

Markov Chains II

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 / 178

The history of Branching Processes

In 1873 Francis Galton asked in Educational Times: what is the probability of dying off of a name, a family dying agnatically? Reverend Henry William Watson answered it and they published a paper together in 1874: On the probability of extinction of families. Thus the correspondent MC is called Galton-Watson process. So we only regard the number of sons in various generations, because they carry on the name.

2 / 178

Branching processes

Let's regard a population, in which the 0^{th} generation only consists of one person and in the n^{th} generation one gives birth to k children (who will be counted in the $(n+1)^{st}$ generation) with probability p_k (independently of each other); with $k = 0, 1, 2, \dots$

Let X_n be the number of individuals in the n^{th} generation. The state space is $\mathbb{N} = \{0, 1, 2, \dots\}$. If Y_1, Y_2, \dots are i.i.d. random variables for which $\mathbb{P}(Y_m = k) = p_k$, then the transition matrix is $p(0, 0) = 1$ and $p(i, j) := \mathbb{P}(Y_1 + \dots + Y_i = j)$ if $i > 0$ and $j \geq 0$.

3 / 178

Special case: The number of children has geometric distribution.

$$p_\ell := \mathbf{P}(\text{number of children} = \ell) = q^\ell p.$$

Then element (k, l) of the transition matrix:

$$p(k, \ell) = \binom{k + \ell - 1}{\ell} p^n q^k.$$

4 / 178

Random walks on \mathbb{Z}^d

Simple symmetric random walk on $S = \mathbb{Z}^d$:

$$(1) \quad p(x, y) := \begin{cases} \frac{1}{2d}, & \text{ha } \|\mathbf{x} - \mathbf{y}\| = 1; \\ 0, & \text{otherwise.} \end{cases}$$

General random walk on $S = \mathbb{Z}^d$:

$p: \mathbb{Z}^d \rightarrow [0, 1]$; $\sum_{\mathbf{x} \in \mathbb{Z}^d} p(\mathbf{x}) = 1$, and the transition matrix $\mathbf{P} = (p(x, y))$:

$$p(x, y) := p(\mathbf{x} - \mathbf{y}).$$

5 / 178

Two stage Markov chains

In this example X_{n+1} is dependent of (X_{n-1}, X_n) .

Basketball chain

Consider a basketball player who makes a shot with the following probabilities:

1/2, if both of his previous shots are missed

2/3, if he has hit one of his last two shots

3/4, if he has hit both of his last two shots.

So let $X_n = S$ denote the success and $X_n = M$ denote the miss.

6 / 178

Two stage Markov chains (cont.)

The state space is: $\{SS, SM, MS, MM\}$ and the transition matrix is:

	SS	SM	MS	MM
SS	3/4	1/4	0	0
SM	0	0	2/3	1/3
MS	2/3	1/3	0	0
MM	0	0	1/2	1/2

Explanation: If $(X_{n-1}, X_n) = (S, M)$, then the probability of $(X_n, X_{n+1}) = (M, S)$ is equal to 2/3.

7 / 178

Stationary distribution for the Basketball chain

Following the rule shown above to compute stationary distribution π , we subtract 1 from transition matrix \mathbf{P} 's diagonal elements and replace the last column with ones.

$$A = \begin{bmatrix} -1/4 & 1/4 & 0 & 1 \\ 0 & -1 & 2/3 & 1 \\ 2/3 & 1/3 & -1 & 1 \\ 0 & 0 & 1/2 & 1 \end{bmatrix}$$

8 / 178

Stationary distribution for the Basketball chain (cont.)

$$\text{Then } A^{-1} = \begin{pmatrix} -\frac{13}{6} & -\frac{5}{16} & \frac{11}{16} & \frac{43}{31} \\ -\frac{1}{6} & -\frac{1}{16} & -\frac{1}{16} & \frac{24}{31} \\ -1 & -\frac{3}{8} & -\frac{3}{8} & \frac{24}{7} \\ \frac{1}{2} & \frac{3}{16} & \frac{3}{16} & \frac{4}{8} \end{pmatrix}.$$

Its last row is π . Hence,

$$\pi = \left(\frac{1}{2}, \frac{3}{16}, \frac{3}{16}, \frac{1}{8}\right).$$

9 / 178

Stationary distribution for the Basketball chain (cont.)

Reminder: the order of components is (SS,SM,MS,MM). (S: success, M: miss.) So, in the long term the ratio of successes is:

$$\pi_{SS} + \pi_{KS} = \pi_1 + \pi_3 = \frac{1}{2} + \frac{3}{16} = \frac{11}{16}.$$

10 / 178

Ehrenfest chain: Stationary distribution

Example 0.1 (π for the Ehrenfest chain)

Recall the definition of the Ehrenfest chain: Consider the Markov Chain with state space $S := \{0, 1, 2, \dots, n\}$ and

- It jumps from 0 to 1 and from n to $n-1$ with probability 1.
- For any $0 < i < n$, it jumps from i to $i-1$ with probability i/n and from i to $i+1$ with probability $1 - \frac{i}{n}$.

11 / 178

Ehrenfest chain: Stationary distribution (cont.)

Now compute the stationary state for this chain. The transition matrix:

$$P := \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{n} & 0 & \frac{n-1}{n} & 0 & \cdots & 0 \\ 0 & \frac{2}{n} & 0 & \frac{n-2}{n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \frac{n-1}{n} & 0 & \frac{1}{n} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

12 / 178

Ehrenfest chain: Stationary distribution (cont.)

For $\pi^T \cdot P = \pi^T$, thus using notation $\pi_{-1} := \pi_{n+1} := 0$ we obtain that:

$$(2) \quad \pi_{k-1} \left(1 - \frac{k-1}{n}\right) + \pi_{k+1} \frac{k+1}{n} = \pi_k, \quad k = 0, 1, \dots, n.$$

We introduce the generating function:

$$(3) \quad g(x) = \sum_{k=0}^n x^k \pi_k.$$

13 / 178

Ehrenfest chain: Stationary distribution (cont.)

Multiply both sides of (2) by n and x^k , then sum it up for k from 1 to n :

$$\sum_{k=1}^n (n-k+1)x^k \pi_{k-1} + \sum_{k=0}^{n-1} \pi_{k+1} (k+1)x^k = n \underbrace{\sum_{k=0}^n x^k \pi_k}_{g(x)}.$$

By obvious manipulations of this formula we obtain:

$$(1+x)g'(x) = ng(x).$$

14 / 178

Ehrenfest chain: Stationary distribution (cont.)

After solving this differential equation we get:

$$g(x) = C(1+x)^n.$$

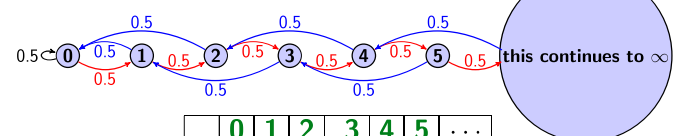
Using that π is a probability vector we get $g(1) = 1$. Hence $C = 2^{-n}$. That is:

$$(4) \quad g(x) = 2^{-n} (1+x)^n = 2^{-n} \sum_{k=0}^n \binom{n}{k} x^k.$$

Compare this to (3) to realize that $\pi_k = 2^{-n} \binom{n}{k}$.

15 / 178

Two steps back, one step forward chain



	0	1	2	3	4	5	...
0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	...
1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	0	...
2	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	0	...
3	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	...
4	0	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$...
...

16 / 178

π for the two steps back one step ahead chain:

From the equation $\pi^T \cdot P = \pi^T$: $\pi_0 = \frac{1}{2}(\pi_0 + \pi_1 + \pi_2)$ and $\forall k \geq 1: \pi_k = \frac{1}{2}(\pi_{k-1} + \pi_{k+2})$. From these two equations it comes by induction that

$$(5) \quad \forall k \geq 0: \pi_k = \pi_{k+1} + \pi_{k+2}.$$

It is obviously satisfied by $\pi_k = (1 - \rho)\rho^k$, $k \geq 0$, where ρ is the golden ratio: $\rho = \frac{\sqrt{5}-1}{2}$. Homework: there is no other stationary distribution. So the process spends most of its time (more than 99%) in the set $\{0, 1, \dots, 9\}$.

17 / 178

Definition 0.2

A MC is doubly stochastic if its probability matrix's column sum equals to 1. $\sum_j p(i, j) = 1, \forall i$.

Theorem 0.3

A MC with finite state space is doubly stochastic iff its stationary distribution is the uniform distribution.

Proof.

Let us assume that $\#S = N$, then

$$\sum_x \pi(x) p(x, y) = \frac{1}{N} \sum_x p(x, y) = \frac{1}{N} = \pi(y).$$

18 / 178

Examples

Example 0.4 (Random walk with periodic boundary conditions)

Recall the definition of the random walk with periodic boundary conditions from slide 144 in File MC I. It is obviously doubly stochastic.

19 / 178

Modulo 6 jumps on a circle

Example 0.5

We roll the finite number series $0, 1, 2, \dots, 5$ on to a circle so that 5 and 0 be neighbours. Then we use such a regular dice which has number

- 1 on **three** sides,
- 2 on **two** sides,
- 3 on **one** side.

We move forward as much as we scored (modulo 6).

20 / 178

Modulo 6 jumps on a circle (cont.)

The transition matrix is:

	0	1	2	3	4	5
0	0	1/2	1/3	1/6	0	0
1	0	0	1/2	1/3	1/6	0
2	0	0	0	1/2	1/3	1/6
3	1/6	0	0	0	1/2	1/3
4	1/3	1/6	0	0	0	1/2
5	1/2	1/3	1/6	0	0	0

It can easily be seen that the elements of transition matrix's third power P^3 are positive. Thus we see that

21 / 178

Modulo 6 jumps on a circle (cont.)

the chain is irreducible and aperiodic, so the conditions of Convergence Theorem (Theorem 6.2 in File MC I) are satisfied (obviously $\pi(i) = 1/6, \forall i$).

22 / 178

1	Recurrence in case of countable infinite state space
2	Detailed balance condition and related topics
	• Detailed balance condition and Reversible Markov Chains
	• Birth and death processes
3	Absorbing Chains
	• Exit distributions through examples
	• Exit time through examples
	• Summary and the general theory
	• Application to Irreducible chains
	• All of these with Mathematica
4	Branching Processes
	• Generator functions
	• Branching Processes
5	Stopping time, Strong Markov property
6	Transient and recurrent states

23 / 178

Recurrence of the simple Symmetric random walk in \mathbb{Z}^d -ben

Theorem 1.1

In \mathbb{R}^d the simple symmetric random walk is **recurrent (zero recurrent) if $d = 1$ or $d = 2$** but **transient for $d \geq 3$** .

We will give the proof in the case when $d = 1$.

