STATS 217: Introduction to Stochastic Processes I

Lecture 23

Continuous time Markov chains

- This week, we will study continuous time Markov chains.
- As before, we will assume that the state space is discrete. (often, it will be finite)
- We will also assume that the Markov chains under consideration are time-homogeneous i.e. that the transition rates do not depend on the time.

Continuous time Markov chains

Recall: stochastic process $(X_t)_{t \in T}$

We say that $(X_t)_{t\geq 0, t\in\mathbb{R}}$ is a **(time-homogeneous) continuous time Markov chain (CTMC)** on the state space Ω if

markov property

$$\mathbb{P}[X_{t+s} = j \mid X_s = i, X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_0} = i_0] = \mathbb{P}[X_{t+s} = j \mid X_s = i]$$

Continuous time Markov chains

We say that $(X_t)_{t\geq 0, t\in\mathbb{R}}$ is a **(time-homogeneous) continuous time Markov chain (CTMC)** on the state space Ω if

$$\begin{aligned}\mathbb{P}[X_{t+s} = j \mid X_s = i, X_{s_{n-1}} = i_{n-1}, \dots, X_{s_0} = i_0] &= \mathbb{P}[X_{t+s} = j \mid X_s = i] \\ &= \mathbb{P}[X_t = j \mid X_0 = i] \\ &=: p^t_{ij}. \quad (P^t)_{ij}\end{aligned}$$

time homogeneity

Continuous time Markov chains

We say that $(X_t)_{t\geq 0, t\in\mathbb{R}}$ is a **(time-homogeneous) continuous time Markov chain (CTMC)** on the state space Ω if

$$\begin{aligned}\mathbb{P}[X_{t+s} = j \mid X_s = i, X_{s_{n-1}} = i_{n-1}, \dots, X_{s_0} = i_0] &= \mathbb{P}[X_{t+s} = j \mid X_s = i] \\ &= \mathbb{P}[X_t = j \mid X_0 = i] \\ &=: p_{ij}^t.\end{aligned}$$

- for all integers $n \geq 0$,
- for all $0 \leq s_0 < s_1 < \dots s_{n-1} < s$,
- for all $0 \leq t$, and
- for all $j, i, i_0, \dots, i_{n-1} \in \Omega$.

Continuous time Markov chains

We say that $(X_t)_{t\geq 0, t\in\mathbb{R}}$ is a **(time-homogeneous) continuous time Markov chain (CTMC)** on the state space Ω if

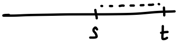
$$\begin{aligned}\mathbb{P}[X_{t+s} = j \mid X_s = i, X_{s_{n-1}} = i_{n-1}, \dots, X_{s_0} = i_0] &= \mathbb{P}[X_{t+s} = j \mid X_s = i] \\ &= \mathbb{P}[X_t = j \mid X_0 = i] \\ &=: p_{ij}^t.\end{aligned}$$

- for all integers $n \geq 0$,
- for all $0 \leq s_0 < s_1 < \dots s_{n-1} < s$,
- for all $0 \leq t$, and
- for all $j, i, i_0, \dots, i_{n-1} \in \Omega$.

As before, we will let P^t denote the $|\Omega| \times |\Omega|$ matrix with $P^t(i,j) = p_{ij}^t$.

Example

\# of points here $\sim \text{Poi}(\lambda(t-s))$



- Let $N(t)$ denote a PPP with rate λ.
- Then, $N(t)$ is a CTMC on the state space $\mathbb{Z}^{\geq 0}$ with transition probabilities

$$\mathbb{P}[X_t = j \mid X_0 = i] = \mathbb{P}[\text{Pois}(\lambda t) = (j - i)].$$

$$= e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}$$

Example

- Let $N(t)$ denote a PPP with rate λ.
- Then, $N(t)$ is a CTMC on the state space $\mathbb{Z}$ with transition probabilities

$$\mathbb{P}[X_t = j \mid X_0 = i] = \mathbb{P}[\text{Pois}(\lambda t) = (j - i)].$$

"continuisation of a DTMC"

- For a very general example, suppose that Y_n is a DTMC with state space Ω and with transition probabilities U_{ij}

