UNIVERSITY OF TORONTO

Statistical Sciences

DoSS Summer Bootcamp Probability
Module 9

Miaoshiqi (Shiki) Liu

University of Toronto

July 27, 2022

Recap

Learnt in last module:

- Convergence of functions of random variables
  - ▷ Slutsky's theorem
  - ▷ Continuous mapping theorem
- Laws of large numbers
  - ▷ WLLN
  - ▷ SLLN
  - ▷ Glivenko-Cantelli theorem
- Central limit theorem

Outline

- Markov Chain
  - Markov Property
- Discrete-time Markov Chain
  - Transition probability
  - Chapman-Kolmogorov equation
- Continuous-time Markov Chain
  - Transition probability
  - Chapman-Kolmogorov equation
  - Generator matrix

Markov chain

Recall:

A sequence of random variables $\{X_n\}_{i=1}^{n}$ are used to describe outcomes of random experiments.

Markov chain

Recall:
A sequence of random variables $\{X_n\}_{i=1}^{n}$ are used to describe outcomes of random experiments.

Remark:
What if the random variables follow some time structure (happen subsequently)?

Markov chain

Recall:
A sequence of random variables $\{X_n\}_{i=1}^{n}$ are used to describe outcomes of random experiments.

Remark:
What if the random variables follow some time structure (happen subsequently)?

Examples:

- Daily weather in Toronto
- Daily Covid-19 cases in Canada

Markov chain

Recall:
A sequence of random variables $\{X_n\}_{i=1}^{n}$ are used to describe outcomes of random experiments.

Remark:
What if the random variables follow some time structure (happen subsequently)?

Examples:

- Daily weather in Toronto
- Daily Covid-19 cases in Canada

Difficulties:

- The possible values of X_i's can vary a lot
- The random structure of X_i's can be complicated

Markov chain

Remark:
Consider a Markov chain to overcome the difficulties.

Markov chain

A Markov chain is specified by three ingredients:

- A state space $\mathcal{S}$, any non-empty finite or countable set.
- Initial probabilities $\{\nu_i\}_{i\in\mathcal{S}}$ where ν_i is the probability of starting at i (at time 0).
- Markov property:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = p_{ij}, \quad \forall i, j \in \mathcal{S},$$

and $\{p_{i,j}\}_{i,j\in\mathcal{S}}$ are transition probabilities.

Markov chain

Original: $\mathbb{P}(X_{n+1} = y \mid X_0 = x_0, \cdots, X_n = x_n)$

$\downarrow$

Countable state space $\mathcal{S}$: $\mathbb{P}(X_{n+1} = a_{i_{n+1}} \mid X_0 = a_{i_0}, \cdots, X_n = a_{i_n})$

$\downarrow$

Markov property: $\mathbb{P}(X_{n+1} = a_{i_{n+1}} \mid X_n = a_{i_n})$

Figure: Simplification by Markov chain

Markov chain

Remark:

The Markov chain we have introduced so far has discrete time index, and is called Discrete-time Markov Chain (DTMC). But there is also Continuous-time Markov chain (CTMC), and is sometimes referred to as "Markov Process".

	Countable state space	Continuous state space
Discrete time	DTMC	
Continuous time	CTMC	Continuous stochastic processes

Table: Types of "Series with Markov Property"

UNIVERSITY OF TORONTO

Discrete-time Markov chain

Representation of DTMC:

- Transition graph

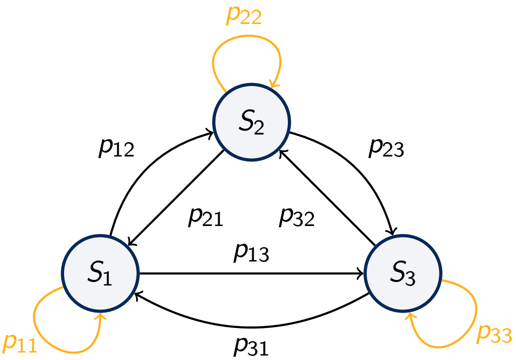

Figure: Example of the transition graph

Discrete-time Markov chain

Representation of DTMC:

