Theory of Stochastic Processes
4. Markov chains

Tomonari Sei
sei@mist.i.u-tokyo.ac.jp

Department of Mathematical Informatics, University of Tokyo

April 27, 2017

http://www.stat.t.u-tokyo.ac.jp/~sei/lec.html

Handouts & Announcements

There are 2 handouts today.

- Slides (this one)
- A copy of Sections 6.1 to 6.6 of PRP.

Announcement:

- Hints for recommended problems were uploaded. Some of them are incomplete. Please let me know if you have a better answer!

Outline today

1. Review of last week's material
   - More about Abel's theorem
   - Branching processes
2. Markov chains
   - Examples
   - Irreducibility and aperiodicity
   - Stationary distributions and the limit theorem
3. Recommended problems

More about Abel's theorem

Abel's theorem we used last week is different from (but related to) the following one.

Abel's theorem (usual one)

Let $a_n \in \mathbb{R}$ and $G(s) = \sum_{n=0}^{\infty} a_n s^n$. Suppose that the radius of convergence is 1, and $\sum_{n=0}^{\infty} a_n$ converges. Then

$$\lim_{s \uparrow 1} G(s) = \sum_{n=0}^{\infty} a_n.$$

Example

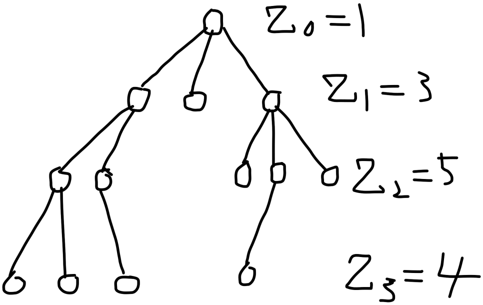
$$\log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

For a proof of the theorem, refer to

- K. A. Ross (2013) *Elementary Analysis*, Springer (available online).

What is the branching process?

The branching process is a tractable model for evolution of a population.



Branching processes

- Let Z_n be the number of members of the nth generation.
- Assumptions:
  (a) $Z_0 = 1$.
  (b) $Z_n = X_1^{(n)} + \cdots + X_{Z_{n-1}}^{(n)}$.
  (c) $X_i^{(n)}$ are independent and have the same probability mass function f and the generating function G.
- Z_n is called a branching process (or Galton-Watson process).

- How to obtain the generating function $G_n(s)$ of Z_n using G?

Generating function of branching processes

Theorem 5.4.1 of PRP

The generating function of Z_n satisfies

$$G_n(s) = G_{n-1}(G(s)) = (\underbrace{G \circ \cdots \circ G}_{n \text{ times}})(s).$$

Proof: Note that $G_0(s) = s$. For $n \geq 1$,

$$\begin{aligned}
G_n(s) &= E[s^{Z_n}] \\
&= E[s^{X_1^{(n)} + \cdots + X_{Z_{n-1}}^{(n)}}] \\
&= E[E[s^{X_1^{(n)} + \cdots + X_{Z_{n-1}}^{(n)}} | Z_{n-1}]] \quad \text{(tower property)} \\
&= E[\textstyle\prod_{i=1}^{Z_{n-1}^{(n)}} E[s^{X_i^{(n)}}]] \quad \text{(independence)} \\
&= E[G(s)^{Z_{n-1}^{(n)}}] \quad \text{(definition of } G) \\
&= G_{n-1}(G(s)).
\end{aligned}$$

Example

Let $f(k) = qp^k$ (geometric distribution). Then

$$G(s) = E[s^X] = \sum_k qp^k s^k = \frac{q}{1-ps},$$

$$G_2(s) = G(G(s)) = \frac{q}{1 - p\frac{q}{1-ps}} = \frac{q - pqs}{1 - pq - ps},$$

$$G_3(s) = G_2(G(s)) = \frac{q - pq\frac{q}{1-ps}}{1 - pq - p\frac{q}{1-ps}} = \frac{q - pq^2 - pqs}{1 - 2pq - p(1-pq)s},$$

$$\dots$$

In general, put $G_n(s) = \frac{a_n + b_n s}{c_n + d_n s}$. Then

$$G_n(s) = G_{n-1}(G(s)) = \frac{a_{n-1} + b_{n-1}\frac{q}{1-ps}}{c_{n-1} + d_{n-1}\frac{q}{1-ps}} = \frac{a_{n-1} + qb_{n-1} - pa_{n-1}s}{c_{n-1} + qd_{n-1} - pc_{n-1}s}.$$

Thus we have the following recurrence equation for a_n, b_n, c_n, d_n.

Example (cont.)