$Y_0, Y_1, Y_2, Y_3, Y_4, \ldots$

- Then, $X_t = Y_{N(t)}$ is a CTMC on the state space Ω

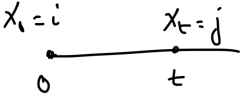
Example

- Let $N(t)$ denote a PPP with rate λ.
- Then, $N(t)$ is a CTMC on the state space $\mathbb{Z}$ with transition probabilities

$$\mathbb{P}[X_t = j \mid X_0 = i] = \mathbb{P}[\text{Pois}(\lambda t) = (j - i)].$$

- For a very general example, suppose that Y_n is a DTMC with state space Ω and with transition probabilities U_{ij}
- Then, $X_t = Y_{N(t)}$ is a CTMC on the state space Ω with transition probabilities

$$\mathbb{P}[X_t = j \mid X_0 = i] = \sum_{\ell \geq 0} \mathbb{P}[X_t = j \mid X_0 = i, N(t) - N(0) = \ell] \cdot \mathbb{P}[N(t) - N(0) = \ell]$$



e.g. exactly two transitions have happened $\iff$ exactly two PP in $[0, t]$

$(U^2)_{ij}$

Example

- Let $N(t)$ denote a PPP with rate λ.
- Then, $N(t)$ is a CTMC on the state space $\mathbb{Z}$ with transition probabilities

$$\mathbb{P}[X_t = j \mid X_0 = i] = \mathbb{P}[\text{Pois}(\lambda t) = (j - i)].$$

- For a very general example, suppose that Y_n is a DTMC with state space Ω and with transition probabilities U_{ij}
- Then, $X_t = Y_{N(t)}$ is a CTMC on the state space Ω with transition probabilities

$$\begin{aligned}\mathbb{P}[X_t = j \mid X_0 = i] &= \sum_{\ell \geq 0} \mathbb{P}[X_t = j \mid X_0 = i, N(t) - N(0) = \ell] \cdot \mathbb{P}[N(t) - N(0) = \ell] \\ &= \sum_{\ell \geq 0} (U^{\ell})_{ij} \left\{ e^{-\lambda t} \frac{(\lambda t)^{\ell}}{\ell!} \right\}.\end{aligned}$$

Heat kernel

- Given the transition matrix U on Ω, the **heat kernel** H_t is defined on $\Omega \times \Omega$ by

$$H_t(i,j) = \sum_{\ell \geq 0} (U^\ell)_{ij} \cdot e^{-\lambda t} \frac{(\lambda t)^\ell}{\ell!}.$$

$$\mathbb{P}[X_t = j \mid X_0 = i]$$

$$X_t = Y_{N(t)}$$

Y has transition matrix U.

Heat kernel

- Given the transition matrix U on Ω, the **heat kernel** H_t is defined on $\Omega \times \Omega$ by
$$H_t(i,j) = \sum_{\ell \geq 0} (U^\ell)_{ij} \cdot e^{-\lambda t} \frac{(\lambda t)^\ell}{\ell!}.$$
- The previous slide shows that H_t is the time t transition matrix of the CTMC $X_t = Y_{N(t)}$, where Y is a DTMC with transition matrix U and $N(t)$ is a PPP with rate λ.

Heat kernel

- Given the transition matrix U on Ω, the **heat kernel** H_t is defined on $\Omega \times \Omega$ by

$$H_t(i,j) = \sum_{\ell \geq 0} (U^{\ell})_{ij} \cdot \underbrace{e^{-\lambda t}} \frac{(\lambda t)^{\ell}}{\ell!}.$$

- The previous slide shows that H_t is the time t transition matrix of the CTMC $X_t = Y_{N(t)}$, where Y is a DTMC with transition matrix U and $N(t)$ is a PPP with rate λ.

$e^{x} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$

- A more compact way to write H_t is as

$$H_t = e^{\lambda t(U-I)} = \underline{\underline{e^{tQ}}},$$

where Q denotes the $\Omega \times \Omega$ matrix $Q = \lambda(\underline{U} - I)$.

