STATS 217: Introduction to Stochastic Processes I

Lecture 24

Last time: jump rates

- Consider a CTMC on Ω with transition probabilities

$$p_{ij}^h = \mathbb{P}[X_{t+h} = j \mid X_t = i].$$

- The **jump rates** are defined by the matrix $Q = (q_{ij})_{i,j\in\Omega}$, where

$$q_{ij} := \lim_{h\to 0} \frac{p_{ij}^h}{h} \quad \forall i \neq j, \left\{ q_{ij} = \lim_{h\to 0} \frac{p_{ij}^h - p_{ij}^0}{h} \right)$$

and

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

- Last time, we saw how to simulate a CTMC with given jump rates.

Last time: Embedded DTMC

- In doing so, we found it useful to look at the **embedded DTMC**, which has transition matrix
$$U_{ij} = \frac{q_{ij}}{\lambda_i} \quad \forall i \neq j,$$
and
$$U_{ii} = 0.$$
- Given the embedded DTMC, we saw that the CTMC can be simulated by staying at each state with a suitable, exponentially distributed waiting time.
- Now, we will see how to recover the transition probabilities from the jump rates.

Example

- Recall the example of the continuization of a DTMC from last time. Namely, $(Y_n)_{n\geq 0}$ is a DTMC on Ω with transition matrix U, $N(t)$ is an independent PPP with rate λ, and $X_t = Y_{N(t)}$ is a CTMC.

Example

- Recall the example of the continuization of a DTMC from last time. Namely, $(Y_n)_{n\geq 0}$ is a DTMC on Ω with transition matrix U, $N(t)$ is an independent PPP with rate λ, and $X_t = Y_{N(t)}$ is a CTMC.
- We saw that the jump rates of X_t are given by

$$Q = \lambda(U - I),$$

where I is the $|\Omega| \times |\Omega|$ identity matrix.

direct computation

$Q_{ij} = \lambda U_{ij}$

$i \neq j$

Example

- Recall the example of the continuization of a DTMC from last time. Namely, $(Y_n)_{n\geq 0}$ is a DTMC on Ω with transition matrix U, $N(t)$ is an independent PPP with rate λ, and $X_t = Y_{N(t)}$ is a CTMC.
- We saw that the jump rates of X_t are given by

$$Q = \lambda(U - I),$$

  where I is the $|\Omega| \times |\Omega|$ identity matrix.
- Also, the transition probabilities of X_t are given by

$$(p^t)_{ij} = (e^{tQ})_{ij}.$$

Example

- Recall the example of the continuization of a DTMC from last time. Namely, $(Y_n)_{n\geq 0}$ is a DTMC on Ω with transition matrix U, $N(t)$ is an independent PPP with rate λ, and $X_t = Y_{N(t)}$ is a CTMC.
- We saw that the jump rates of X_t are given by

$$Q = \lambda(U - I),$$

  where I is the $|\Omega| \times |\Omega|$ identity matrix.
- Also, the transition probabilities of X_t are given by

$$(p^t)_{ij} = (e^{tQ})_{ij}.$$

$$P^t = e^{tQ}$$

- We will see that this holds more generally.

$$(*) \left\{ \frac{dP^t}{dt}\Big|_{t_0} = Q e^{t_0 Q} = Q P^{t_0} \right\}$$

Chapman-Kolmogorov equations

$(*)$ + initial condition

$P^0 = \mathrm{Id}$

unique soln $\Rightarrow$ $P^t = e^{tQ}$

- Recall that for a DTMC, we have the Chapman-Kolmogorov equations

$$p_{ij}^{n+m} = \sum_{k \in \Omega} p_{ik}^n p_{kj}^m.$$

$$P^{n+m} = P^n P^m \quad \text{(in matrix notation)}$$

Chapman-Kolmogorov equations

- Recall that for a DTMC, we have the Chapman-Kolmogorov equations

$$p_{ij}^{n+m} = \sum_{k \in \Omega} p_{ik}^{n} p_{kj}^{m}.$$

- The same argument also applies to a CTMC and shows that

$$p_{ij}^{s+t} = \sum_{k \in \Omega} p_{ik}^{s} p_{kj}^{t}.$$

we want to say something about

$\frac{d}{dt} P^t \big|_{t=0}$

Kolmogorov's forward equation $\left\{ \lim_{h \to 0} \frac{P^{t_0+h} - P^{t_0}}{h} \right\}$.