- Transition matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Properties:

- $p_{ij} \geq 0, \quad i, j \in \mathcal{S}$
- $\sum_{j \in \mathcal{S}} p_{ij} = 1, \quad i \in \mathcal{S}$

Discrete-time Markov chain

Representation of DTMC:

- Transition matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Properties:

- $p_{ij} \geq 0, \quad i, j \in \mathcal{S}$
- $\sum_{j \in \mathcal{S}} p_{ij} = 1, \quad i \in \mathcal{S}$

Remark:

We don't have $\sum_{i \in \mathcal{S}} p_{ij} = 1, \quad j \in \mathcal{S}$.

Discrete-time Markov chain

Computation of joint probability:

$$\mathbb{P}(X_0 = i, X_1 = j) = \mathbb{P}(X_0 = i) \cdot \mathbb{P}(X_1 = j \mid X_0 = i) = \nu_i \cdot p_{ij}$$

$$\begin{aligned}
\mathbb{P}(X_0 = i, X_1 = j, X_2 = k) &= \mathbb{P}(X_0 = i, X_1 = j) \cdot \mathbb{P}(X_2 = k \mid X_0 = i, X_1 = j) \\
&= \mathbb{P}(X_0 = i, X_1 = j) \cdot \mathbb{P}(X_2 = k \mid X_1 = j) \qquad \text{(Markov Property)} \\
&= \nu_i \cdot p_{ij} \cdot p_{jk} \\
&\vdots
\end{aligned}$$

Discrete-time Markov chain

Computation of joint probability:

$$\mathbb{P}\left(X_0 = i, X_1 = j\right) = \mathbb{P}\left(X_0 = i\right) \cdot \mathbb{P}\left(X_1 = j \mid X_0 = i\right) = \nu_i \cdot p_{ij}$$

$$\begin{aligned}
\mathbb{P}\left(X_0 = i, X_1 = j, X_2 = k\right) &= \mathbb{P}\left(X_0 = i, X_1 = j\right) \cdot \mathbb{P}\left(X_2 = k \mid X_0 = i, X_1 = j\right) \\
&= \mathbb{P}\left(X_0 = i, X_1 = j\right) \cdot \mathbb{P}\left(X_2 = k \mid X_1 = j\right) \qquad \text{(Markov Property)} \\
&= \nu_i \cdot p_{ij} \cdot p_{jk} \\
&\vdots
\end{aligned}$$

Remark:

From the transition graph: the joint probability is just specifying the path we are taking.

Discrete-time Markov chain

Computation of transition probability after n transitions:

n-transition probability

$p_{ij}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i) = \mathbb{P}(X_{m+n} = j \mid X_m = i)$ is the probability that the state after n transitions is j if the original state is i. As a special case, $p_{ij}^{(1)} = p_{ij}$.

UNIVERSITY OF TORONTO

Discrete-time Markov chain

Computation of transition probability after n transitions:

n-transition probability

$p_{ij}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i) = \mathbb{P}(X_{m+n} = j \mid X_m = i)$ is the probability that the state after n transitions is j if the original state is i. As a special case, $p_{ij}^{(1)} = p_{ij}$.

$$
\begin{aligned}
p_{ij}^{(2)} = \mathbf{P}\left(X_2 = j \mid X_0 = i\right) &= \sum_{k \in S} \mathbf{P}\left(X_2 = j, X_1 = k \mid X_0 = i\right) \\
&= \sum_{k \in S} \mathbf{P}\left(X_2 = j \mid X_1 = k, X_0 = i\right) \cdot \mathbf{P}\left(X_1 = k \mid X_0 = i\right) \\
&= \sum_{k \in S} \mathbf{P}\left(X_2 = j \mid X_1 = k\right) \cdot \mathbf{P}\left(X_1 = k \mid X_0 = i\right) \\
&= \sum_{k \in S} p_{ik} p_{kj} = \left(P^2\right)[i, j]
\end{aligned}
$$

