MS-C2111 Stochastic Processes

Lecture 9

Continuous-time Markov chains

Jukka Kohonen
Aalto University

Pre-Christmas problem

A taxi company has three cabs. Customers arrive at rate 2 per hour, and rides take on average 20 min. If all cabs are busy, a customer goes elsewhere. What is the (invariant) probability that all cabs are busy?

Contents

Contents

General Markov processes

Stochastic process = Random function $X : T \to S$ with state space S and time range $T \subset \mathbb{R}$, defined on some measurable space with probability measure $\mathbb{P}$

Markov process = Stochastic process which satisfies the Markov property

$$\mathbb{P}(X_t \in B \,|\, X_s = x,\, H_s) \;=\; \mathbb{P}(X_t \in B \,|\, X_s = x)$$

for all $x \in S$, $B \subset S$, all time instants $s < t$, and all events H_s which are determined by the past values $\{X_r : r \leq s\}$

Discrete-time Markov chain
Markov process with S = countable set, $T = \mathbb{Z}_+$

Continuous-time Markov chain
Markov process with S = countable set, $T = \mathbb{R}_+$

Time homogeneity: $\mathbb{P}(X_t \in B \,|\, X_s = x) = \mathbb{P}(X_{t-s} \in B \,|\, X_0 = x)$

All Markov processes are assumed time-homogeneous unless otherwise mentioned

Continuous-time Markov chain

Stochastic process $X = (X_t)_{t \in \mathbb{R}_+}$ with a countable state space S, which satisfies the Markov property

$$\mathbb{P}(X_{t+h} = y \,|\, X_t = x,\, H_t) \;=\; \mathbb{P}(X_{t+h} = y \,|\, X_t = x)$$

for all $x, y \in S$, all $t, h \geq 0$, and all events H_t which are determined by the past values $\{X_s : s \leq t\}$

Transition matrix of the chain = ?

$$\mathbb{P}(X_{t+h} = y \,|\, X_t = x) \;=\; \mathbb{P}(X_h = y \,|\, X_0 = x) \;=\; P_h(x, y)$$

where the P_h is the h-step transition matrix of the chain.

A continuous-time MC is characterized by infinitely many transition matrices P_h, $h \geq 0$

Distribution at time t

$$\mu_t(x) \ = \ \mathbb{P}(X_t = x)$$

Theorem

The distribution at time t is obtained from the initial distribution μ_0 and the t-step transition matrix P_t via $\mu_t = \mu_0 P_t$.

Proof.

$$\mathbb{P}(X_t = y) \ = \ \sum_{x \in S} \mathbb{P}(X_0 = x)\mathbb{P}(X_t = y \,|\, X_0 = x) \ = \ \sum_{x \in S} \mu_0(x) P_t(x, y)$$

□

Note: In discrete time: $P_t = P^t$, the t-th power of $P = P_1$

We have already seen one example of a continuous-time Markov chain with a countably infinite state space. Which one?

Example: Poisson process

$(N_t)_{t\in\mathbb{R}_+}$ = Poisson process with intensity $\lambda > 0$

$$\begin{aligned}
\mathbb{P}(N_{t+h} = \ell \mid N_t = k,\ H_t) &= \mathbb{P}(N_{t+h} - N_t = \ell - k \mid N_t = k,\ H_t) \\
&= \mathbb{P}(N_{t+h} - N_t = \ell - k) \\
&= \mathbb{P}(N_h = \ell - k) \\
&= P_h(k, \ell)
\end{aligned}$$

where

$$P_h(k,\ell) = \begin{cases} \left(e^{-\lambda h}\right) \frac{(\lambda h)^{\ell - k}}{(\ell - k)!}, & \text{if } 0 \le k \le \ell \\ 0, & \text{else} \end{cases}$$

$\Longrightarrow$ $(N_t)_{t\in\mathbb{R}_+}$ is a continuous-time Markov chain on the countably infinite state space $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$

Example: Satellite

A satellite launched in space has an Exp(λ)-distributed operational time T with mean $1/\lambda = 10$ years. The state of the satellite is

$$X_t = \begin{cases} 1, & \text{if the satellite is operational at time } t \\ 0, & \text{else} \end{cases}$$