24 / 178

One needs to be careful

In the case of countably infinite state space it can happen that there are no recurrent states as the following trivial example shows

Example 1.2 (Monotone increasing MC)

Let S be the set of non-negative integers and $p(i, i+1) := 1$ for all $i \in S$.

25 / 178

Detailed balance condition

π satisfies **detailed balance condition**, if $\forall x, y$

$$(6) \quad \pi(x)p(x, y) = \pi(y)p(y, x)$$

If we sum both sides for y , we get that

$$\sum_y \pi(y)p(y, x) = \pi(x) \underbrace{\sum_y p(x, y)}_{=1} = \pi(x).$$

So, if a probability measure satisfies formula (6), then it is a stationary distribution. There exist stationary

27 / 178

Detailed balance condition (cont.)

$\pi(3) = 0$. This and formula (6) yield that $\pi(2) = \pi(1) = 0$ which is impossible. On the other hand, \mathbf{P} is a doubly stochastic matrix for which there is a stationary distribution (the uniform distribution): $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

So, it can happen that there is a stationary distribution but it does not satisfy (6). In spite of this, if we have a guess about a probability vector that it could be the stationary distribution, we can check it easily by substituting it into formula (6).

29 / 178

Reversible MC (cont.)

It comes easily from formula (6) that:

$$(10) \quad \pi(x_0)p(x_0, x_1) \cdots p(x_{n-1}, x_n) = \pi(x_n)p(x_n, x_{n-1}) \cdots p(x_1, x_0).$$

Using notation $\mathbf{x} = (x_0, \dots, x_n)$ this implies that:

$$(11) \quad \mathbb{P}_\pi(X_0^n = \mathbf{x}) = \mathbb{P}_\pi(X_0^n = \overleftarrow{\mathbf{x}}).$$

So if MC (X_n) has stationary distribution, **and** it satisfies **detailed balance condition**, then **the distribution of**

31 / 178

- 1 Recurrence in case of countable infinite state space
- 2 Detailed balance condition and related topics
 - Detailed balance condition and Reversible Markov Chains
 - Birth and death processes
- 3 Absorbing Chains
 - Exit distributions through examples
 - Exit time through examples
 - Summary and the general theory
 - Application to Irreducible chains
 - All of these with Mathematica
- 4 Branching Processes
 - Generator functions
 - Branching Processes
- 5 Stopping time, Strong Markov property
- 6 Transient and recurrent states

26 / 178

Detailed balance condition (cont.)

distributions which do not satisfy the detailed balance condition (6). For example, consider the MC whose probability matrix is:

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.1 & 0.6 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}.$$

Then the stationary distribution π of P does not satisfy (6). To get contradiction, assume that π satisfies (6). From this and from the fact that $p(1, 3) = 0$ we get

28 / 178

Reversible MC

Now we use [4, chapter 1.6]. **Notation:** For the MC (X_n) we introduce

$$(7) \quad X_0^n := (X_0, \dots, X_n).$$

So for $\mathbf{x} := (x_0, \dots, x_n)$

$$(8) \quad \{X_0^n = \mathbf{x}\} = \{X_0 = x_0, \dots, X_n = x_n\}$$

and for an $\mathbf{x} = (x_0, \dots, x_n)$ let

$$(9) \quad \overleftarrow{\mathbf{x}} := (x_n, x_{n-1}, \dots, x_1, x_0).$$

30 / 178

Reversible MC (cont.)

(X_0, \dots, X_n) is the same as the distribution of (X_n, \dots, X_0) .

Definition 2.1 (reversible MC)

A MC X_n is **reversible** if it has stationary distribution π and **π satisfies the detailed balance condition**, that is formula (6) holds.

32 / 178

Reversible MC (cont.)

Example 2.2 (Simple random walk on graphs, slide 137 in File MC I)

Let us regard a simple random walk on graph $G = (V, E)$. Using notation of slide 137 in File MC I, the stationary distribution is: $\pi(y) = \deg(y)/2\#E$. It can be easily seen (homework) that it satisfies detailed balance condition:

$$\pi(x)p(x, y) = \pi(y)p(y, x), \quad \forall x, y \in S.$$

33 / 178

Definition 2.4 (Chain with reversed time)

Given an irreducible MC X_n with transition matrix \mathbf{P} and stationary distribution π . Let us define the matrix $\hat{\mathbf{P}} = (\hat{p}(x, y))$:

$$(12) \quad \hat{p}(x, y) := \frac{\pi(y)p(y, x)}{\pi(x)}.$$

Then $\hat{\mathbf{P}}$ is a stochastic matrix (every element is non-negative, the row-sums are 1.) So $\hat{\mathbf{P}}$ determines a MC (\hat{X}_n) , which we call **time reversal** of (X_n) .

Obviously, if (X_n) is reversible, then $\mathbf{P} = \hat{\mathbf{P}}$.

35 / 178

Time reversal (cont.)

Proof.

Firstly we prove part (a):

$$\sum_y \pi(y) \hat{p}(y, x) = \sum_y \pi(y) \frac{\pi(x)p(x, y)}{\pi(y)} = \pi(x).$$

Now we see part (b):

$$\begin{aligned} \mathbb{P}_\pi(X_0^n = \mathbf{x}) &= \pi(x_0)p(x_0, x_1)p(x_1, x_2) \cdots p(x_{n-1}, x_n) \\ &= \pi(x_n)\hat{p}(x_n, x_{n-1}) \cdots \hat{p}(x_2, x_1)\hat{p}(x_1, x_0) \\ &= \mathbb{P}_\pi(\hat{X}_0^n = \hat{\mathbf{x}}). \end{aligned}$$

37 / 178

Birth and death processes (cont.)

Then the transition matrix \mathbf{P} is:

$$\begin{aligned} p(x, x+1) &= p_x \text{ if } x < n \\ p(x, x-1) &= q_x \text{ if } x > k \\ p(x, x) &= 1 - p_x - q_x \text{ if } k \leq x \leq n. \end{aligned}$$

and all other $p(x, y) = 0$. Warning: $p + q \neq 1$ is possible here!

39 / 178

Reversible MC (cont.)

Example 2.3 (Random walk with periodic boundary condition)

Reminder: finite state space (with cardinality N) rolled onto a circle. We jump 1 clockwise with probability p and anticlockwise with probability $q = 1 - p$. The chain is double stochastic, so $\pi = (\frac{1}{N}, \dots, \frac{1}{N})$. But

$\pi(k)p(k, k+1) = \frac{p}{N}$ and $\frac{q}{N} = \pi(k+1)p(k+1, k)$ and they are equal only if $p = q$. So in other instances the detailed balance condition is not satisfied.

34 / 178

Time reversal

Theorem 2.5

Using notation of Definition 2.4:

- (a) π is stationary distribution not only for (X_n) but for (\hat{X}_n) , too, and
- (b) for all \mathbf{x} :

$$(13) \quad \mathbf{P}_\pi(X_0^n = \mathbf{x}) = \mathbf{P}_\pi(\hat{X}_0^n = \hat{\mathbf{x}}),$$

where $\hat{\mathbf{x}}$ was defined in (9).

36 / 178

Birth and death processes

Birth and death processes are those MCs, whose state space are

$$S := \{k, k+1, \dots, n\}.$$

and we cannot jump more than 1. So the possible jumps are: $-1, 0, 1$. The transition probability:

$$p(x, y) = 0 \text{ if } |x - y| > 1:$$

38 / 178

Birth and death processes (cont.)

Theorem 2.6

All birth and death processes are reversible.

40 / 178

Birth and death processes (cont.)

Proof

We need to see that we can find a probability measure π on S which satisfies formula (6), thus for $x < n$ it must be true for π :

$$\pi(x+1) \underbrace{p(x+1, x)}_{q_{x+1}} = \pi(x) \underbrace{p(x, x+1)}_{p_x}$$

So, for (6), it is needed that

$$(14) \quad \pi(x+1) = \frac{p_x}{q_{x+1}} \pi(x).$$

Iterating this for every $1 \leq i \leq n-k$

41 / 178

We have computed the stationary distribution for the Ehrenfest Chain (see slide 11). We got that $\pi(k) = 2^{-N} \binom{N}{k}$, but we needed an unpleasant reduction involving generator functions. Now we can easily get this from formula (15) because the Ehrenfest Chain is obviously a birth and death process.

43 / 178

Proof Cont.

$$(15) \quad \pi(k+i) = \pi(k) \cdot \underbrace{\frac{p_{k+i-1} \cdot p_{k+i-2} \cdots p_{k+1} \cdot p_k}{q_{k+i} \cdot q_{k+i-1} \cdots q_{k+2} \cdot q_{k+1}}}_{r_i}$$

It is easy to see that if we choose $\pi(k)$ such way that

$$(16) \quad \pi(k) \cdot \left(1 + \sum_{i=1}^{n-k} r_i\right) = 1,$$

then π is a stationary distribution which satisfies the detailed balance condition, so the chain is reversible.

42 / 178

Example 2.7 (π for the Ehrenfest chain)

Here: $S = \{0, 1, \dots, N\}$. From formula (15) we get that

$$r_i = \binom{N}{i} \text{ if } 1 \leq i \leq N.$$

Using that $1 + \sum_{i=1}^N r_i = 2^N$ we obtain that for $i = 0, \dots, N$:

$$\pi(i) = 2^{-N} \binom{N}{i}.$$

44 / 178

45 / 178

Two year collage

Example 3.1 (Two year collage)

At a two year collage the first year students are called freshmen the second year students are the sophomores.

- **Freshmen:** 60% of them become sophomores, 25% of them remain freshmen, 15% of them exit (E) so leave the school.
- **Sophomores:** 70% of them complete the courses with Success (S), 20% of them remain sophomores and 10% of them exit.

46 / 178

Two year collage (cont.)

Then if $S = \{1, 2, G, D\}$ (freshmen, sophomores, Graduate, Drop out) and X_n shows that a student is in which state after n years, then X_n is a MC whose state space is S and its transition matrix:

	1	2	G	D
1	0.25	0.6	0	0.15
2	0	0.2	0.7	0.1
G	0	0	1	0
D	0	0	0	1

47 / 178

Two year collage (cont.)

Let $h(x)$, $x \in S$ be the probability that a student in state x eventually graduates. Then we apply the one step reasoning method. Namely, we do not know $h(1)$ and $h(2)$ but after making one step on the chain the following equations hold:

$$\begin{aligned} h(1) &= 0.25h(1) + 0.6h(2) \\ h(2) &= 0.2h(2) + 0.7. \end{aligned}$$

From this $h(2) = 7/8$ and $h(1) = 0.7$.