In particular,

chapman-kolmogorov.

$$p_{ij}^{t+h} - p_{ij}^{t} = \left(\sum_{k \in \Omega} p_{ik}^{t} p_{kj}^{h} \right) - p_{ij}^{t}$$

Kolmogorov's forward equation

In particular,

$$
\begin{aligned}
p_{ij}^{t+h} - p_{ij}^{t} &= \left(\sum_{k \in \Omega} p_{ik}^{t} p_{kj}^{h} \right) - p_{ij}^{t} \\
&= \left(p_{ij}^{t} p_{jj}^{h} + \sum_{j \neq k} p_{ik}^{t} p_{kj}^{h} \right) - p_{ij}^{t}
\end{aligned}
$$

Kolmogorov's forward equation

In particular,

$$
\begin{aligned}
p_{ij}^{t+h} - p_{ij}^{t} &= \left(\sum_{k \in \Omega} p_{ik}^{t} p_{kj}^{h} \right) - p_{ij}^{t} \\
&= \left(p_{ij}^{t} p_{jj}^{h} + \sum_{j \neq k} p_{ik}^{t} p_{kj}^{h} \right) - p_{ij}^{t} \\
&= p_{ij}^{t}(p_{jj}^{h} - 1) + \sum_{j \neq k} p_{ik}^{t} p_{kj}^{h}.
\end{aligned}
$$

Kolmogorov's forward equation

So,

$$\frac{d}{dt} p_{ij}^t = \lim_{h \to 0} \frac{p_{ij}^{t+h} - p_{ij}^t}{h}$$

Kolmogorov's forward equation

So,

$$
\begin{aligned}
\frac{d}{dt} p_{ij}^t &= \lim_{h \to 0} \frac{p_{ij}^{t+h} - p_{ij}^t}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (p_{jj}^h - 1) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h}
\end{aligned}
$$

$$
= p^t_{ij} \left[\lim_{h \to 0} \frac{(p^h_{jj} - 1)}{h} \right] + \sum_{j \neq k} p^t_{ik} \underbrace{\left[\lim_{h \to 0} \frac{p^h_{kj}}{h} \right]}_{= \, q_{kj}}
$$

$$
p^h_{jj} = 1 - \sum_{j \neq k} p^h_{jk}
$$

Kolmogorov's forward equation

So,

$$\begin{aligned}
\frac{d}{dt}p_{ij}^t &= \lim_{h\to 0}\frac{p_{ij}^{t+h}-p_{ij}^t}{h} \\
&= \lim_{h\to 0}\frac{p_{ij}^t(p_{jj}^h-1)+\sum_{j\neq k}p_{ik}^t p_{kj}^h}{h} \\
&= \lim_{h\to 0}\frac{p_{ij}^t(-\sum_{k\neq j}p_{jk}^h)+\sum_{j\neq k}p_{ik}^t p_{kj}^h}{h}
\end{aligned}$$

Kolmogorov's forward equation

So,

$$\begin{aligned}
\frac{d}{dt} p_{ij}^t &= \lim_{h \to 0} \frac{p_{ij}^{t+h} - p_{ij}^t}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (p_{jj}^h - 1) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (-\sum_{k \neq j} p_{jk}^h) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h} \\
&= p_{ij}^t \left(-\sum_{j \neq k} q_{jk} \right) + \sum_{j \neq k} p_{ik}^t q_{kj}
\end{aligned}$$

Kolmogorov's forward equation

So,

$$
\begin{aligned}
\frac{d}{dt} p_{ij}^t &= \lim_{h \to 0} \frac{p_{ij}^{t+h} - p_{ij}^t}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (p_{jj}^h - 1) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (- \sum_{k \neq j} p_{jk}^h) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h} \\
&= p_{ij}^t \left(- \sum_{j \neq k} q_{jk} \right) + \sum_{j \neq k} p_{ik}^t q_{kj} \\
&= p_{ij}^t q_{jj} + \sum_{j \neq k} p_{ik}^t q_{kj}
\end{aligned}
$$

Kolmogorov's forward equation

So,

$$
\begin{aligned}
\frac{d}{dt} p_{ij}^t &= \lim_{h \to 0} \frac{p_{ij}^{t+h} - p_{ij}^t}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (p_{jj}^h - 1) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h} \\
&= \lim_{h \to 0} \frac{p_{ij}^t (-\sum_{k \neq j} p_{jk}^h) + \sum_{j \neq k} p_{ik}^t p_{kj}^h}{h} \\
&= p_{ij}^t \left(-\sum_{j \neq k} q_{jk} \right) + \sum_{j \neq k} p_{ik}^t q_{kj} \\
&= p_{ij}^t q_{jj} + \sum_{j \neq k} p_{ik}^t q_{kj} \\
&= \sum_{k \in \Omega} p_{ik}^t q_{kj}. \quad = \quad (P^t Q)_{ij}
\end{aligned}
$$