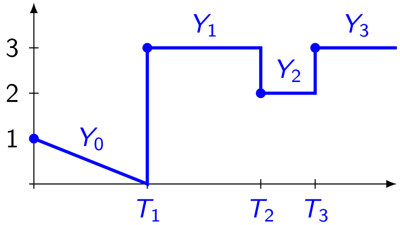
Then

$$\begin{aligned}
\mathbb{P}(X_{t+h} = 1 \,|\, X_t = 1, H_t) &= \mathbb{P}(T > t + h \,|\, T > t) = e^{-\lambda h} \\
\mathbb{P}(X_{t+h} = 0 \,|\, X_t = 1, H_t) &= 1 - e^{-\lambda h} \\
\mathbb{P}(X_{t+h} = 1 \,|\, X_t = 0, H_t) &= 0 \\
\mathbb{P}(X_{t+h} = 0 \,|\, X_t = 0, H_t) &= 1
\end{aligned}$$

$\implies$ $(X_t)_{t \in \mathbb{R}_+}$ is a continuous-time Markov chain on state space $\{0, 1\}$ with transition matrices

$$\begin{bmatrix} P_h(0,0) & P_h(0,1) \\ P_h(1,0) & P_h(1,1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 - e^{-\lambda h} & e^{-\lambda h} \end{bmatrix}, \qquad h \geq 0$$

Poisson modulated discrete-time chain



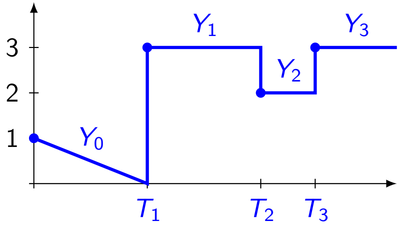
$(Y_n)_{n\in\mathbb{Z}_+}$ = discrete-time Markov chain with transition matrix P

$(N(t))_{t\in\mathbb{R}_+}$ = Poisson process with intensity λ

When $(Y_n)_{n\in\mathbb{Z}_+}$ and $(N(t))_{t\in\mathbb{R}_+}$ are independent, $X_t = Y_{N(t)}$ is a continuous-time Markov chain. WHY?

- Next state after $X_t = x$ is y with probability $P(x, y)$, and previously visited states are irrelevant in predicting this
- Information on how long the chain has been in state $X_t = x$ is irrelevant in predicting how long it will still stay in x, by the memoryless property of $\mathrm{Exp}(\lambda)$

Poisson modulated discrete-time chain



$(Y_n)_{n\in\mathbb{Z}_+}$ = discrete-time Markov chain with transition matrix P

$(N(t))_{t\in\mathbb{R}_+}$ = Poisson process with intensity λ

When $(Y_n)_{n\in\mathbb{Z}_+}$ and $(N(t))_{t\in\mathbb{R}_+}$ are independent, $X_t = Y_{N(t)}$ is a continuous-time Markov chain with transition matrices

$$\begin{aligned} P_t(x,y) &= \sum_{n=0}^{\infty} \mathbb{P}(N(t)=n)\,\mathbb{P}(Y_n = y \mid Y_0 = x) \\ &= \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!} P^n(x,y) \end{aligned}$$

Examples of Poisson modulated discrete-time chains

(i) $(N(t))_{t\in\mathbb{R}_+}$ is a Poisson process with intensity λ
$(Y_n)_{n\in\mathbb{Z}_+}$ is a DTMC on $\mathbb{Z}_+$ with $Y_0 = 0$ and transition matrix

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \\ \vdots & & \ddots & \ddots & \ddots \end{bmatrix}$$

Then $X_t := Y_{N(t)} = N(t) =$ Poisson process (Here $Y_n = n$)

(ii) $(N(t))_{t\in\mathbb{R}_+}$ is a Poisson process with intensity λ
$(Y_n)_{n\in\mathbb{Z}_+}$ is a DTMC on $\{0,1\}$ with $Y_0 = 1$ and transition matrix

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

Then $X_t := Y_{N(t)} =$ the CTMC of Example *Satellite*.