48 / 178

Theorem 3.2

A MC is given with a finite state space S . Let $a, b \in S$ and $C := S \setminus \{a, b\}$. Let $h : S \rightarrow \mathbb{R}^+$ be a function satisfying:

$$(17) \quad h(a) = 1, \quad h(b) = 0, \quad \forall x \in C : \quad h(x) = \sum_{y \in S} p(x, y) h(y).$$

Put

$$V_y = \min \{n \geq 0 : X_n = y\}.$$

Assume that $\forall x \in C : \mathbb{P}_x(V_a \wedge V_b < \infty) = 1$. Then

$$h(x) = \mathbb{P}_x(V_a < V_b).$$

49 / 178

Proof

We frequently use the shorthand notation

$$a \wedge b := \min \{a, b\}.$$

Let $T := V_a \wedge V_b$. By assumption

$$(18) \quad \forall x \in C, \quad \mathbb{P}_x(T < \infty) = 1.$$

50 / 178

Proof (cont.)

First we express the probability $\mathbb{P}_x(V_a < V_b)$ in terms of the expectation of a random variable. Namely, note that by definition,

$$h(X_T) = \begin{cases} 1, & \text{if } V_a < V_b; \\ 0, & \text{if } V_b < V_a. \end{cases}$$

That is

$$(19) \quad h(X_T) = \mathbb{1}_{\{V_a < V_b\}}$$

51 / 178

Proof (cont.)

Hence, for all $x \in C$ we have

$$(20) \quad \mathbb{P}_x(V_a < V_b) = \mathbb{E}_x[h(X_T)]$$

Now we prove that

$$(21) \quad \mathbb{E}_x[h(X_T)] = \lim_{n \rightarrow \infty} \mathbb{E}_x[h(X_{T \wedge n})].$$

52 / 178

Proof (cont.)

To see this, recall that we assumed that the state space $\#S < \infty$. So, $M := \max_{x \in S} h(x) < \infty$. That is, on the one hand, for all $x \in C$,

$$(22) \quad h(X_{T \wedge n}) < M \text{ holds for all } n.$$

On the other hand, using (18) (which says that T is almost surely finite) we have that

$$(23) \quad \lim_{n \rightarrow \infty} h(X_{T \wedge n}) = h(X_T).$$

53 / 178

Proof (cont.)

Putting together (23) and (22), we obtain that (21) holds by **Lebesgue Dominated Convergence Theorem**. Finally, we verify that

$$(24) \quad \mathbb{E}_x h(X_{T \wedge n}) = h(x), \quad \forall n > 1, \quad \forall x \in C.$$

$$u_1(x, a) := \mathbb{P}_x(X_1 = a) = p(x, a) \text{ and for } k \geq 2$$

$$\begin{aligned} u_k &= \mathbb{P}_x(X_k = a, T = k) \cdot \underbrace{h(a)}_1 \\ &= \sum_{x_1, \dots, x_{k-1} \in C} p(x, x_1) p(x_1, x_2) \cdots p(x_{k-1}, a). \end{aligned}$$

54 / 178

Proof (cont.)

Moreover, let $S_0 := h(x)$ and

$$S_k := \sum_{x_1, \dots, x_k \in C} p(x, x_1) p(x_1, x_2) \cdots p(x_{k-1}, x_k) h(x_k).$$

A careful case analysis yields that by (17) for $k \geq 1$:

$$(25) \quad S_k = S_{k-1} - u_k.$$

Observe that for a $k \leq n$ we have

$$(26) \quad \mathbb{E}_x[h(X_{T \wedge n}), T = k] = \mathbb{P}_x(X_k = a, T = k).$$

55 / 178

Proof (cont.)

Using (25), (26), a telescoping sum in the third step and the fact that $S_0 = h(x)$ we obtain:

$$\mathbb{E}_x[h(X_{T \wedge n})] = \mathbb{E}_x[h(X_{T \wedge n}), T > n] + \sum_{k=1}^n \mathbb{E}_x[h(X_{T \wedge n}), T=k]$$

$$(27) \quad = S_n + \sum_{k=1}^n u_k$$

$$(28) \quad = \underbrace{h(x)}_{S_0} + \sum_{k=1}^n \underbrace{(S_k - S_{k-1})}_{-u_k} + \sum_{k=1}^n u_k$$

$$(29) \quad = h(x). \blacksquare$$

56 / 178

Wright-Fisher model

This was introduced on slide 131 in File MC I.

The state space: $S = \{0, 1, \dots, 2N\}$. The absorbing states: 0 and $2N$. Question: what is the probability of ending up in $2N$, or in the model's language: what is the probability that once every gene becomes type a ?

The transition matrix:

$$p(x, y) = \underbrace{\binom{2N}{y} \left(\frac{x}{2N}\right)^y \left(1 - \frac{x}{2N}\right)^{2N-y}}_{\text{Binomial}(2N, x/2N)}.$$

57 / 178

Wright-Fisher model (cont.)

That is: the distribution of $y \in \{0, 1, \dots, 2N\}$ where the Markov chain jumps to from $x \in \{0, 1, \dots, 2N\}$ is a $\text{Binomial}(2N, x/2N)$ random variable. We know that expected value of a $\text{Binomial}(2N, x/2N)$ r.v. is equal to x . The same in formula:

$$(30) \quad x = \sum_{y=0}^{2N} p(x, y) \cdot y$$

58 / 178

Wright-Fisher model (cont.)

Let us define a function: $h(t) := \frac{t}{2N}$, then by (30):

$$h(x) = \sum_{y=0}^{2N} p(x, y) h(y).$$

Let $a = 2N$ and $b = 0$. Then $h(a) = 1$ and $h(b) = 0$. Obviously:

$$\mathbb{P}_x(V_a \wedge V_b < \infty) > 0, \quad \forall 0 < x < N.$$

59 / 178

Wright-Fisher model (cont.)

So, we can use Theorem 3.2, thus we get:

$$\mathbb{P}_x(V_{2N} < V_0) = h(x) = \frac{x}{N}. \blacksquare$$

In summary: here we guessed the exit probability function $h(x)$ and to verify our guess we used Theorem 3.2.

60 / 178

Example: Gambler's ruin, unfair case

Now we use the notation introduced on slide 4 in File MC I, where the Gambler's ruin example was introduced with the modification that now $p \neq 1/2$ is arbitrary. Let

$$h(x) = \mathbb{P}_x(V_N < V_0).$$

That is $h(x)$ is the probability that a gambler starting with $\$x$ eventually wins, that is reaches $\$N$ earlier than $\$0$. Obviously, $h(N) = 1$ and $h(0) = 0$. As usual, let

61 / 178

Example: Gambler's ruin, unfair case (cont.)

$q := 1 - p$ and let $0 < x < N$. Yet again we use the one-step argument: After one step:

$$X_{n+1} = \begin{cases} x + 1, & \text{with probability } p; \\ x - 1, & \text{with probability } q. \end{cases}$$

So, for $0 < x < N$:

$$(31) \quad h(x) = ph(x+1) + qh(x-1).$$

62 / 178

Example: Gambler's ruin, unfair case (cont.)

Obvious manipulations yield:

$$p(h(x+1) - h(x)) = q(h(x) - h(x-1)).$$

Hence,

$$(32) \quad h(x+1) - h(x) = \frac{q}{p}(h(x) - h(x-1))$$

63 / 178

Example: Gambler's ruin, unfair case (cont.)

Let $c := h(1) - h(0)$. So, from formula (32) for $x \geq 1$

$$(33) \quad h(x) - h(x-1) = c \left(\frac{q}{p}\right)^{x-1}.$$

Using that $h(N) = 1$, $h(0) = 0$ and a telescopic sum in the second step and (33) in the last step:

$$1 = h(N) - h(0) = \sum_{x=1}^N h(x) - h(x-1) = c \sum_{x=1}^N \left(\frac{q}{p}\right)^{x-1}.$$

64 / 178

Example: Gambler's ruin, unfair case (cont.)

Put $\theta = q/p$. Then $c = (1 - \theta)/(1 - \theta^N)$. So

$$(34) \quad h(x) = h(x) - h(0) = c \sum_{i=0}^{x-1} \theta^i = \frac{1 - \theta^x}{1 - \theta^N}.$$

From here if $N \rightarrow \infty$ we get that

$$(35) \quad p > \frac{1}{2} \Rightarrow \mathbb{P}_x(V_0 = \infty) = 1 - \left(\frac{q}{p}\right)^x.$$

65 / 178

Example: Gambler's ruin, unfair case (cont.)

Corollary 3.3

Consider a random walk on \mathbb{Z} , in which starting from all $x > 0$ we go forward one step with probability $p > \frac{1}{2}$ and we go backward one step with probability $q = 1 - p$. Then the probability that starting from an arbitrary $x > 0$ we never reach 0 is $1 - \left(\frac{q}{p}\right)^x > 0$. That is every state is transient.

66 / 178

Example: Gambler's ruin, fair case

We consider the Gambler's ruin example with $p = 1/2$. We use the unfair case ($p \neq 1/2$)'s notation. The argument is the same until formula (32).

But in case of $p = 1/2$ formula (32) shows that the gradient of function $h(x)$ is constant and $h(0) = 0$, $h(N) = 1$ so if $p = 1/2$

$$\mathbb{P}_x(V_N < V_0) = h(x) = x/N.$$

67 / 178

Tennis

The following problem is from [1, p.44].

68 / 178

Tennis (cont.)

Example 3.4

In tennis a player wins the game if either she gets 4 points when the other player has not more than 2 points. If the score is 4 - 3 then the winner is the player who makes a two points advantage first. Assume that

- The server wins the point with 0.6 probability,
- Successive points are independent.

Question: What is the probability that the server wins if the score now is 3 - 3?

69 / 178

Tennis (cont.)

Solution: Let X_n be the difference of the points scored from the point of the server after 3 - 3 until one of the player has a 2 point advantage so that the game ends. That is the state space is $S := \{-2, -1, 0, 1, 2\}$. Then the transition matrix:

	2	1	0	-1	-2
2	1	0	0	0	0
1	0.6	0	0.4	0	0
0	0	0.6	0	0.4	0
-1	0	0	0.6	0	0.4
-2	0	0	0	0	1

70 / 178

Tennis

Let $h(x)$ be the probability that the server wins when starting from $X_0 = x$. Obviously now the absorbing states are $\{-2, 2\}$ and $C = \{-1, 0, 1\}$. Clearly,

$$h(2) = 1 \text{ and } h(-2) = 0.$$

From the one-step reasoning:

$$(36) \quad h(x) = \sum_y p(x, y)h(y), \quad \forall x \in C.$$

71 / 178

Tennis (cont.)

$$(37) \quad \begin{aligned} h(1) &= 0.6 \cdot \underbrace{h(2)}_1 + 0.4h(0) = 0.4h(0) + 0.6 \\ h(0) &= 0.6h(1) + 0.4h(-1) \\ h(-1) &= 0.6h(0) + 0.4 \cdot \underbrace{h(-2)}_0 = 0.6h(0). \end{aligned}$$

Let $\mathbf{R} = (r(x, y))_{x, y \in C}$ be the restriction of matrix \mathbf{P} to rows and columns of C , and let $\hat{\mathbf{h}}$ be the vector which

72 / 178

Tennis (cont.)

we get by ignoring those coordinates of \mathbf{h} which are outside C . Then formula (37):

$$(38) \quad \hat{\mathbf{h}} - \mathbf{R} \cdot \hat{\mathbf{h}} = \begin{bmatrix} 0.6 \\ 0 \\ 0 \end{bmatrix}$$

Which is:

$$\underbrace{\begin{bmatrix} 1 & -0.4 & 0 \\ -0.6 & 1 & -0.4 \\ 0 & -0.6 & 0 \end{bmatrix}}_{I-R} \cdot \begin{bmatrix} h(1) \\ h(0) \\ h(-1) \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0 \\ 0 \end{bmatrix}$$

73 / 178

Tennis (cont.)