Kolmogorov's backward equation

- Written in matrix form, we have **Kolmogorov's forward equation**

$$\left(\frac{d}{dt}P^t\right)(t_0) = P^{t_0}Q. \qquad \frac{d}{dt}P^t\Big|_{t_0} = QP^{t_0}$$

Kolmogorov's backward equation

- Written in matrix form, we have **Kolmogorov's forward equation**

$$\left(\frac{d}{dt}P^t\right)(t_0) = P^{t_0}Q.$$

earlier: $p_{ij}^{t+h} = \sum_{k \in \Omega} p_{ik}^{t} p_{kj}^{h}$

- Similarly, by writing

$$p_{ij}^{t+h} - p_{ij}^{t} = \left(\sum_{k\in\Omega} p_{ik}^{h} p_{kj}^{t}\right) - p_{ij}^{t},$$

and computing as before, we have **Kolmogorov's backward equation**

$$\left(\frac{d}{dt}P^t\right)(t_0) = QP^{t_0}$$

Computing transition probabilities from jump rates

- We have shown that

$$P^{t_0} Q = \left(\frac{d}{dt} P^t \right)(t_0) = Q P^{t_0}.$$

Computing transition probabilities from jump rates

- We have shown that

$$P^{t_0} Q = \left(\frac{d}{dt} P^t \right) (t_0) = Q P^{t_0}.$$

- The solution to this matrix ordinary differential equation with the initial condition $P^0 = \mathrm{Id}$ is given by

$$P^t = e^{tQ} := \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!}.$$

e.x. → Y_n is a DTMC w/ transition matrix U
→ $N(t)$ is PPP w/ rate λ
→ $X_t = Y_{N(t)}$: $Q = \lambda(U - \mathrm{Id})$

Irreducibility

embedded DTMC

$$\begin{cases} v_{ij} = q_{ij}/\lambda_i & j \neq i \\ v_{ii} = 0 \end{cases}$$

- We say that a CTMC $(X_t)_{t\geq 0}$ on Ω is **irreducible** if the embedded DTMC is irreducible.

- natural defn: $(X_t)_{t\geq 0}$ is irred if $\forall$ $i, j \in \Omega$, $\exists\, t > 0$ s.t.

$$p^t_{ij} > 0.$$

Irreducibility

- We say that a CTMC $(X_t)_{t \geq 0}$ on Ω is **irreducible** if the embedded DTMC is irreducible.
- By definition of the embedded DTMC, this amounts to the following: for any $i, j \in \Omega$, there exists a finite sequence of states

  $$k_0 = i, k_1, \ldots, k_{n-1}, k_n = j$$

  such that

  $$q_{k_{m-1} k_m} > 0 \quad \forall 1 \leq m \leq n.$$

Irreducibility

- We say that a CTMC $(X_t)_{t\geq 0}$ on Ω is **irreducible** if the embedded DTMC is irreducible.
- By definition of the embedded DTMC, this amounts to the following: for any $i, j \in \Omega$, there exists a finite sequence of states

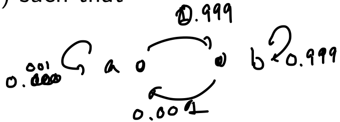
$$k_0 = i, k_1, \ldots, k_{n-1}, k_n = j$$

such that

$$q_{k_{m-1}k_m} > 0 \quad \forall 1 \leq m \leq n.$$

- Clearly, $(X_t)_{t\geq 0}$ is irreducible if and only if for any pair of states $i, j \in \Omega$, there exists some t (possibly depending on i, j) such that

$$p^t_{i,j} > 0.$$



Levy's dichotomy

- In fact, if P is irreducible, then for any pair of states $i, j \in \Omega$ and for **every** $t > 0$,
$$p_{i,j}^t > 0.$$

Levy's dichotomy

- In fact, if P is irreducible, then for any pair of states $i, j \in \Omega$ and for **every** $t > 0$,

$$p^t_{i,j} > 0.$$

- This is the consequence of **Levy's dichotomy**: for a CTMC and for any two states $i, j \in \Omega$, exactly one of the following holds:
  - $P^t_{i,j} > 0$ for all $t > 0$.
  - $P^t_{i,j} = 0$ for all $t > 0$.

