RW) on this graph. From every vertex we move to a uniformly chosen neighbour. (Described more precisely on the next slide.)

Let $G = (V, E)$ be a simple graph (no loops, no double edges), where as usual, V is the set of vertices and E is the set of edges. We denote the degree of vertex $x \in V$ by $\deg(x)$. The simple random walk on G is Markov chain on state space S which is defined by the following transition matrix:

$$p(x, y) = \begin{cases} \frac{1}{\deg(x)}, & (x, y) \in E; \\ 0, & \text{otherwise.} \end{cases} \tag{32}$$

Example 7.10 (Cont.)

Using the mathematica 11 code on the next slide we obtain that the stationary distribution: $\pi = \left(\frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{1}{6}, \frac{1}{12}\right)$ (the last command on the next slide results the 5-th component of π). The mean first passage matrix is $M = (m_{i,j})_{i,j=1}^{5}$, where $m_{i,j}$ is the expected number (≥ 1) of steps to get from i to j for the first time.

$$M = \begin{pmatrix} 6 & \frac{11}{4} & \frac{9}{4} & 6 & \frac{53}{4} \\ \frac{19}{4} & 4 & \frac{5}{2} & \frac{19}{4} & \frac{27}{2} \\ \frac{21}{4} & \frac{7}{2} & 3 & \frac{21}{4} & 11 \\ 6 & \frac{11}{4} & \frac{9}{4} & 6 & \frac{53}{4} \\ \frac{25}{4} & \frac{9}{2} & 1 & \frac{25}{4} & 12 \end{pmatrix}.$$

Mean First Passage Time Matrix

$M = (m_{i,j})$ and we know the diagonal: $m_{i,i} = \frac{1}{\pi_i}$. In general we need to solve the system of equations for all $i \neq j$:

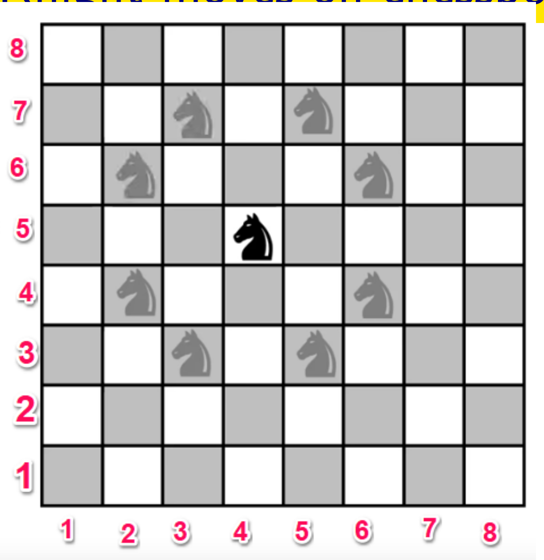
$$m_{i,j} = p_{i,j} \cdot 1 + \sum_{k \neq j} p_{i,k} \cdot (1 + m_{k,j}) = 1 + \sum_{k \neq j} p_{i,k} \cdot m_{k,j}.$$

Example 7.10 (Cont.)

```
In[173]:= 𝒫 = DiscreteMarkovProcess[{0, 0, 1, 0, 0},
  {{0, 1/2, 1/2, 0, 0},
   {1/3, 0, 1/3, 1/3, 0},
   {1/4, 1/4, 0, 1/4, 1/4},
   {0, 1/2, 1/2, 0, 0},
   {0, 0, 1, 0, 0}}]

Out[173]= DiscreteMarkovProcess[{0, 0, 1, 0, 0},
  {{0, 1/2, 1/2, 0, 0}, {1/3, 0, 1/3, 1/3, 0}, {1/4, 1/4, 0, 1/4, 1/4}, {0, 1/2, 1/2, 0, 0}, {0, 0, 1, 0, 0}}]

In[174]:= 𝒟 = FirstPassageTimeDistribution[𝒫, 4];
          Mean[𝒟]

Out[175]= 21/4

In[176]:= PDF[StationaryDistribution[𝒫], 5]

Out[176]= 1/12
```

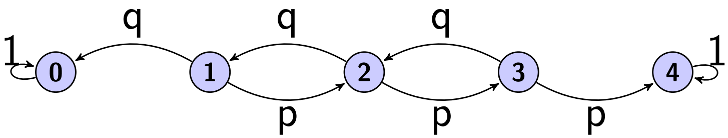
The second and third commands computes the value $m_{3,4} = \frac{21}{4}$. The last command yields that $\pi(5) = \frac{1}{12}$.

Knight moves on chessboard



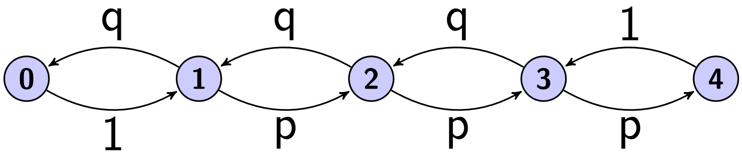
Simple RW on the graph G, where $G = (E, V)$:
$V := \{1, \ldots, 8\}^2$,
and for $(i_1, j_1), (i_2, j_2) \in V$
$((i_1, j_1), (i_2, j_2)) \in E$
iff either:
$|i_1 - i_2| = 2 \& |j_1 - j_2| = 1$
or
$|j_1 - j_2| = 2 \& |i_1 - i_2| = 1$

RW with absorbing boundary :



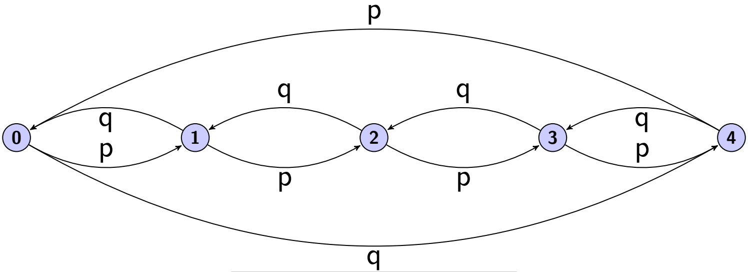
	0	**1**	**2**	**3**	**4**
0	1	0	0	0	0
1	q	0	p	0	0
2	0	q	0	p	0
3	0	0	q	0	p
4	0	0	0	0	1

RW with reflecting boundary



	0	**1**	**2**	**3**	**4**
0	0	1	0	0	0
1	q	0	p	0	0
2	0	q	0	p	0
3	0	0	q	0	p
4	0	0	0	1	0

RW with periodic boundary conditions



	0	**1**	**2**	**3**	**4**
0	0	p	0	0	q
1	q	0	p	0	0
2	0	q	0	p	0
3	0	0	q	0	p
4	p	0	0	q	0

[1] Sven Erick Alm
Sven Erick Alm
Simple random walk. Click here for the online version.

[2] R. Durrett
Essentials of Stochastic Processes, Second edition
Springer, 2012. Click here to download from the authors website the almost-ready version.

[3] R. Durrett
Probability: Theory and Examples
2019. Click here to download from the authors website

[4] CHARLES M. GRINSTEAD, J. LAURIE SNELL
Itroduction to Probability
Click here for the book.

[5] S. KARLIN, H.M. TAYLOR
A first course in stochastic processes
Academic Press,2014

[6] D.A. LEVIN, Y. PERES,E.L. WILMER
Markov chains and mixing times
American Mathematical Society, 2009.
Click here for the book.

[7] G. Lawler
Intoduction to Stochastic Processes
Chapman & Hall 1995.

[8] E. Senata
Non-negative Matrices and Markov Chains
Springer 1973.