So

$$\begin{bmatrix} h(1) \\ h(0) \\ h(-1) \end{bmatrix} = (I - R)^{-1} \cdot \begin{bmatrix} 0.6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.8769 \\ 0.6923 \\ 0.4154 \end{bmatrix}$$

74 / 178

Exit time from the two year collage

Consider the two-year collage example on slide 46. There, we asked what was the probability of a $k = 1, 2$ -year-student of graduating ever. Now, for the same example we ask:

Question: On average, how much time is needed for a student to get out of the school either by completing it successfully or drop out (unsuccessfully).

Let $g(x)$ be the expected number of years that an $x \in \{1, 2\}$ -year student leaves the school either because

75 / 178

Exit time from the two year collage (cont.)

she graduates or because she drops out. We define $g(G) = g(D) = 0$. Again, we use the one-step reasoning:

$$\begin{aligned} g(1) &= 1 + 0.25g(1) + 0.6g(2) \\ g(2) &= 1 + 0.2g(2). \end{aligned}$$

This yields: $g(2) = 1.25$ and $g(1) = 2.333$.

76 / 178

Exit time

Theorem 3.5

Let X_n be a MC with a finite state space S . Let $A \subset S$ and $C := S \setminus A$, and $V_A := \min \{n \geq 0 : X_n \in A\}$. Let $g : S \rightarrow \mathbb{R}^+$ be a function which satisfies:

- (a) $\mathbb{P}_x(V_A < \infty) > 0$, $\forall x \in C$,
- (b) $g(a) = 0$, $\forall a \in A$,

77 / 178

Exit time (Cont.)

Theorem 3.5 (Cont.)

(c) $\forall x \in C$

$$(39) \quad g(x) = 1 + \sum_y p(x, y)g(y).$$

Then this function g is the expected exit time. That is

$$(40) \quad g(x) = \mathbb{E}_x[V_A].$$

Proof.

The proof goes similarly as the proof of Theorem 3.2. \square

78 / 178

Waiting time for TT

Example 3.6

We flip a fair coin until we get two Tails (TT) in a row. Question: what is the expected value of the number of flips?

Solution: We call T the Tails and H the Heads. Let T_{TT} be the (random) number of flips until we get the two Tails (the TT). Now we associate a MC (X_n) with state space $S := \{0, 1, 2\}$, where X_n is the number of consecutive Tails after the n^{th} flip. So, if the n^{th} flip

79 / 178

Waiting time for TT (cont.)

results in a Head, then $X_n = 0$, if it is a Tail, then $X_n = 1$ or $X_n = 2$ depending on X_{n-1} (if it was Head or Tail). State 2 is absorbing because we only flip the coin until this happens. So, the transition matrix:

	0	1	2
0	1/2	1/2	0
1	1/2	0	1/2
2	0	0	1

80 / 178

Waiting time for TT (cont.)

Let

$$V_2 := \min \{n \geq 0 : X_n = 2\} \text{ and } g(x) := \mathbb{E}_x[V_2].$$

Then from the one-step reasoning:

$$(41) \quad \begin{aligned} g(0) &= 1 + 0.5g(0) + 0.5g(1) \\ g(1) &= 1 + 0.5g(0). \end{aligned}$$

81 / 178

Waiting time for TT (cont.)

Let $\mathbf{1}$ be the vector in \mathbb{R}^2 , having both components equal to 1. Then $g(0) = 0$ by formula (41):

$$(42) \quad (I - R) \cdot \hat{\mathbf{g}} = \mathbf{1},$$

where, as before, R is the matrix we get from \mathbf{P} by deleting the rows and columns corresponding to the absorbing states (now the only absorbing state is 2) and $\hat{\mathbf{g}}$ is the vector we get from vector \mathbf{g} by deleting the

82 / 178

Waiting time for TT (cont.)

components belonging to the absorbing states which is 2 as mentioned before. Hence from (42) we get

$$\hat{\mathbf{g}} = \begin{pmatrix} g(0) \\ g(1) \end{pmatrix} = (I - R)^{-1} \cdot \mathbf{1} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \cdot \mathbf{1} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

So, by Theorem 3.5, we have

$$\mathbb{E}_0[V_2] = g(0) = \hat{g}(0) = 6.$$

83 / 178

Tennis at 3 – 3

Consider the Tennis problem on slide 68 again.

Question: How long the game lasts if now the score is 4 – 3, 3 – 3 and 3 – 4 from the point of the server?

Solution: Let $g(x)$ be the expected time of the game if $x \in \{1, 0, -1\}$. As we discussed, the absorbing states are $A := \{-2, 2\}$ and the state space is $S := \{-2, -1, 0, 1, 2\}$. So, $C : A \setminus A = \{1, 0, -1\}$. Using notation analogue to the previous problem:

$$\mathbf{R} = \begin{bmatrix} 0 & 0.4 & 0 \\ 0.6 & 0 & 0.4 \\ 0 & 0.6 & 0 \end{bmatrix}$$

84 / 178

Tennis at 3 – 3 (cont.)

and from here:

$$I - R = \begin{bmatrix} 1 & -0.4 & 0 \\ -0.6 & 1 & -0.4 \\ 0 & -0.6 & 1 \end{bmatrix}$$

So, like the previous problem:

$$\begin{pmatrix} g(1) \\ g(0) \\ g(-1) \end{pmatrix} = (I - R)^{-1} \mathbf{1} = \begin{bmatrix} 19/13 & 10/13 & 4/13 \\ 15/13 & 25/13 & 10/13 \\ 9/13 & 15/13 & 19/13 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

85 / 178

Tennis at 3 – 3 (cont.)

So, at 3 – 3 the expected play-time:

$$(43) \quad g(0) = \frac{15 + 25 + 10}{13} = 3.846.$$

86 / 178

Tennis at 3 – 3 (cont.)

Remark 3.7

Consider an absorbing MC with state space S , absorbing states A and transient states $C := S \setminus A$. Let $y \in C$ and we denote the total number of visit to y including the time 0 if we started from y by $N(y)$. The

$$N(y) = \sum_{n=0}^{\infty} \mathbb{1}_{X_n=y}. \text{ In this way}$$

87 / 178

Tennis at 3 – 3 (cont.)

Remark 3.7 (Cont.)

$$(44) \quad \mathbb{E}_x[N(y)] = \sum_{n=0}^{\infty} R^n(x, y) = (I - R)^{-1}(x, y).$$

Let T be the duration until the chain gets into an absorbing state. This is equal to the total time the MC spends at all of the transient states together. That is

$$T = \sum_{y \in C} N(y).$$

Hence by (44)

88 / 178

Tennis at 3 – 3 (cont.)

Remark 3.7 (Cont.)

$$(45) \mathbb{E}_x[T] = \sum_{y \in C} \mathbb{E}_x[N(y)] = \sum_{y \in C} (I - R)^{-1}(x, y),$$

which is the x -th component of the vector

$$(I - R)^{-1} \cdot \mathbf{1}.$$

With this argument we proved that $(I - R)^{-1}(x, y)$ is equal to the expectation of the number of visits to y (counting the initial state if $x = y$) starting from x .

89 / 178

Tennis at 3 – 3 (cont.)

As a Corollary of this Remark we can see that in (43) the summands

$$\frac{15}{13}, \frac{25}{13}, \frac{10}{13}$$

are the expected number of cases when the score is 1, 0, -1 respectively, before the game ends.

90 / 178

Gambler's ruin, $p = 1/2$: How long does it last?

So: $p(i, i+1) = p(i, i-1) = 1/2$. $A := \{0, N\}$,

$$V_A := \min \{n \geq 0 : X_n \in A\}.$$

Let $g(x) := \mathbb{E}_x[V_A]$. Obviously

$$(46) \quad g(0) = g(N) = 0.$$

91 / 178

Gambler's ruin, $p = 1/2$: How long does it last? (cont.)

If $0 < x < N$:

$$g(x) = 1 + \frac{1}{2}g(x+1) + \frac{1}{2}g(x-1)$$

$$g(x+1) - g(x) = g(x) - g(x-1) - 2.$$

If $c = g(1) - g(0)$, then

$$(47) \quad g(k) - g(k-1) = c - 2(k-1)$$

92 / 178

Gambler's ruin, $p = 1/2$: How long does it last? (cont.)

Using that $g(N) = 0$ and summing the previous equations for $1 \leq k \leq N$, we get telescopic sums. From these:

$$\begin{aligned} 0 &= g(N) = \sum_{k=1}^N (g(k) - g(k-1)) \\ &= \sum_{k=1}^N (c - 2(k-1)) = cN - 2\frac{N(N-1)}{2}. \end{aligned}$$

93 / 178

Gambler's ruin, $p = 1/2$: How long does it last? (cont.)

Hence, $c = N - 1$. Substituting this back to formula (47) and summing it up we obtain that:

$$g(x) = x(N - x).$$

94 / 178

Gambler's ruin, $p \neq 1/2$: How long does it last?

So, in this case: $p(i, i+1) = p \neq 1/2$ and $p(i, i-1) = 1 - p =: q$. Let $A := \{0, N\}$, $C := \{1, \dots, N-1\}$,

$$V_A := \min \{n \geq 0 : X_n \in A\}.$$

Let $g(x) := \mathbb{E}_x[V_A]$. Obviously

$$(48) \quad g(0) = g(N) = 0.$$

95 / 178

Gambler's ruin, $p \neq 1/2$: How long does it last? (cont.)

From the one-step reasoning:

$$(49) \quad g(x) = 1 + p \cdot g(x+1) + q \cdot g(x-1), \quad x \in C.$$

These are $N-1$ equations for the $N-1$ unknowns: $(g(1), \dots, g(N-1))$. This system of equation is the same that appeared in formula (39). Thus, from Theorem 3.5 its solution can only be $g(x) = \mathbb{E}_x[V_A]$.

96 / 178

Gambler's ruin, $p \neq 1/2$: How long does it last? (cont.)