The recurrence equation in a matrix form is

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 1 & q \\ -p & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}, \quad \begin{pmatrix} c_n \\ d_n \end{pmatrix} = \begin{pmatrix} 1 & q \\ -p & 0 \end{pmatrix} \begin{pmatrix} c_{n-1} \\ d_{n-1} \end{pmatrix}.$$

With the initial condition $G_0(s) = s = \frac{0+1s}{1+0s}$, we have

$$\begin{pmatrix} a_n & c_n \\ b_n & d_n \end{pmatrix} = \begin{pmatrix} 1 & q \\ -p & 0 \end{pmatrix}^n \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Use the spectral decomposition to obtain

$$\begin{pmatrix} a_n & c_n \\ b_n & d_n \end{pmatrix} = \frac{1}{p-q} \begin{pmatrix} q(p^n - q^n) & p^{n+1} - q^{n+1} \\ -pq(p^{n-1} - q^{n-1}) & -p(p^n - q^n) \end{pmatrix}.$$

Compare with p.172 of PRP (a copy was provided last week).

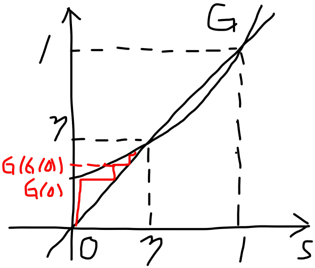
Extinction

Ultimate extinction is the event $\{Z_n = 0 \text{ for some } n\}$.

Theorem 5.4.5

As $n \to \infty$, $P(Z_n = 0) \to P(\text{ultimate extinction}) = \eta$, where η is the smallest non-negative root of $s = G(s)$.

A sketch of proof: $P(Z_n = 0) = G_n(0) = (\underbrace{G \circ \cdots \circ G}_{n \text{ times}})(0)$.



Phase transition

Let $\mu = G'(1) = E[Z_1]$ and $\eta = P(\text{ultimate extinction})$.

Corollary ("phase transition")

- $\mu > 1 \;\Rightarrow\; \eta < 1.$
- $\mu < 1 \;\Rightarrow\; \eta = 1.$

Exercise (5 min)

Let f be a trinomial distribution

$$f(k) = \begin{cases} p & \text{if } k = 2, \\ q & \text{if } k = 1, \\ r = 1 - p - q & \text{if } k = 0, \end{cases}$$

where $p, q, r > 0$. Find the condition that $P(\text{ultimate extinction}) < 1$.

Outline today

1. Review of last week's material
   - More about Abel's theorem
   - Branching processes
2. Markov chains
   - Examples
   - Irreducibility and aperiodicity
   - Stationary distributions and the limit theorem
3. Recommended problems

Markov chains

A process with the Markov property is called a Markov chain, that is,

Definition

A process $X = \{X_n\}_{n \geq 0}$ is called a Markov chain if

$$P(X_n = s \mid X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) = P(X_n = s \mid X_{n-1} = x_{n-1})$$

for all $n \geq 1$ and for all $s, x_1, \ldots, x_{n-1} \in S$. Here S denotes the state space.

Again, the Markov property means

the future is independent of the past conditional on the present.

Examples

Examples

Examples of Markov chains:

- random walks
- branching processes
- card shuffling

$$\spadesuit, \heartsuit, \diamondsuit, \clubsuit$$

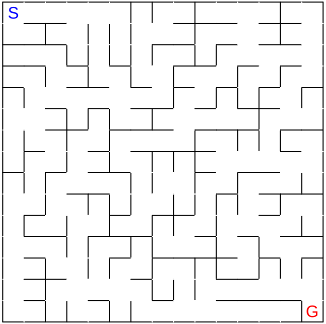
The size of the state space is $52! \approx 8.066 \times 10^{67}$.

Examples of not a Markov chain:

- Let $X_0, X_1, \ldots$ be a Bernoulli trial, and let $Y_n = X_n + X_{n+1}$ for $n \geq 0$. Then Y_n is not a Markov chain. $\rightarrow$ Why?
- A broad class of non-Markov chains is the hidden Markov model (HMM). This is quite important for application, but not discussed.

Applications

How to construct a maze randomly?



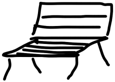
Markov chain can do it!
$\longrightarrow$ will be explained later.

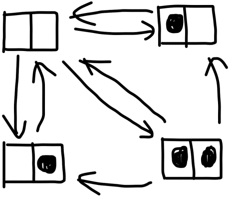
Propp and Wilson (1998). *J. Algorithm*, **27**, 170–217.

Transition diagram

If the state space is small or simple, the transition diagram is useful.

1. A bench:





2. A simple random walk:

$$\cdots \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \cdots$$

Notation

We briefly mention the limiting behavior of Markov chains without proofs. Let's begin!