$\left\{ \text{we saw that for an irr + aperiodic DTMC } \exists \text{ some } r_0 \geq 1 \text{ s.t. } p^{r_0}(i,j) > 0 \quad \forall i,j \right\}$

Levy's dichotomy

- In fact, if P is irreducible, then for any pair of states $i, j \in \Omega$ and for **every** $t > 0$,

$$p^t_{i,j} > 0.$$

- This is the consequence of **Levy's dichotomy**: for a CTMC and for any two states $i, j \in \Omega$, exactly one of the following holds:
  - $P^t_{i,j} > 0$ for all $t > 0$.
  - $P^t_{i,j} = 0$ for all t $\cancel{\neq 0}$. $\underline{> 0}$

i.e. if $\exists\, t_0$ s.t. $p^{t_0}_{ij} > 0$

$\Downarrow$

then $\forall\, t > 0$: $p^t_{ij} > 0$.

- In particular, for CTMC, we don't have to worry about (a)periodicity.

Levy's dichotomy

Here's the idea:

- If $P_{i,j}^{t_0} > 0$ for some $t_0 > 0$, then there must exist some $k \geq 0$ such that it is possible for the embedded chain to go from i to j in exactly k steps.

$$i \to s_1 \to s_2 \dots \to s_{k-1} \to j$$

suppose the embedded DTMC comes up as

$$i \quad s_1 \quad s_2 \quad \dots \quad s_{k-1} \quad j$$

CTMC waits at z for $\text{Exp}(\lambda_z)$ time

imagine $t = 0.01$. w.p. positive probability

exactly $\text{Exp}(\lambda_i) \dots \text{Exp}(\lambda_{k-1})$ have occured in $[0, 0.01]$.

Levy's dichotomy

Here's the idea:

- If $P^{t_0}_{i,j} > 0$ for some $t_0 > 0$, then there must exist some $k \geq 0$ such that it is possible for the embedded chain to go from i to j in exactly k steps.
- However, for any $t > 0$, there is a positive probability that there are exactly k transitions in the time interval $[0, t]$ (recall that each transition happens after an independent waiting time, which is exponentially distributed).

Stationary distributions

$\pi P = \pi$

$\pi Q \overset{\times}{=} \pi$

- Recall that for a DTMC with transition matrix P, we defined a stationary distribution to be a probability distribution π satisfying

$$\pi P = \pi.$$

- A consequence of this is that

$$\pi P^t = \pi \quad \forall t \geq 0.$$

Stationary distributions

- Recall that for a DTMC with transition matrix P, we defined a stationary distribution to be a probability distribution π satisfying

$$\pi P = \pi.$$

for DTMC.

- A consequence of this is that

$$\pi P^t = \pi \quad \forall t \geq 0.$$

- For a CTMC, we will take this second statement to be the definition of a stationary distribution.

Stationary distributions

for DTMC, we avoided this by noting that this is equiv. to $\pi P = \pi$

- However, the condition $\pi P^t = \pi$ for all t is typically hard to check in practice since it requires checking a condition for every $t \geq 0$.
- Therefore, it is typically more convenient to use the following characterization of the stationary distribution in terms of the jump rates:

Stationary distributions

- However, the condition $\pi P^t = \pi$ for all t is typically hard to check in practice since it requires checking a condition for every $t \geq 0$.
- Therefore, it is typically more convenient to use the following characterization of the stationary distribution in terms of the jump rates:

$$\pi Q = 0.$$

Stationary distributions

- However, the condition $\pi P^t = \pi$ for all t is typically hard to check in practice since it requires checking a condition for every $t \geq 0$.
- Therefore, it is typically more convenient to use the following characterization of the stationary distribution in terms of the jump rates:

$$\pi Q = 0.$$

- Why are these two definitions equivalent?

$$\pi P^t = \pi \quad \forall \quad t \geq 0$$

$$\pi Q = 0.$$

Stationary distributions

- However, the condition $\pi P^t = \pi$ for all t is typically hard to check in practice since it requires checking a condition for every $t \geq 0$.
- Therefore, it is typically more convenient to use the following characterization of the stationary distribution in terms of the jump rates:

$$\pi Q = 0.$$

- Why are these two definitions equivalent?
- If $\pi P^t = \pi$ for all $t \geq 0$, then

$$0 = \frac{d}{dt}\pi P^t|_{t=0} = \pi \frac{d}{dt} P^t|_{t=0}$$

$= \pi Q P^0$

$= \pi Q$

Kolmogorov eqns

$\frac{d}{dt} P^t = P^t Q$

$= Q P^t$

Stationary distributions

- However, the condition $\pi P^t = \pi$ for all t is typically hard to check in practice since it requires checking a condition for every $t \geq 0$.
- Therefore, it is typically more convenient to use the following characterization of the stationary distribution in terms of the jump rates:

$$\pi Q = 0.$$

- Why are these two definitions equivalent?
- If $\pi P^t = \pi$ for all $t \geq 0$, then

$$\begin{aligned} 0 &= \frac{d}{dt}\pi P^t|_{t=0} = \pi \frac{d}{dt} P^t|_{t=0} \\ &= \pi Q P^0 \end{aligned}$$