$e^{t\lambda(Q-I)} = e^{-t\lambda I} \, e^{t\lambda U}$

$|\Omega| \times |\Omega|$ identity matrix

Jump rates

- For a DTMC, the transition matrix encodes the probability of transitioning from one state to another in the first step.

is there something like this for CTMC?

for DTMC:

$$\underline{P}_{ij} = \mathbb{P}[X_1 = j \mid X_0 = i]$$

Jump rates

- For a DTMC, the transition matrix encodes the probability of transitioning from one state to another in the first step.
- In continuous time, the ‘first step is infinitesimally small’.

Jump rates

- For a DTMC, the transition matrix encodes the probability of transitioning from one state to another in the first step.
- In continuous time, the 'first step is infinitesimally small'.
- Accordingly, we define the **jump rates** by

Recall:

$p^h_{ij} = \mathbb{P}[X_h = j \mid X_0 = i]$

$$q_{ij} := \lim_{h \to 0} \frac{p^h_{ij}}{h} \quad \forall i \neq j.$$

Jump rates

- For a DTMC, the transition matrix encodes the probability of transitioning from one state to another in the first step.
- In continuous time, the 'first step is infinitesimally small'.
- Accordingly, we define the **jump rates** by

$$q_{ij} := \lim_{h \to 0} \frac{p_{ij}^h}{h} \quad \forall i \neq j.$$

- We will set

$$q_{ii} = -\sum_{j \neq i} q_{ij} =: -\lambda_i.$$

because of this

$$\sum_j q_{ij} = 0$$

Jump rates

- For a DTMC, the transition matrix encodes the probability of transitioning from one state to another in the first step.
- In continuous time, the 'first step is infinitesimally small'.
- Accordingly, we define the **jump rates** by

$$q_{ij} := \lim_{h \to 0} \frac{p_{ij}^h}{h} \quad \forall i \neq j.$$

$$\lambda_i = \sum_{j \neq i} q_{ij}$$

- We will set

$$q_{ii} = -\sum_{j \neq i} q_{ij} =: -\lambda_i.$$

- In particular, for any $i \in \Omega$,

$$r_{ij} := \frac{q_{ij}}{\lambda_i} \quad j \neq i$$

is a probability distribution on $\Omega \setminus \{i\}$.

Example

$$q_{ij} = \lim_{h \to 0} \frac{(P^h)_{ij}}{h}$$

- For a PPP with rate λ, the only non-zero jump rates are $q_{i,i+1} = \lambda$ for all $i \geq 0$ and $q_{ii} = -\lambda$ for all $i \geq 0$.

$$(P^h)_{ij} = \mathbb{P}[X_h = j \mid X_0 = i]$$

$$= \mathbb{P}[\mathrm{Poi}(h) = (j-i)]$$

$$j = i+1 \qquad = \underbrace{e^{-\lambda h}}_{\to 1} \frac{(\lambda \cancel{h})^1}{1!} \frac{1}{\cancel{h}}$$

Example

- For a PPP with rate λ, the only non-zero jump rates are $q_{i,i+1} = \lambda$ for all $i \geq 0$ and $q_{ii} = -\lambda$ for all $i \geq 0$.
- For $X_t = Y_{N(t)}$ as before, where Y_n is a DTMC on the state space Ω with transition probabilities U_{ij}, the jump rates are (for $i \neq j$)

$(P^h)_{ij} = (e^{hQ})_{ij}$

$Q = \lambda(U - I)$

$$q_{ij} = \lim_{h \to 0} \frac{p^h_{ij}}{h}$$

$$= \lim_{h \to 0} \frac{(e^{hQ})_{ij}}{h}$$

Example

- For a PPP with rate λ, the only non-zero jump rates are $q_{i,i+1} = \lambda$ for all $i \geq 0$ and $q_{ii} = -\lambda$ for all $i \geq 0$.
- For $X_t = Y_{N(t)}$ as before, where Y_n is a DTMC on the state space Ω with transition probabilities U_{ij}, the jump rates are (for $i \neq j$)

$$
\begin{aligned}
q_{ij} &= \lim_{h \to 0} \frac{p_{ij}^h}{h} \\
&= \lim_{h \to 0} \frac{(e^{hQ})_{ij}}{h} \\
&= Q_{i,j} \\
&= \lambda (U_{ij} - I_{ij}) \\
&= \lambda \cdot U_{ij} \quad \text{since } i \neq j.
\end{aligned}
$$