We can easily check that $g(1), \dots, g(N-1)$ is the solution of the system of equation (49) if

$$g(x) = \frac{x}{q-p} - \frac{N}{q-p} \cdot \frac{1-(q/p)^x}{1-(q/p)^N}, \quad 0 < x < N-1.$$

So, the expected time of the game for $0 < x < N-1$:

$$(50) \quad \mathbb{E}_x[V_A] = \frac{x}{q-p} - \frac{N}{q-p} \cdot \frac{1-(q/p)^x}{1-(q/p)^N}.$$

97 / 178

Gambler's ruin, $p \neq 1/2$: How long does it last? (cont.)

From now we always assume that $x \in C$. We use (50) and distinguish two cases: if $p < q$, then

$$(51) \quad \lim_{N \rightarrow \infty} \frac{N}{1-(q/p)^N} = 0 \text{ thus } g(x) \approx \frac{x}{q-p}.$$

On the other hand, if $p > q$, then $(q/p)^N \rightarrow 0$, thus

$$(52) \quad g(x) \approx \frac{N-x}{p-q} [1 - (q/p)^x] + \frac{x}{p-q} (q/p)^x.$$

98 / 178

This Subsection is based on Charles M. Grinstead, J. Laurie Snell's book. [2]. [Click here for the book.](#)

In this Section (unless we say otherwise) X_n is supposed to be an absorbing MC on a finite state space S with

- transition matrix P ,
- absorbing states $A \subset S$ and
- transient states $C := S \setminus A$.

We write $a := \#A$ and $c := \#C$.

We will answer the following questions in general terms:

99 / 178

Questions answered on this Subsection in general terms

- (Q1) What is the probability that the process will end up in a given absorbing state? (Theorem 3.11.)
- (Q2) What is expected exit time (expectation of the time to get to any of the absorbing states)? (Theorem 3.9.)
- (Q3) What is the expected number of visits to a transient state before finally getting to an absorbing state. (Theorem 3.8.)

100 / 178

We always assume that the $c + a$ states of S are arranged as follows: the first c states are the transient states and the last a states are the absorbing states. Then the transition matrix P is in the canonical form:

$$(53) \quad \begin{array}{c|cc} & C & A \\ \hline C & R & Q \\ A & \mathbf{0}_{a,c} & \mathbf{I}_a \end{array}, \text{ that is } P = \begin{pmatrix} R & Q \\ \mathbf{0}_{a,c} & \mathbf{I}_a \end{pmatrix}.$$

where

- R is a $c \times c$ matrix,
- Q is a non-zero $c \times a$ matrix
- $\mathbf{0}_{a,c}$ is an $a \times c$ zero matrix (all elements are zero),
- \mathbf{I}_a is an $a \times a$ identity matrix,

101 / 178

The powers of P

Clearly,

$$(54) \quad P^n = \begin{pmatrix} R^n & \star \\ \mathbf{0}_{a,c} & \mathbf{I}_a \end{pmatrix},$$

where \star is a $c \times a$ matrix. We have actually proved that

$$(55) \quad \lim_{n \rightarrow \infty} R^n = \mathbf{0}_{c,c}$$

The following Theorem answers question Q3. In special cases we have already seen its proof. Alternatively, for the proof see [2, p. 418, Theorem 11.4].

102 / 178

The fundamental matrix

Theorem 3.8

As always in this Subsection, we assume that X_n is an absorbing MC. Then

- (a) $\mathbf{I}_c - R$ has an inverse $\mathbf{N} := (\mathbf{I}_c - R)^{-1}$ which is called the fundamental matrix.
- (b) $\mathbf{N} = \mathbf{I}_c + R + R^2 + R^3 + \dots$.
- (c) $\mathbf{N} = (n_{ij})_{i,j=1}^c$ then n_{ij} is the expected values of the times the chain starting from $i \in C$ visits $j \in C$ before the absorption happens. Initial state is counted if $i = j$.

103 / 178

Time to absorption

Let X_n be as in Theorem 3.8. We write

$$V_A := \min \{n \geq 0 : X_n \in A\}.$$

We define the vector $\mathbf{g} = (g(x))_{x \in C}$, where

$$g(x) := \mathbb{E}_x[V_A], \quad \text{where } x \in C$$

That is the $x \in C$ -th component $g(x)$ of the vector \mathbf{g} is the expected number of steps until the absorption happens if the MC starts from x .

104 / 178

Time to absorption (cont.)

Theorem 3.9

Let X_n be an absorbing MC. We denote the column vector with all components equal to 1 by $\mathbf{1} \in \mathbb{R}^c$. Then

$$(56) \quad \mathbf{g} = \mathbf{N} \cdot \mathbf{1}.$$

We have actually proved this in the previous subsection in special cases. For a proof see [2, p. 420, Theorem 11.5]. This theorem answers [Question Q2](#). In the following slides we will answer [Question Q1](#).

105 / 178

An auxiliary lemma

We often need the following simple lemma.

Lemma 3.10

Let X be a non-negative integer valued r.v.. Then

$$(57) \quad \mathbb{E}[X] = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k).$$

Proof.

Observe that $X = \sum_{k=1}^{\infty} \mathbb{1}_{\{X \geq k\}}$. Then

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \mathbb{E}[\mathbb{1}_{\{X \geq k\}}] = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k). \quad \square$$

106 / 178

Absorption probabilities

Let $\mathbf{B} = (b_{i,j})_{i \in C, j \in A}$ be a $c \times a$ matrix whose elements are defined as follows: for an $i \in C$ and $j \in A$ we write

$$b_{i,j} := \mathbb{P}(\text{the chain starting from } i \text{ is absorbed at } j)$$

Theorem 3.11

Let X_n be an absorbing MC. Then

$$(58) \quad \mathbf{B} = \mathbf{N} \cdot \mathbf{Q}.$$

Now we present the proof in a shorter form than we repeat in a more detailed form.

107 / 178

Proof in short

Proof of Theorem 3.11 in short.

Let $\mathbf{R}^0 := \mathbf{I}$. Then

$$\begin{aligned} b_{i,j} &\stackrel{(57)}{=} \sum_{n=0}^{\infty} \sum_{k \in C} r_{i,k}^{(n)} \cdot q_{k,j} \\ &= \sum_{k \in C} \sum_{n=0}^{\infty} r_{i,k}^{(n)} \cdot q_{k,j} \\ &= \sum_{k \in C} n_{i,k} \cdot q_{k,j} \\ &= (\mathbf{N} \cdot \mathbf{R})_{i,j}. \quad \square \end{aligned}$$

108 / 178

Proof of Theorem 3.11 with details

Proof of Theorem 3.11 with details

Fix an arbitrary $i \in C$ and $j \in A$. Imagine that we start from i and finally arrive at j on such a path which stay within C before arriving at j . Let m be the length of this path. Observe that $m = 2$ means no states in between i and j on the path and for $m > 2$ there are $n - 2$ states in between i and j on the path and all of them must be in C . So such a path is describe with $c_1, \dots, c_{m-2} \in C$.

109 / 178

Proof of Theorem 3.11 with details (cont.)

Proof of Theorem 3.11 with details (cont.)

Let us call the probability that such a path is realized $w_{i,c_1,\dots,c_{m-2},j}$, where the word c_1, \dots, c_{m-2} is the empty word if $m = 2$. Below we write $n = m - 1$ from the two but last step:

110 / 178

Proof of Theorem 3.11 with details (cont.)

Proof (cont.)

$$\begin{aligned} b_{i,j} &= \sum_{m=2}^{\infty} \sum_{c_1, \dots, c_{m-2} \in C} w_{i,c_1, \dots, c_{m-2},j} \\ &= \sum_{m=2}^{\infty} \sum_{c_1, \dots, c_{m-2} \in C} p_{i,c_1} \cdot \prod_{k=1}^{m-2} p_{c_k, c_{k+1}} \cdot p_{c_{m-1}, j} \cdot p_{c_{m-1}, j} \\ &= \sum_{m=2}^{\infty} \sum_{c_1, \dots, c_{m-2} \in C} r_{i,c_1} \cdot \prod_{k=1}^{m-2} r_{c_k, c_{k+1}} \cdot q_{c_{m-1}, j} \\ &= \sum_{n=0}^{\infty} \sum_{k \in C} r_{i,k}^{(n)} \cdot q_{k,j} \\ &= \left(\sum_{n=0}^{\infty} \mathbf{R}^n \cdot \mathbf{Q} \right)_{i,j}, \end{aligned}$$

111 / 178

Proof of Theorem 3.11 with details (cont.)

Proof (cont.)

where $\prod_{k=1}^{m-2} r_{c_k, c_{k+1}} := 1$ if $m = 2$. Hence

$$(59) \quad \mathbf{B} = \sum_{n=0}^{\infty} \mathbf{R}^n \cdot \mathbf{Q}$$

Recall that according to part (b) of Theorem 3.8 we have $\mathbf{N} = \sum_{n=0}^{\infty} \mathbf{R}^n$. Hence, by (59) we obtained that

$$(60) \quad \mathbf{B} = \mathbf{N} \cdot \mathbf{Q}. \quad \blacksquare$$

112 / 178

The problems considered

In this subsection X_n is an irreducible chain on the finite state space S with transition matrix P and we assume that $\#S \geq 3$. Let $i, j, k \in S$ be three distinct elements of S . We pose the following questions:

- (Q4) What is the probability that the chain **starting from** $i \in S$ visits $j \in S$ earlier than $k \in S$?
- (Q5) What is the probability that the chain **starting from** $j \in S$ returns to j earlier than it visits $k \in S$.

113 / 178

The answer to question Q4

We prepare an absorbing MC from X_n by declaring some of the states absorbing. Namely, let \mathbf{e}_j and \mathbf{e}_k be the coordinate unit vectors in $\mathbb{R}^{\#S}$ which contains a 1 in their j and k -th position respectively, and all other components are zero. We replace of the j -th and k -th rows of P by \mathbf{e}_j and \mathbf{e}_k respectively. The transition probability matrix obtained in this way is denoted by $P^{(j,k)}$ and the corresponding MC is denoted by $X_n^{(j,k)}$.

114 / 178

The answer to question Q4 (cont.)

Clearly, $X_n^{(j,k)}$ is an absorbing MC with absorbing states $A := \{j, k\}$ transient states $C := S \setminus A$. Let

$$P^{(j,k)} = \begin{pmatrix} \mathbf{R}^{(j,k)} & \mathbf{Q}^{(j,k)} \\ \mathbf{0}_{a,c} & \mathbf{I}_a \end{pmatrix}$$

be the canonical form of $P^{(j,k)}$ and let $\mathbf{N}^{(j,k)}$ be the corresponding fundamental matrix:

$$\mathbf{N}^{(j,k)} = (\mathbf{I} - \mathbf{R}^{(j,k)})^{-1},$$

115 / 178

The answer to question Q4 (cont.)

where \mathbf{I} is the $(\#S - 2) \times (\#S - 2)$ identity matrix. Now we apply Theorem 3.11 for the MC $X_n^{(j,k)}$. That is we define the $(\#S - 2) \times 2$ matrix

$$(61) \quad \mathbf{B}^{(j,k)} = \mathbf{N}^{(j,k)} \cdot \mathbf{Q}^{(j,k)},$$

where the rows are indexed by the elements of C and the columns are indexed by $\{j, k\}$.