Stationary distributions

- However, the condition $\pi P^t = \pi$ for all t is typically hard to check in practice since it requires checking a condition for every $t \geq 0$.
- Therefore, it is typically more convenient to use the following characterization of the stationary distribution in terms of the jump rates:

$$\pi Q = 0.$$

- Why are these two definitions equivalent?
- If $\pi P^t = \pi$ for all $t \geq 0$, then

$$\begin{aligned} 0 &= \frac{d}{dt} \pi P^t |_{t=0} = \pi \frac{d}{dt} P^t |_{t=0} \\ &= \pi Q P^0 = \pi Q. \end{aligned}$$

Stationary distributions

- Conversely, suppose that $\pi Q = 0$. Then,

$$\begin{aligned}\frac{d}{dt}\pi P^t\big|_{t=t_0} &= \pi \frac{d}{dt} P^t\big|_{t=t_0} \\ &= \pi Q P^{t_0} \\ &= 0 P^{t_0} \\ &= 0\end{aligned}$$

$\Rightarrow$ πP^t is constant over t

$\Rightarrow$ $\pi P^t = \pi P^0 = \pi \, \mathrm{Id} = \pi.$

Stationary distributions

- Conversely, suppose that $\pi Q = 0$. Then,

$$\begin{aligned}\frac{d}{dt}\pi P^t|_{t=t_0} &= \pi \frac{d}{dt} P^t|_{t=t_0} \\ &= \pi Q P^{t_0} \\ &= 0.\end{aligned}$$

Stationary distributions

- Conversely, suppose that $\pi Q = 0$. Then,

$$\begin{aligned}\frac{d}{dt}\pi P^t|_{t=t_0} &= \pi \frac{d}{dt} P^t|_{t=t_0} \\ &= \pi Q P^{t_0} \\ &= 0.\end{aligned}$$

- Therefore, πP^t is constant for $t \geq 0$ so that

$$\pi P^t = \pi P^0 = \pi.$$

Detailed balance conditions

- The condition $\pi Q = 0$ may still be hard to verify in practice, and in many interesting examples, one finds a stationary distribution/verifies the stationarity condition using the **detailed balance condition**, which now takes the form

$$\boxed{\pi_i q_{ij} = \pi_j q_{ji} \quad \forall i, j}.$$

for DTMC:

$\pi_i P_{ij} = \pi_j P_{ji}$

Detailed balance conditions

- The condition $\pi Q = 0$ may still be hard to verify in practice, and in many interesting examples, one finds a stationary distribution/verifies the stationarity condition using the **detailed balance condition**, which now takes the form

$$\pi_i q_{ij} = \pi_j q_{ji} \quad \forall i, j.$$

wts : $\Rightarrow \pi Q = 0$

- This implies that π is a stationary distribution since

$$(\pi Q)_j = \sum_i \pi_i q_{ij}$$
$$= \sum_i \pi_j q_{ji}$$

$$= \pi_j \sum_i q_{ji}$$

$$= \pi_j \cdot 0 = 0. \quad \blacksquare$$

Detailed balance conditions

- The condition $\pi Q = 0$ may still be hard to verify in practice, and in many interesting examples, one finds a stationary distribution/verifies the stationarity condition using the **detailed balance condition**, which now takes the form

$$\pi_i q_{ij} = \pi_j q_{ji} \quad \forall i, j.$$

- This implies that π is a stationary distribution since

$$\begin{aligned}(\pi Q)_j &= \sum_i \pi_i q_{ij} \\ &= \sum_i \pi_j q_{ji} \\ &= \pi_j \sum_i q_{ji}\end{aligned}$$

Detailed balance conditions

- The condition $\pi Q = 0$ may still be hard to verify in practice, and in many interesting examples, one finds a stationary distribution/verifies the stationarity condition using the **detailed balance condition**, which now takes the form

$$\pi_i q_{ij} = \pi_j q_{ji} \quad \forall i, j.$$

- This implies that π is a stationary distribution since

$$\begin{aligned} (\pi Q)_j &= \sum_i \pi_i q_{ij} \\ &= \sum_i \pi_j q_{ji} \\ &= \pi_j \sum_i q_{ji} \\ &= 0. \end{aligned}$$