Example

- For a PPP with rate λ, the only non-zero jump rates are $q_{i,i+1} = \lambda$ for all $i \geq 0$ and $q_{ii} = -\lambda$ for all $i \geq 0$.
- For $X_t = Y_{N(t)}$ as before, where Y_n is a DTMC on the state space Ω with transition probabilities U_{ij}, the jump rates are (for $i \neq j$)

$$\begin{aligned} q_{ij} &= \lim_{h \to 0} \frac{p_{ij}^h}{h} \\ &= \lim_{h \to 0} \frac{(e^{hQ})_{ij}}{h} \\ &= Q_{i,j} \\ &= \lambda \cdot (U - I)_{i,j} \\ &= \lambda \cdot U_{ij}. \end{aligned}$$

Example

- Typically, it is easier to describe a CTMC using jump rates and then compute the transition probabilities from the jump rates.

- intermediate q: if you have the jump rates, how do you simulate the process

Example

- Typically, it is easier to describe a CTMC using jump rates and then compute the transition probabilities from the jump rates.
- As an example, consider a continuous-time branching process where each individual independently dies at rate μ and gives birth to a new individual at rate λ.
- This corresponds to a CTMC with the jump rates

$$\left.\begin{aligned} q(n, n+1) &= \lambda n \\ q(n, n-1) &= \mu n. \end{aligned}\right\}$$

convenient + intuitive.
representation

Example

- Typically, it is easier to describe a CTMC using jump rates and then compute the transition probabilities from the jump rates.
- As an example, consider a continuous-time branching process where each individual independently dies at rate μ and gives birth to a new individual at rate λ.
- This corresponds to a CTMC with the jump rates

$$q(n, n+1) = \lambda n$$
$$q(n, n-1) = \mu n.$$

- We will now discuss how to construct a CTMC from the jump rates.

Constructing a CTMC from jump rates

How can we construct/simulate (say on a computer) a CTMC with given jump rates q_{ij}?

Constructing a CTMC from jump rates

How can we construct/simulate (say on a computer) a CTMC with given jump rates q_{ij}?

- Recall that for a DTMC (Y_n) on the state space Ω with transition probabilites U_{ij} and for $N(t)$ a PPP with rate λ, the jump rates of $X_t = Y_{N(t)}$ are

$$\longrightarrow \quad q_{ij} = U_{ij}\lambda, \quad i \neq j.$$

$$q_{ii} = -\sum_{j \neq i} q_{ij}$$

Constructing a CTMC from jump rates

How can we construct/simulate (say on a computer) a CTMC with given jump rates q_{ij}?

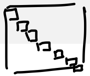
- Recall that for a DTMC (Y_n) on the state space Ω with transition probabilites U_{ij} and for $N(t)$ a PPP with rate λ, the jump rates of $X_t = Y_{N(t)}$ are

$$q_{ij} = U_{ij}\lambda, \quad i \neq j.$$

- We will basically reverse this process.

we have q_{ij}
we will try to find U_{ij} & λ
s.t. U_{ij} is a valid transition matrix &
$q_{ij} = U_{ij}\lambda$

Constructing a CTMC from jump rates



- First, suppose that $\Lambda := \sup_i \lambda_i < \infty$, where recall that $\lambda_i = \sum_{j\neq i} q_{ij}$.

e.g. $\left.\begin{array}{l} q(n, n+1) = \lambda n \\ q(n, n-1) = \mu n \end{array}\right\}$

does not satisfy $\Lambda < \infty$

Constructing a CTMC from jump rates

- First, suppose that $\Lambda := \sup_i \lambda_i < \infty$, where recall that $\lambda_i = \sum_{j \neq i} q_{ij}$.
- Let $N(t)$ be a PPP with rate Λ.