Now we can answer question Q4:

We introduce:

116 / 178

The answer to question Q4 (cont.)

$$(62) \quad \eta_{i,j,k} := \mathbb{P}(\text{the chain starting from } i \text{ visits } j \text{ earlier than } k)$$

Then by Theorem 3.11 and by the definition of matrix \mathbf{B} we obtain that

$$(63) \quad \eta_{i,j,k} = \mathbf{b}_{i,j}^{(j,k)},$$

117 / 178

The answer to question Q4 (cont.)

where $\mathbf{b}_{i,j}^{(j,k)}$ is the j -th element of the i -th row of the matrix $\mathbf{B}_{i,j}^{(j,k)}$ defined in (61) and this answers question Q4.

118 / 178

The answer to question Q5

Fix an arbitrary distinct $j, k \in S$. Let $\tau_{j,k}$ be the probability that the chain **starting from** $j \in S$ returns to j earlier than it visits $k \in S$. We can use the **one-step reasoning**. Namely, if the chain starting from j returns to j for the first time before visiting k then the chain starting from j cannot make its first step to k . So, in the first step the chain either remains in j (with probability $p(j,j)$) and then it has arrived back to j without visiting k) or it jumps to an $i \notin \{j, k\}$ and then it will continue starting now from $i \notin \{j, k\}$ and visits j earlier than k .

119 / 178

The answer to question Q5 (cont.)

The probability of this is (by definition) $\eta_{i,j,k}$. So, the one-step reasoning yields:

$$(64) \quad \tau_{j,k} = p_{j,j} + \sum_{i \notin \{j,k\}} p(j,i) \cdot \eta_{i,j,k}.$$

120 / 178

Absorbing Chains

All of these with Mathematica

Example 3.12 (Exercise 1.13 from Lawler's book [5])

Let X_n be a MC on $S = \{1, 2, 3, 4, 5\}$ with

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & \frac{4}{5} \\ 0 & 0 & 0 & \frac{2}{5} & \frac{3}{5} \\ 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

(a) Is this chain irreducible? Is it aperiodic? (b) Find π . (c) What is the expected number of steps to return to 1 for the first time if the chain starts from 1? (d) What is the expected number of steps to get to 4 for the first time, if the chain starts from 1? (e) What is the probability that the chain visits 5 earlier than 3 if the chain starts from 1? (f) What is the probability that the chain starting from 3 returns to 3 earlier than it visits 5?

121 / 178

Absorbing Chains

All of these with Mathematica

```

In[71]:=
Clear[P, p]

P = {
  {0, 1/2, 1/2, 0, 0},
  {0, 0, 0, 1/5, 4/5},
  {0, 0, 0, 2/5, 3/5},
  {1, 0, 0, 0, 0},
  {1/2, 0, 0, 0, 1/2}
}

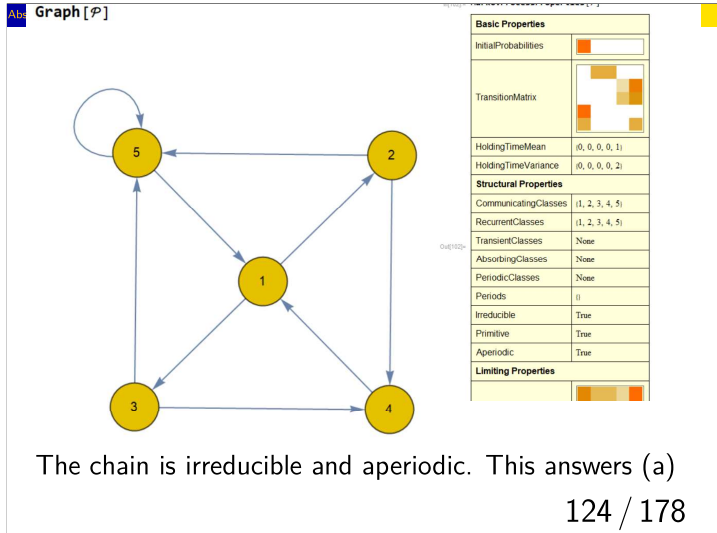
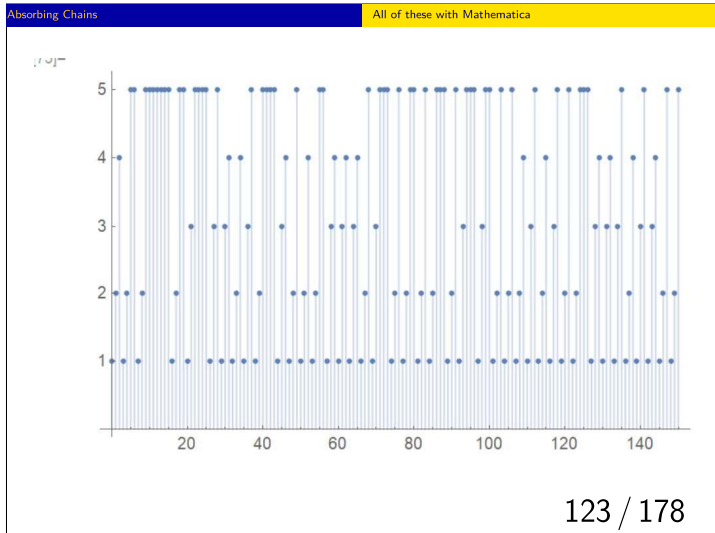
P = DiscreteMarkovProcess[1, p]

In[74]:= data = RandomFunction[P, {0, 150}]
Out[74]=
TemporalData[
  Time: 0 to 150
  Data points: 151
  Paths: 1
]

In[75]:= ListPlot[data, Filling -> Axis, Ticks -> {Automatic, {1, 2, 3, 4, 5}}]

```

122 / 178



Absorbing Chains

All of these with Mathematica

```

InvertMrep =
Inverse[ReplacePart[p - IdentityMatrix[Length[p[[1]]], {i_, Length[p[[1]]]} -> 1]]]

InvertMrep[[Length[p[[1]]]]]
{10/37, 5/37, 5/37, 3/37, 14/37}

f[i_, j_] := Mean[FirstPassageTimeDistribution[DiscreteMarkovProcess[i, p], j]]
Array[f, {5, 5}] // MatrixForm

MatrixForm=
{
  {37, 23, 24, 34, 23},
  {10, 5, 5, 3, 7},
  {14, 37, 38, 35, 13},
  {5, 5, 5, 3, 7},
  {13, 36, 37, 9, 19},
  {5, 5, 5, 3, 7},
  {1, 28, 29, 37, 38},
  {5, 5, 5, 3, 7},
  {2, 33, 34, 48, 37},
  {5, 5, 5, 3, 14}
}

```

So, (b): $\pi = (\frac{10}{37}, \frac{5}{37}, \frac{5}{37}, \frac{3}{37}, \frac{14}{37})$. (c): $\frac{37}{10}$. (d): $\frac{34}{3}$.

125 / 178

Absorbing Chains

All of these with Mathematica

```

In[92]:= MatrixForm=
{
  {1/2, 1/2, 0, 0, 0},
  {0, 0, 0, 1/5, 4/5},
  {0, 0, 0, 2/5, 3/5},
  {1, 0, 0, 0, 0},
  {1/2, 0, 0, 0, 1/2}
}

In[94]:= r = p[[{1, 2, 4}, {1, 2, 4}]]

In[95]:= q = p[[{1, 2, 4}, {3, 5}]]

In[96]:= id = IdentityMatrix[3]

In[97]:= n = Inverse[id - r]

In[98]:= b = n.q // MatrixForm

MatrixForm=
{
  {5, 4, 4},
  {9, 9, 9},
  {5, 4, 9}
}

```

(e) The answer is 4/9. (The element in the first row since we start from 1 and the column which corresponds to 5 (this is a the second column)).

$$B^{3,5} = \begin{pmatrix} 1 & 5 & 4 \\ 2 & 9 & 9 \\ 4 & 9 & 9 \end{pmatrix}$$

126 / 178

Absorbing Chains

All of these with Mathematica

That is

$$(65) \quad \eta_{1,3,5} = \frac{5}{9}, \quad \eta_{2,3,5} = \frac{1}{9}, \quad \eta_{4,3,5} = \frac{5}{9}.$$

Now we can answer question (f) that is we compute $\tau_{3,5}$ which was defined as the probability that the chain starting from 3 returns to 3 earlier than it visits 5. Namely, by (64) we have

$$\begin{aligned}
 \tau_{3,5} &= p_{3,3} + p(3,1)\eta_{1,3,5} + p(3,2)\eta_{2,3,5} + p(3,4)\eta_{4,3,5} \\
 &= 0 + 0 \cdot \frac{5}{9} + 0 \cdot \frac{1}{9} + \frac{2}{5} \cdot \frac{5}{9} \\
 &= \frac{2}{9}.
 \end{aligned}$$

127 / 178

Absorbing Chains

All of these with Mathematica

Umbrellas example [1, Exercice 1.37]

Example 3.13

An individual has three umbrellas, some at her office, and some at home. If she is leaving home in the morning (or leaving work at night) and it is raining, she will take an umbrella, if one is there. Otherwise, she gets wet. Assume that independent of the past, it rains on each trip with probability 0.2.

Question 1: Which percentage of time does she get wet?

128 / 178

Umbrellas example (cont.)

We approach this problem in the language of Markov chains. The only idea:

Let $S := \{0, 1, 2, 3\}$ and we write X_n for the number of umbrellas at the current location.

Then the transition matrix P is:

	0	1	2	3
0	0	0	0	1
1	0	0	0.8	0.2
2	0	0.8	0.2	0
3	0.8	0.2	0	0

129 / 178

Umbrellas example (cont.)

```
In[57]:= Clear[p, p]
```

$$\text{In[58]:= } p = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{8}{10} & \frac{2}{10} \\ 0 & \frac{8}{10} & \frac{2}{10} & 0 \\ \frac{8}{10} & \frac{2}{10} & 0 & 0 \end{pmatrix}$$

```
In[59]:= invmatrep =
```

```
Inverse[ReplacePart[p - IdentityMatrix[Length[p[[1]]],
{i_, Length[p[[1]]]} -> 1]]]
```

```
invmatrep[[Length[p[[1]]]]]
```

```
Out[60]=
```

$$\left\{ \frac{4}{19}, \frac{5}{19}, \frac{5}{19}, \frac{5}{19} \right\}$$

130 / 178

Umbrellas example (cont.)