Discrete-time Markov chain

Remark:

In general, we have

$$p_{ij}^{(n)} = (P^n)[i,j].$$

Chapman-Kolmogorov equation / inequality

- $p_{ij}^{(m+n)} = \sum_{k\in\mathcal{S}} p_{ik}^{(m)} p_{kj}^{(n)}$ and $p_{ij}^{(m+s+n)} = \sum_{k\in\mathcal{S}} \sum_{l\in\mathcal{S}} p_{ik}^{(m)} p_{kl}^{(s)} p_{sj}^{(n)}$;
- $p_{ij}^{(m+n)} \geq p_{ik}^{(m)} p_{kj}^{(n)}$ and $p_{ij}^{(m+s+n)} \geq p_{ik}^{(m)} p_{kl}^{(s)} p_{sj}^{(n)}$ for any fixed state $k, l \in \mathcal{S}$.

Proof:

Discrete-time Markov chain

Example:

Consider a Markov chain with $\mathcal{S} = 1, 2, 3$, and $\nu = (\frac{1}{3}, \frac{2}{3}, 0)$, and

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}.$$

- Compute $\mathbb{P}(X_0 = 2)$;
- Compute $\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 2)$;
- Compute $p_{12}^{(3)}$.

Continuous-time Markov chain

Generalize the time index to be continuous:

Continuous-time Markov chain

A Continuous-time Markov chain $\{X(t)\}_{t\geq 0}$ is specified by three ingredients:

- A state space $\mathcal{S}$, any non-empty finite or countable set.
- Initial probabilities $\{\nu_i\}_{i\in\mathcal{S}}$ where ν_i is the probability of starting at $t=0$.
- Markov property: $\forall i,j \in \mathcal{S}$, $s,t\geq 0$,

$$\mathbb{P}(X(t+s)=j \mid X(s)=i, X(u)=x(u), 0\leq u\leq s) = \mathbb{P}(X(t+s)=j \mid X(s)=i).$$

Remark:

The process is called time-homogeneous when this probability does not depend on s.
Throughout the module, we will assume this time-homogeneity as a default.

Continuous-time Markov chain

Remark:
For time-homogeneous CTMC, we can define transition probability

$$p_{ij}^{(t)} = \mathbb{P}(X(s+t) = j \mid X(s) = i) = \mathbb{P}(X(t) = j \mid X(0) = i).$$

Continuous-time Markov chain

Remark:
For time-homogeneous CTMC, we can define transition probability

$$p_{ij}^{(t)} = \mathbb{P}(X(s+t) = j \mid X(s) = i) = \mathbb{P}(X(t) = j \mid X(0) = i).$$

Representation of CTMC:

- Transition graph after time t;
- Transition probability matrix:

$$P^{(t)} = \begin{bmatrix} p_{11}^{(t)} & p_{12}^{(t)} & \cdots \\ p_{21}^{(t)} & p_{22}^{(t)} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Continuous-time Markov chain

Properties:

- $p_{ij}^{(t)} \geq 0, \quad i, j \in \mathcal{S}$
- $\sum_{j \in \mathcal{S}} p_{ij}^{(t)} = 1, \quad i \in \mathcal{S}$
- $\mathbb{P}\left(X(0) = i_0, X\left(t_1\right) = i_1, \ldots, X\left(t_n\right) = i_n\right) = v_{i_0} p_{i_0 i_1}^{(t_1)} \ldots p_{i_{n-1} i_n}^{(t_n - t_{n-1})}$, for $0 < t_1 < \cdots < t_n$.