Constructing a CTMC from jump rates

- First, suppose that $\Lambda := \sup_i \lambda_i < \infty$, where recall that $\lambda_i = \sum_{j \neq i} q_{ij}$.
- Let $N(t)$ be a PPP with rate Λ.
- Let (Y_n) be a DTMC on the state space Ω with transition probabilities U_{ij} where

$$q_{ij} = U_{ij} \cdot \Lambda$$

$$U_{ij} = \frac{q_{ij}}{\Lambda} \qquad j \neq i$$

$$\sum_{i \neq j} U_{ij} = \frac{\sum_{i \neq j} q_{ij}}{\Lambda} = \frac{\lambda_i}{\Lambda} \leq 1$$

Constructing a CTMC from jump rates

- First, suppose that $\Lambda := \sup_i \lambda_i < \infty$, where recall that $\lambda_i = \sum_{j \neq i} q_{ij}$.
- Let $N(t)$ be a PPP with rate Λ.
- Let (Y_n) be a DTMC on the state space Ω with transition probabilities U_{ij} where

$$\sum_{j \neq i} U_{ij} = \frac{\lambda_i}{\Lambda} \rightarrow$$

$$U_{ij} = q_{ij}/\Lambda \quad i \neq j$$
$$U_{ii} = 1 - \lambda_i/\Lambda.$$

Constructing a CTMC from jump rates

- First, suppose that $\Lambda := \sup_i \lambda_i < \infty$, where recall that $\lambda_i = \sum_{j \neq i} q_{ij}$.
- Let $N(t)$ be a PPP with rate Λ.
- Let (Y_n) be a DTMC on the state space Ω with transition probabilities U_{ij} where

$$\begin{aligned} U_{ij} &= q_{ij}/\Lambda \quad i \neq j \\ U_{ii} &= 1 - \lambda_i/\Lambda. \end{aligned}$$

- Then, $X_t = Y_{N(t)}$ is a CTMC with jump rate from i to j for $i \neq j$ given by

$$\Lambda U_{ij} = q_{ij},$$

as desired.

Constructing a CTMC from jump rates

- What if $\Lambda = \infty$? For instance, this is the case for the branching process example we discussed earlier.

Constructing a CTMC from jump rates

- What if $\Lambda = \infty$? For instance, this is the case for the branching process example we discussed earlier.
- Given jump rates q_{ij}, let

$$\lambda_i = \sum_{j \neq i} q_{ij}$$

$$U_{ij} := r_{ij} = \frac{q_{ij}}{\lambda_i} \quad i \neq j.$$

Constructing a CTMC from jump rates

- What if $\Lambda = \infty$? For instance, this is the case for the branching process example we discussed earlier.
- Given jump rates q_{ij}, let

$$U_{ij} := r_{ij} = \frac{q_{ij}}{\lambda_i} \quad i \neq j.$$

- Recall that r_{ij} and hence U_{ij} is a probability distribution on $\Omega \setminus \{i\}$.

Constructing a CTMC from jump rates

- Let $Y_0, Y_1, \ldots$ denote a DTMC with

$$= \frac{q_{ij}}{\lambda_i}$$

$$\mathbb{P}[Y_{n+1} = j \mid Y_n = i] = U_{ij} \quad i \neq j$$
$$\mathbb{P}[Y_{n+1} = i \mid Y_n = i] = 0.$$

Constructing a CTMC from jump rates

- Let $Y_0, Y_1, \ldots$ denote a DTMC with

$$\mathbb{P}[Y_{n+1} = j \mid Y_n = i] = U_{ij} \quad i \neq j$$
$$\mathbb{P}[Y_{n+1} = i \mid Y_n = i] = 0.$$

- Given $Y_0 = i_0, Y_1 = i_1, Y_2 = i_2, \ldots,$ generate independent random variables $t_1, t_2, \ldots$ with

$$t_i \sim \text{Exp}(\lambda_{i-1}).$$

Constructing a CTMC from jump rates

- Let $Y_0, Y_1, \ldots$ denote a DTMC with

$$\mathbb{P}[Y_{n+1} = j \mid Y_n = i] = U_{ij} \quad i \neq j$$
$$\mathbb{P}[Y_{n+1} = i \mid Y_n = i] = 0.$$