This yields that the stationary distribution is

$$(66) \quad \pi = \left(\frac{4}{19}, \frac{5}{19}, \frac{5}{19}, \frac{5}{19} \right).$$

Hence it happens with probability $\frac{4}{19}$ that the individual does not have any umbrellas at her current location. However, she does not necessarily get wet at all of these occasions, since there is a rain only every 5th days (independently of everything). So, she gets wet with probability $\frac{4}{19 \cdot 5} = 0.04210526\dots$ Remark: the stationary distribution could be computed by hands easily since the system of equations is very simple.

131 / 178

Umbrellas example (cont.)

Namely, we want to find a probability vector

$\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$ such that

(67)

$$(\pi_0, \pi_1, \pi_2, \pi_3) \cdot \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{8}{10} & \frac{2}{10} \\ 0 & \frac{8}{10} & \frac{2}{10} & 0 \\ \frac{8}{10} & \frac{2}{10} & 0 & 0 \end{pmatrix} = (\pi_0, \pi_1, \pi_2, \pi_3)$$

This yields the system of equations:

132 / 178

Umbrellas: answer of Question 1

$$(68) \quad 0.8\pi(3) = \pi(0)$$

$$0.8\pi(2) + 0.2\pi(3) = \pi(1)$$

$$0.8\pi(1) + 0.2\pi(2) = \pi(2)$$

$$\pi(0) + \pi(1) + \pi(2) + \pi(3) = 1$$

As on slide 37 in File MC I, we throw away the last equation and substituted it by the condition that the sum of the components of π is equal to one, since the last equation of the original system would give no more information than the retained first three equations do. The solution of the system (68) is really obvious high school mathematics.

133 / 178

Doubly stochastic Markov Chains

1

Recurrence in case of countable infinite state space

2

Detailed balance condition and related topics

Detailed balance condition and Reversible Markov Chains

Birth and death processes

3

Absorbing Chains

Exit distributions through examples

Exit time through examples

Summary and the general theory

Application to Irreducible chains

All of these with Mathematica

4

Branching Processes

Generator functions

Branching Processes

5

Stopping time, Strong Markov property

6

Transient and recurrent states

134 / 178

Notation used in this Section

- In this Section we always that X is a such r.v. which takes only non-negative integers.
- $\forall k \in \mathbb{N}$ -re let $p_k := \mathbb{P}(X = k)$.
- The generator function of the r.v. X is

$$g_X(s) := \mathbb{E}[s^X] = \sum_{k=0}^{\infty} p_k \cdot s^k.$$

The most basic properties of generator functions (in short: g.f.)

135 / 178

Generator functions

- A generator function uniquely determines the cumulative distribution function.
- The generator function of the sum of two independent r.v. which take only non-negative integers, is the product of the generator functions of these r.v..

136 / 178

Generator functions (cont.)

- (c) Let $g(x)$ be the generator function of the r.v. X . Then

$$\mathbb{E}[X(X-1)\cdots(X-k)] = g^{(k+1)}(1),$$

where $g^{(k+1)}$ is the $k+1$ -th derivative of g . Hence by a simple calculation we get:

(69)

$$\mathbb{E}[X] = g'(1) \text{ és } \mathbb{E}[X^2] = g''(1) + g'(1).$$

137 / 178

Generator functions (cont.)

- (d) $g(1) = 1$ since (p_k) is a probability vector.

138 / 178

Generator functions (cont.)

Lemma 4.1

Let X and N be independent non-negative integer valued r.v. with generator functions g_X és g_N . Moreover, let X_1, X_2, \dots be i.i.d. r.v. having the same distribution as X . We define the r.v.:

$$R := X_1 + \cdots + X_N.$$

Then the generator function of R is:

$$(70) \quad g_R(s) = g_N(g_X(s)).$$

Before the proof of the Lemma we remark that an important corollary of Lemma 4.1 is as follows: Using

139 / 178

Branching Processes with more details (cont.)

$\{1, 2, \dots, X_n\}$ give birth to $Y_1^{(n)}, \dots, Y_{X_n}^{(n)}$ individuals. So, the number of level $n+1$ individuals is:

$$(72) \quad X_{n+1} = Y_1^{(n)} + \cdots + Y_{X_n}^{(n)}.$$

We always assume that $\{Y_m^{(n)}\}_{m,n}$ are i.i.d. r.v. with

$$Y_m^{(n)} \stackrel{d}{=} Y.$$

141 / 178

Branching Processes with more details (cont.)

That is

$$\mathbb{P}(Y_m^{(n)} = k) = p_k.$$

We can consider (X_n) as a Markov Chain with state space $S = \{0, 1, 2, \dots\}$ and the transition matrix $P = (p_{i,j})$ is given by

(73)

$$p(i,j) = P(Y_1 + \cdots + Y_i = j) \quad \text{for } i > 0 \text{ and } j \geq 0,$$

where $\{Y_i\}_{i=1}^\infty$ are i.i.d. with $Y_k \stackrel{d}{=} Y$.

142 / 178

Branching Processes with more details (cont.)

Let

$$g_n := \mathbb{E}[s^{X_n}],$$

That is g_n is the generator function of X_n , (which was defined as the number of level n individuals). Let

$$g(s) := g_1(s) := g_Y(s) = \sum_{n=0}^{\infty} p_n \cdot s^n.$$

143 / 178

Branching Processes with more details (cont.)

Clearly, for all m , the generator function of Y_m is the same:

$$g(s) = g_{Y_m}(s) \quad \forall m.$$

To get a better understanding of the generator function g_n we apply Lemma 4.1 with the following substitutions:

$$X_n \rightarrow N, Y_i \rightarrow X_i, X_{n+1} \rightarrow R.$$

144 / 178

Branching Processes with more details (cont.)

The we obtain from Lemma 4.1 that

$$g_{n+1} = g_n(g(s)).$$

From here, we obtain by mathematical induction that

$$(74) \quad g_n(s) = \underbrace{g \circ \dots \circ g}_n(s) =: g^n(s).$$

145 / 178

Branching Processes with more details (cont.)

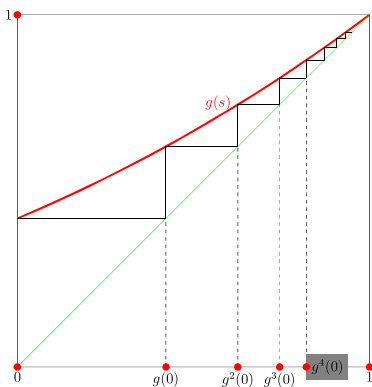
Apply this for $s = 0$ to get:

$$(75) \quad \mathbb{P}(X_n = 0) = g^n(0).$$

Hence $\mathbb{P}(\text{Extinction}) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0)$, where Extinction is the event the the Brancing Process dies out in finitely many steps.

146 / 178

$$\mathbb{E}[Y] = g'(1) < 1 \implies \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = 1$$



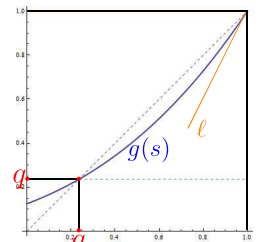
147 / 178

Summary: p_n is the probability that $0 \leq g'(q) < 1$. So, for an individual has exactly n offsprings. $g^n := \underbrace{g \circ \dots \circ g}_n$, we have $g^n(0) \rightarrow q$. That is by (75) q is the probability of extinction.

Then the expected number of offsprings of an individual is $m := \sum_{n=1}^{\infty} p_n \cdot n$. Consider the generator function:

$$g(s) := \sum_{n=0}^{\infty} p_n \cdot s^n.$$

The graph of g goes through $(1, 1)$. Let ℓ be the tangent line to g at $s = 1$. The slope of ℓ is $g'(1) = m$. If $m > 1$ then $\ell \cap [0, 1]^2$ is below the line $y = x$. Hence $\exists a \in [0, 1)$ with $g(a) = a$. Looking at the Figure:



148 / 178

- 1 Doubly stochastic Markov Chains
- 2 Recurrence in case of countable infinite state space
- 3 Detailed balance condition and related topics
 - Detailed balance condition and Reversible Markov Chains
 - Birth and death processes
- 4 Absorbing Chains
 - Exit distributions through examples
 - Exit time through examples
 - Summary and the general theory
 - Application to Irreducible chains
 - All of these with Mathematica
- 5 Branching Processes
 - Generator functions
 - Branching Processes
- 6 Stopping time, Strong Markov property
- 7 Transient and recurrent states

149 / 178

Notation

We study discrete time Markov chain X_n on the countable (finite or countably infinite) state space S with transition matrix $P = (p(i, j))_{i, j \in S}$.

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

\mathbb{E}_x notates the expected value for the probability \mathbb{P}_x . We frequently use the hitting time:

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

150 / 178

Notation (cont.)

The probability that the chain of starting at x will ever get to y :

$$\rho_{xy} := \mathbb{P}_x(T_y < \infty)$$

Intuitively: we feel that ρ_{yy}^2 is the probability of the event that {starting from y , we will come back to y twice} because we feel that whatever happens after we got back to y first is independent of what had happened before. To make this feeling precise we introduce the notion of **stopping time** or **Markov-time**.

151 / 178

Notation (cont.)

Definition 5.1 (Stopping time)

T is a **stoppingtime** if we can decide whether the event $\{T = n\}$ (we stop at time n) occur or does not occur by looking at the values X_0, \dots, X_n .

152 / 178

Stopping time

We can see easily that T_y is a stopping time, because

$$\{T_y = n\} = \{X_1 \neq y, \dots, X_{n-1} \neq y, X_n = y\}.$$

153 / 178

Stopping time (cont.)

Lemma 5.3

The

- sum
- maximum
- minimum

of two stopping times is stopping time.

155 / 178

Strong Markov property (cont.)

We only prove now that

$$(76) \quad \mathbb{P}(X_{T+1} = z | X_T = y, T = n) = p(y, z).$$

For an arbitrary $\mathbf{x} = (x_0, \dots, x_n)$, where $x_i \in S$, let $X_0^n(\mathbf{x})$ be and event defined by

$$X_0^n(\mathbf{x}) = \{X_0 = x_0, \dots, X_n = x_n\}.$$

We define

157 / 178

Strong Markov property (cont.)

$$\begin{aligned} \mathbb{P}(X_{T+1} = z, X_T = y, T = n) &= \\ &= \sum_{\mathbf{x} \in V_n} \mathbb{P}(X_{n+1} = z, X_0^n(\mathbf{x})) = \\ &= \sum_{\mathbf{x} \in V_n} \underbrace{\mathbb{P}(X_{n+1} = z | X_0^n(\mathbf{x}))}_{p(y, z)} \cdot \mathbb{P}(X_0^n(\mathbf{x})) = \\ &= p(y, z) \sum_{\mathbf{x} \in V_n} \mathbb{P}(X_0^n(\mathbf{x})) = \\ &= p(y, z) \cdot \mathbb{P}(T = n, X_T = y). \end{aligned}$$

We divide both sides by $\mathbb{P}(T = n, X_T = y)$ and this yields (76).