- Let S be a countable subset, called the state space.
- We only consider homogeneous Markov chains, that is,

$$P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i) \quad \text{for all } n, i, j.$$

- Define the transition matrix $\mathbf{P} = (p_{ij})$ of a Markov chain X by

$$p_{ij} = P(X_1 = j \mid X_0 = i).$$

- Define the n-step transition matrix $\mathbf{P}_n = (p_{ij}(n))$ by

$$p_{ij}(n) = P(X_n = j \mid X_0 = i).$$

- A stationary distribution of $\mathbf{P}$ is a mass function $\boldsymbol{\pi} = (\pi_i)$ such that

$$\sum_{i \in S} \pi_i p_{ij} = \pi_j \quad \text{for all } j.$$

  In matrix notation, $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$.

Classification of states and chains (I)

- A Markov chain is said to be irreducible if, for any states i and j, there exists $n \geq 1$ such that $p_{ij}(n) > 0$.

- A state i is said to be aperiodic if there exists n_0 such that $p_{ii}(n) > 0$ for any $n \geq n_0$.
- If all the states are aperiodic, then the chain is called aperiodic.

Example

A simple random walk is irreducible but not aperiodic (i.e. periodic).

$$\cdots \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \cdots$$

For more details, see Sections 6.2 and 6.3.

Exercises

Blackboard

- Is card shuffling irreducible/aperiodic?
- Show that the example of mazes is irreducible and aperiodic.
  $\longrightarrow$ refer to the paper by Propp and Wilson mentioned before.

Classification of states and chains (II)

- Let $T_i = \inf\{n \geq 1 \mid X_n = i\}$ for a state i. Note that T_i may be ∞.
- The mean recurrence time μ_i of a state i is defined as

$$\mu_i = E[T_i \mid X_0 = i].$$

- A state i is called persistent if $P(T_i < \infty) = 1$.
- A persistent state i is called non-null if $\mu_i < \infty$; null otherwise.

- If all the states are persistent, then the chain is called persistent.

Example

A symmetric simple random walk is null persistent. (last week's material)

$$\cdots \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \bigcirc \leftrightarrows \cdots$$

For more details, see Sections 6.2 and 6.3.

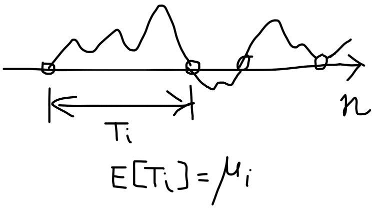
Stationary distributions

Now we state main theorems without proofs.

Theorem 6.4.3

An irreducible Markov chain has a stationary distribution $\boldsymbol{\pi}$ if and only if all the states are non-null persistent; in this case, $\pi_i = 1/\mu_i$.

The theorem is reasonable: since the chain comes the state i once in each period μ_i on average, the probability that the chain is at i will be $1/\mu_i$.



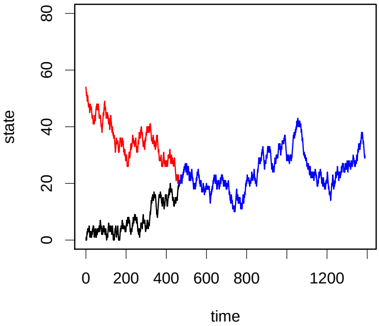
The limit theorem

Theorem 6.4.17

For an irreducible aperiodic non-null persistent Markov chain, we have that

$$p_{ij}(n) \to \frac{1}{\mu_j} \quad \text{as} \quad n \to \infty, \quad \text{for all } i \text{ and } j.$$

The proof is based on a coupling argument.



Finite state space

Lemma 6.4.5

Let S be finite. If a Markov chain is irreducible, then it is non-null persistent. In other words, $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ has a unique solution.

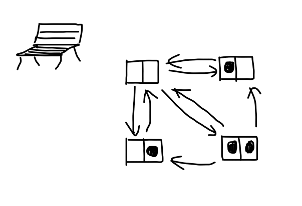
This is a corollary of the Perron-Frobenius theorem (Theorem 6.6.1), a proof of which is found in

- J. Liesen and V. Mehrmann (2015). *Linear Algebra*, Springer (available online).

Exercise

Find the stationary distribution of

$$\mathbf{P} = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1-2\delta & \delta & \delta & 0 \end{pmatrix}.$$



Recommended problems

Recommended problems:

- §6.1, Problem 2, 3, 8, 10.
- §6.4, Problem 4, 6.
- §6.15, Problem 1, 9*.

The asterisk (*) shows difficulty.

The next Thursday is "Green Day".
See you the week after next!