- Given $Y_0 = i_0, Y_1 = i_1, Y_2 = i_2, \ldots,$ generate independent random variables $t_1, t_2, \ldots$ with

$$t_i \sim \text{Exp}(\lambda_{i-1}).$$

- Let $T_0 := 0$ and

$$T_n := t_1 + \cdots + t_n.$$

Constructing a CTMC from jump rates

- Let $Y_0, Y_1, \ldots$ denote a DTMC with

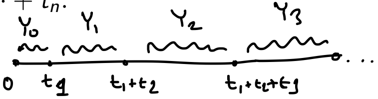
$$\mathbb{P}[Y_{n+1} = j \mid Y_n = i] = U_{ij} \quad i \neq j$$
$$\mathbb{P}[Y_{n+1} = i \mid Y_n = i] = 0.$$

- Given $Y_0 = i_0, Y_1 = i_1, Y_2 = i_2, \ldots,$ generate independent random variables $t_1, t_2, \ldots$ with

$$t_i \sim \text{Exp}(\lambda_{i-1}).$$

- Let $T_0 := 0$ and

$$T_n := t_1 + \cdots + t_n.$$



- Finally, let

$$X(t) = Y_n \quad \forall T_n \leq t < T_{n+1}.$$

Constructing a CTMC from jump rates

- Why does this work?

Constructing a CTMC from jump rates

- Why does this work? $\rightarrow$ wts $i \neq j$

$$q_{ij} = \lim_{h \to 0} \frac{p_{ij}^h}{h}$$

- For any $i \neq j$, we have

$$\lim_{h\to 0} \frac{p_{ij}^h}{h} = \lim_{h \to 0} \frac{1}{h} \mathbb{P}[\mathrm{Exp}(\lambda_i) \leq h] \cdot \mathbb{P}[Y_1 = j \mid Y_0 = i]$$

- Why does this work?
- For any $i \neq j$, we have

$$\lim_{h\to 0}\frac{p_{ij}^h}{h} = \lim_{h\to 0}\frac{1}{h}\mathbb{P}[\mathrm{Exp}(\lambda_i)\le h]\cdot\mathbb{P}[Y_1=j\mid Y_0=i]$$

$$= \lim_{h\to 0}\frac{1-e^{-\lambda_i h}}{h}\cdot\underbrace{\mathbb{P}[Y_1=j\mid Y_0=i]}_{=\,v_{ij}}$$

$$\approx \frac{\lambda_i \cancel{h}}{\cancel{h}}$$

$$= \lambda_i v_{ij} = \cancel{\lambda_i}\,\frac{q_{ij}}{\cancel{\lambda_i}}$$

Constructing a CTMC from jump rates

- Why does this work?
- For any $i \neq j$, we have

$$\begin{aligned}
\lim_{h \to 0} \frac{p_{ij}^h}{h} &= \lim_{h \to 0} \frac{1}{h} \mathbb{P}[\mathrm{Exp}(\lambda_i) \leq h] \cdot \mathbb{P}[Y_1 = j \mid Y_0 = i] \\
&= \lim_{h \to 0} \frac{1 - e^{-\lambda_i h}}{h} \cdot \mathbb{P}[Y_1 = j \mid Y_0 = i] \\
&= \lambda_i \cdot U_{ij}
\end{aligned}$$

Constructing a CTMC from jump rates

- Why does this work?
- For any $i \neq j$, we have

$$\begin{aligned}
\lim_{h\to 0} \frac{p_{ij}^h}{h} &= \lim_{h\to 0} \frac{1}{h}\mathbb{P}[\mathsf{Exp}(\lambda_i) \leq h] \cdot \mathbb{P}[Y_1 = j \mid Y_0 = i] \\
&= \lim_{h\to 0} \frac{1 - e^{-\lambda_i h}}{h} \cdot \mathbb{P}[Y_1 = j \mid Y_0 = i] \\
&= \lambda_i \cdot U_{ij} \\
&= \lambda_i \cdot \frac{q_{ij}}{\lambda_i} \\
&= q_{ij},
\end{aligned}$$

as desired.