Continuous-time Markov chain

Properties:

- $p_{ij}^{(t)} \geq 0, \quad i, j \in \mathcal{S}$
- $\sum_{j \in \mathcal{S}} p_{ij}^{(t)} = 1, \quad i \in \mathcal{S}$
- $\mathbb{P}\left(X(0) = i_0, X\left(t_1\right) = i_1, \ldots, X\left(t_n\right) = i_n\right) = v_{i_0} p_{i_0 i_1}^{(t_1)} \ldots p_{i_{n-1} i_n}^{(t_n - t_{n-1})}$, for $0 < t_1 < \cdots < t_n$.

Chapman-Kolmogorov Equation

For a Continuous-time Markov chain $\{X_t\}_{t \geq 0}$ with transition probability matrix $P^{(t)}$,

$$P^{(s+t)} = P^{(s)} P^{(s)}.$$

Proof:

Continuous-time Markov chain

Generator and generator matrix

Given a Markov process, its generator is

$$g_{ij} = \lim_{t \to 0} \frac{p_{ij}^{(t)} - \delta_{ij}}{t},$$

where $\delta_{ij} = p_{ij}^{(0)} = 1$ if $i = j$, and 0 otherwise. The generator matrix is defined by

$$G = \lim_{t \to 0} \frac{P^{(t)} - I}{t}.$$

Properties:

- For t small. $P^{(t)} \approx I + tG$;
- Row sums of G is 0.

Continuous-time Markov chain

Continuous-time transition theorem

If a continuous-time markov chain has generator martix G, then for $t \geq 0$

$$P^{(t)} = \exp(tG) = I + tG + \frac{t^2 G^2}{2!} + \cdots$$

Proof:

UNIVERSITY OF TORONTO

Continuous-time Markov chain

Remark:

Suppose the eigendecomposition of G is $G = UDU^{-1}$, where D is a diagonal matrix with diagonal entries $\{d_1, d_2, \cdots\}$, then

$$P^{(t)} = U \exp(tD) U^{-1}.$$

Continuous-time Markov chain

Remark:

Suppose the eigendecomposition of G is $G = UDU^{-1}$, where D is a diagonal matrix with diagonal entries $\{d_1, d_2, \cdots\}$, then

$$P^{(t)} = U \exp(tD) U^{-1}.$$

Example:

Let

$$P^{(t)} = \begin{bmatrix} 1-3t & 3t \\ 5t & 1-5t \end{bmatrix}.$$

- Find G;
- Find the exact form of $P^{(t)}$.

UNIVERSITY OF TORONTO

Problem Set

Problem 1: (Bernoulli Process) Let $0 < p < 1$, repeatedly flp a coin with head probability p. Let X_n be the number of heads on the first n flips.

- Verify that $\{X_n\}$ is a Markov chain, specify the state space, initial probability and transition probability;
- Draw a sketch of the transition graph;
- For $p = \frac{1}{4}$, compute $\mathbb{P}(X_0 = 0, X_1 = 1, X_2 = 1, X_3 = 2)$.

Problem 2: Suppose a fair six-sided die is repeatedly rolled at times $0, 1, \cdots$ Let $X_0 = 0$, and for $n \geq 1$ let X_n be the largest value that appears among all of the rolls up to time n.

- Verify that $\{X_n\}$ is a Markov chain, specify the state space, initial probability and transition probability;
- Compute two-step transitions $\{p_{35}^{(2)}\}$.

Problem Set

Problem 3: Let $\{X(t)\}_{t\geq 0}$ be a continuous-time Markov chain on the state space $\mathcal{S} = \{1, 2, 3\}$, suppose that as $t \to 0$, the transition probabilities are given by

$$P^{(t)} = \begin{pmatrix} 1-7t & 7t & 0 \\ 0 & 1-3t & 3t \\ t & 2t & 1-3t \end{pmatrix} + o(t),$$

Compute the generator matrix G.