It can be shown that *every* finite-state CTMC can be represented as a Poisson modulated discrete-time chain.

Contents

Transition semigroup

Theorem

The transition matrices of a continuous-time Markov chain satisfy $P_{s+t} = P_s P_t$ *for all* $s, t \geq 0$.

Proof.

$$\begin{aligned} P_{s+t}(x,z) &= \mathbb{P}(X_{s+t} = z \mid X_0 = x) \\ &= \sum_{y \in S} \mathbb{P}(X_{s+t} = z,\, X_s = y \mid X_0 = x) \\ &= \sum_{y \in S} \mathbb{P}(X_s = y \mid X_0 = x)\, \mathbb{P}(X_{s+t} = z \mid X_s = y, X_0 = x) \\ &= \sum_{y \in S} P_s(x,y) P_t(y,z). \end{aligned}$$

The set $(P_t)_{t \geq 0}$ is called a transition semigroup. $\square$

In discrete time, the transition semigroup $(P_t)_{t \in \mathbb{Z}_+}$ is generated by the 1-step transition matrix $P = P_1$ via the formula $P_t = P^t$.

Is it possible to generate the transition semigroup $(P_t)_{t \in \mathbb{R}_+}$ of a continuous-time Markov chain using just one matrix?

Determining the transition semigroup

Is it sufficient to know the transition matrices P_h on a small time interval $h \in (0, \epsilon)$? YES

- For any t, choose a sufficiently large integer n so that $t/n < \epsilon$
- Then $P_t = P_{n \cdot (t/n)} = P_{t/n}^n$ by the semigroup property

All transition matrices are hence determined from $(P_h)_{h \in (0,\epsilon)}$

Sufficient to know $\lim_{h \to 0} P_h$? NO (Because $\lim_{h \to 0} P_h = I$)

However, for sufficiently regular chains there is a limit:

$$\lim_{h \to 0+} \frac{P_h - P_0}{h} = \lim_{h \to 0+} \frac{P_h - I}{h} = Q$$

Generator matrix

The generator matrix of a CTMC is defined by

$$Q(x,y) \;=\; \left[\frac{d}{dh}P_h(x,y)\right]_{h=0} = \lim_{h\to 0+} \frac{P_h(x,y) - I(x,y)}{h},$$

if the limit on right exists for all $x, y \in S$.

For $x \neq y$, $Q(x,y) = \lim_{h\to 0+} \frac{P_h(x,y)}{h}$ is the jump rate from x to y

The row sums of Q equal zero (because P_h, I have unit row sums)

Hence

$$Q(x,x) \;=\; -\sum_{y\neq x} Q(x,y)$$

Example: Poisson process

$$P_h(k,\ell) = \begin{cases} \left(e^{-\lambda h}\right) \frac{(\lambda h)^{\ell-k}}{(\ell-k)!}, & \text{if } 0 \le k \le \ell \\ 0, & \text{else} \end{cases}$$

$$\frac{P_h(k,k) - P_0(k,k)}{h} = \frac{e^{-\lambda h} - 1}{h} \to -\lambda$$

$$\frac{P_h(k,\ell) - P_0(k,\ell)}{h} = \frac{P_h(k,\ell)}{h} \to \begin{cases} \lambda, & \ell = k+1, \\ 0, & \ell \notin \{k, k+1\}. \end{cases}$$

Generator matrix

$$Q = \lim_{h \to 0+} \frac{P_h - I}{h} = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ \vdots & & & \ddots & \ddots & \ddots \end{bmatrix}$$

Example: Satellite

$$P_t = \begin{bmatrix} 1 & 0 \\ 1 - e^{-\lambda t} & e^{-\lambda t} \end{bmatrix}, \qquad t \geq 0.$$

$$\lim_{h \to 0} \frac{P_h(0,0) - I(0,0)}{h} = \lim_{h \to 0} \frac{1 - 1}{h} = 0$$