159 / 178

Stopping time (cont.)

Example 5.2

- $T \equiv k$ constant time is stopping time.
- The first time when X_n enters a given set A .
 $T(A) := \min \{n : X_n \in A\}$ is a stopping time.
- For a fixed k : the first time when the process enters into a given $A \subset S$ set for the k^{th} time is also a stopping time. (We will prove this later.)

Counter example: The last time when the process enters a given set is not a stopping time because we need to know the whole future to check it.

154 / 178

Strong Markov property

Theorem 5.4

Let X_n be Markov chain with transition matrix:

$\mathbf{P} = (p(i, j))$ and T be a stopping time. Assuming that $T = n$ and $X_T = y$, every further piece of information about X_0, \dots, X_T is irrelevant for the future (to estimate values of X_{T+k}) and for $k \geq 0$: X_{T+k} behaves like the original Markov chain started from y .

In the case of $T \equiv k$ we get back the Markov property.

156 / 178

Strong Markov property (cont.)

$$V_n := \{\mathbf{x} : X_0^n(\mathbf{x}) \implies (T = n \text{ and } X_T = y)\}.$$

In other words: V_n is the set of those $\mathbf{x} = (x_0, \dots, x_n)$, for which:

$$X_0 = x_0, \dots, X_n = x_n \implies T = n \text{ and } X_T = y$$

158 / 178

Strong Markov property (cont.)

- Doubly stochastic Markov Chains
- 1 Recurrence in case of countable infinite state space
- 2 Detailed balance condition and related topics
 - Detailed balance condition and Reversible Markov Chains
 - Birth and death processes
- 3 Absorbing Chains
 - Exit distributions through examples
 - Exit time through examples
 - Summary and the general theory
 - Application to Irreducible chains
 - All of these with Mathematica
- 4 Branching Processes
 - Generator functions
 - Branching Processes
- 5 Stopping time, Strong Markov property
- 6 Transient and recurrent states

160 / 178

Recurrent and transient states

Let $T_y^1 := T_y$ and

$$T_y^k := \min \{n > T_y^{k-1} : X_n = y\}$$

the time of the k^{th} return to y . Because of the strong Markov property

$$\mathbb{P}_y(T_y^k < \infty) = \rho_{yy}^k.$$

161 / 178

Recurrent and transient states (cont.)

- If $\rho_{yy} < 1$, then the probability of the event that the chain process comes back to y : $\rho_{yy}^k \rightarrow 0$. Thus, there's a time when the process no longer gets back to y . These y states are called **transient**.
- If $\rho_{yy} = 1$. Then for $\forall k$: $\rho_{yy}^k = 1$. Thus the process gets back to y infinitely many times. Then these y states are called **recurrent**.

162 / 178

Recurrent and transient states (cont.)

The following simple observation will be useful:

Lemma 6.1

If $\mathbb{P}_x(T_y \leq k) \geq \alpha > 0 \forall x \in S$, then

$$\mathbb{P}_x(T_y > nk) \leq (1 - \alpha)^n.$$

Namely, the probability that in the first n steps we have not visited y is less than $1 - \alpha$, the same is true for the subsequent $n - 1$ blocks of paths of length k .

163 / 178

Recurrent and transient states (cont.)

Definition 6.2

We say that x **communicates with** y ($x \rightsquigarrow y$) if the probability of reaching y from x in some (not necessarily in one) steps is positive. In other words:

$$x \rightsquigarrow y \text{ if } \rho_{xy} = \mathbb{P}_x(T_y < \infty) > 0.$$

It follows from Markov property that

$$(77) \quad \text{If } x \rightsquigarrow y \text{ and } y \rightsquigarrow z \text{ then } x \rightsquigarrow z.$$

164 / 178

Recurrent and transient states (cont.)

Lemma 6.3

If $\rho_{xy} > 0$ and $\rho_{yx} < 1$, then x is transient.

This is trivial, because since the event {starting from x we can get to y in finitely many steps} has positive probability and the event {from y we don't get back to x } also has positive probability. By Markov property: {starting from x we never get back to x } has also positive probability, so x is transient.

165 / 178

Recurrence and transience

Unless we say otherwise, we do not assume that $\#S < \infty$. Recall:

$$T_y^k = \min \{n > T_y^{k-1} : X_n = y\}$$

$$\text{and } \rho_{xy} = \mathbb{P}_x(T_y < \infty).$$

From the strong Markov property:

$$(78) \quad \mathbb{P}_x(T_y^k < \infty) = \rho_{xy} \cdot \rho_{yy}^{k-1}$$

166 / 178

Recurrence and transience (cont.)

Let

$$N(y) := \# \{n \geq 1 : X_n = y\}.$$

Obviously,

$$(79) \quad \{N(y) \geq k\} = \{T_y^k < \infty\}.$$

Hence, whenever $\rho_{yy} < 1$ (that is y is transient) we have

$$\begin{aligned} \mathbb{E}_x N(y) &= \sum_{k=1}^{\infty} \mathbb{P}_x(N(y) \geq k) = \sum_{k=1}^{\infty} \mathbb{P}_x\{T_y^k < \infty\} \\ &\stackrel{(78)}{=} \rho_{xy} \sum_{k=1}^{\infty} \rho_{yy}^{k-1} = \frac{\rho_{xy}}{1 - \rho_{yy}} \end{aligned}$$

167 / 178

Recurrence and transience (cont.)

So have obtained that

$$(80) \quad \rho_{yy} < 1 \implies \mathbb{E}_x N(y) = \frac{\rho_{xy}}{1 - \rho_{yy}}.$$

That's why $\mathbb{E}_y N(y) < \infty$ iff $\rho_{yy} < 1$. On the other hand we will prove that

Lemma 6.4

$$\mathbb{E}_x N(y) = \sum_{n=1}^{\infty} p^n(x, y).$$

168 / 178

Recurrence and transience (cont.)

Proof.

$N(y) = \sum_{n=1}^{\infty} \mathbb{1}_{X_n=y}$. Taking expected value:

$$\begin{aligned} \mathbb{E}_x N(y) &= \sum_{n=1}^{\infty} \mathbb{E}_x [\mathbb{1}_{X_n=y}] = \sum_{n=1}^{\infty} \underbrace{\mathbb{P}_x(X_n=y)}_{p^n(x,y)} \\ &= \sum_{n=1}^{\infty} p^n(x,y). \end{aligned}$$

□

169 / 178

Recurrence and transience (cont.)

As a corollary of Lemma 6.4 and (80) we get:

Theorem 6.5

An element $y \in S$ is recurrent if and only if:

$$\sum_{n=1}^{\infty} p^n(y,y) = \mathbb{E}_y[N(y)] = \infty.$$

170 / 178

Now we prove, using Theorem 6.5 that the **Simple Symmetric Random Walk (SSRW)** on \mathbb{Z} is recurrent. Recall that SSRW is defined on \mathbb{Z} by the transition probability matrix:

$$p(i, i+1) = p(i, i-1) = \frac{1}{2}, \quad \text{for all } i \in \mathbb{Z}.$$

Theorem 6.6

SSRW is **null-recurrent** on \mathbb{Z} . (The same is true on \mathbb{Z}^2 , but the SSRW is transient in \mathbb{Z}^d for $d \geq 3$.)

We use **Stirling-formula** in the proof:

171 / 178

$$(81) \quad 1 < \frac{n!}{\sqrt{2\pi n} \cdot (n/e)^n} < e^{1/(12n)}.$$

Hence we get

$$(82) \quad \binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}},$$

where \sim means that the ratio of the two sides tends to 1.

Proof

First we prove that SSRW is recurrent on \mathbb{Z} .

Remark: Starting from 0 we get to 0 in $2n$ steps iff we make n steps to the right and n steps to the left. The probability of each of these paths is $(1/2)^{2n}$ and the number of these paths is $\binom{2n}{n}$.

172 / 178

Proof (Cont.)

Hence,

$$\begin{aligned} p^{2n}(0,0) &= \binom{2n}{n} (1/2)^{2n} \\ &\sim \frac{1}{\sqrt{\pi n}}, \end{aligned}$$

where we used the formula given in (82). So,

$$\sum_{n=1}^{\infty} p^n(0,0) \geq \sum_{n=1}^{\infty} p^{2n}(0,0) = \text{const} \cdot \sum_{n=1}^{\infty} n^{-1/2} = \infty.$$

Now we use Theorem 6.5 to conclude that the simple symmetric random walk on \mathbb{Z} is recurrent.

173 / 178

Proof (Cont.)

Now we prove null-recurrence: Let E_k be the expected number of steps required to reach k starting from 0 for the first time. By definition, E_0 is not zero but the expected number of steps of the first return to 0. If we want to get into $k > 1$ from 0, first we have to reach 1, then 2, and so on; and the expected number of getting from i to $i+1$ is the same for all $i \in \mathbb{Z}$. Hence,

$$E_k = kE_1.$$

174 / 178

Proof (Cont.)

From the 1-step argument:

$$E_1 = 1 + \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot E_2,$$

because from -1 we can get into 1 in two steps. From this:

$$E_1 = 1 + E_1 \text{ so } E_1 = \infty.$$

Then by the 1-step argument we get

$$E_0 = 1 + \frac{1}{2}E_{-1} + \frac{1}{2}E_1,$$

So $E_0 = \infty$, thus the chain is null-recurrent.

175 / 178

- [1] R. DURRETT
Essentials of Stochastic Processes, Second edition
Springer, 2012. [Click here for the almost-ready version.](#)
- [2] CHARLES M. GRINSTEAD, J. LAURIE SNELL
Introduction to Probability
[Click here for the book.](#)
- [3] S. KARLIN, H.M. TAYLOR
A first course in stochastic processes
Academic Press, 2014

176 / 178

[4] D.A. LEVIN, Y. PERES, E.L. WILMER

Markov chains and mixing times

American Mathematical Society, 2009.

[Click here for the book.](#)

[5] G. LAWLER

Introduction to Stochastic Processes

Chapman & Hall 1995.

[6] E. SENATA

Non-negative Matrices and Markov Chains

Springer 1973.

Examples

Branching processes, 3

General random walk on $S = \mathbb{Z}^d$, 5

Modulo 6 jumps on a circle, 20–22

Simple Symmetric Random Walk (SSRW) on $S = \mathbb{Z}^d$, 5

Tennis, 68–70

Two stage Markov chains, 6, 7

Two steps back, one step forward chain, 16

Two year collage, 46–48

Umbrellas example, 128

Waiting time for TT , 79–83