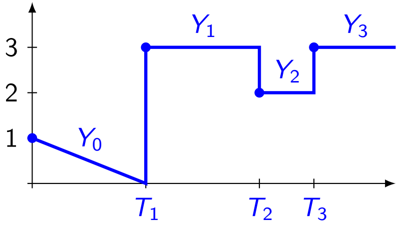
$$\lim_{h \to 0} \frac{P_h(0,1) - I(0,1)}{h} = \lim_{h \to 0} \frac{0 - 0}{h} = 0$$

$$\lim_{h \to 0} \frac{P_h(1,0) - I(1,0)}{h} = \lim_{h \to 0} \frac{1 - e^{-\lambda h}}{h} = \lambda$$

$$\lim_{h \to 0} \frac{P_h(1,1) - I(1,1)}{h} = \lim_{h \to 0} \frac{e^{-\lambda h} - 1}{h} = -\lambda$$

Generator matrix $\quad Q = \lim_{h \to 0+} \frac{P_h - I}{h} = \begin{bmatrix} 0 & 0 \\ \lambda & -\lambda \end{bmatrix}$

Poisson modulated discrete-time chain



$$P_t(x,y) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} P^n(x,y)$$

For $x \neq y$,

$$\begin{aligned}
\frac{d}{dt} P_t(x,y) &= \frac{d}{dt} \sum_{n=1}^{\infty} \left(e^{-\lambda t} \frac{(\lambda t)^n}{n!} \right) P^n(x,y) \\
&= \sum_{n=1}^{\infty} \left(-\lambda e^{-\lambda t} \frac{(\lambda t)^n}{n!} + e^{-\lambda t} \frac{\lambda (\lambda t)^{n-1}}{(n-1)!} \right) P^n(x,y)
\end{aligned}$$

$$\text{Jump rate} \quad Q(x,y) = \left[\frac{d}{dt} P_t(x,y) \right]_{t=0} = \lambda P(x,y)$$

Generator matrix generates the transition matrices

Theorem

For any transition semigroup $(P_t)_{t\geq 0}$ of a continuous-time Markov chain on a finite state space, the generator matrix $Q = \lim_{h\to 0+} \frac{P_h - I}{h}$ (i) exists, (ii) satisfies Kolmogorov's backward and forward differential equations

$$\frac{d}{dt}P_t \;=\; QP_t, \qquad \frac{d}{dt}P_t \;=\; P_tQ,$$

and (iii) determines the transition matrices of the chain via

$$P_t \;=\; \sum_{n=0}^{\infty} \frac{t^n Q^n}{n!}.$$

The above matrix equations are defined entry by entry.
The infinite sum of matrix powers equals the matrix exponential $\exp(tQ)$.

Contents

Invariant distribution

A probability distribution π is an invariant for a continuous-time Markov chain with transition semigroup $(P_t)_{t\in\mathbb{R}_+}$ if

$$\pi P_t \;=\; \pi \quad \text{for all } t \geq 0.$$

If $(X_t)_{t\in\mathbb{R}_+}$ is a continuous-time Markov chain with initial distribution π which invariant, then

$$\begin{aligned}
\mathbb{P}(X_t = y) &= \sum_x \mathbb{P}(X_0 = x)\,\mathbb{P}(X_t = y \,|\, X_0 = x) \\
&= \sum_x \pi(x) P_t(x,y) \\
&= \pi P_t(y) \;=\; \pi(y)
\end{aligned}$$

Hence the distribution of X_t remains invariant over time.

Theorem

π is an invariant distribution of a finite-state continuous-time Markov chain with generator matrix Q if and only if $\pi Q = 0$.

Proof.

(i) If π is invariant, then $\pi P_t = \pi$ for all t. Then $(\pi P_t)' = 0$.
By Kolmogorov's forward differential equation $P_t' = P_t Q$,

$$0 = (\pi P_t)' = \pi P_t' = \pi(P_t Q) = (\pi P_t)Q = \pi Q.$$

(ii) If $\pi Q = 0$, then

$$\pi P_t \;=\; \pi \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n \;=\; \pi + \sum_{n=1}^{\infty} \frac{t^n}{n!} (\pi Q) Q^{n-1} \;=\; \pi.$$

$\square$

Reversibility

A generator matrix Q is reversible with respect to a probability distribution π if it satisfies the detailed balance conditions

$$\pi(x)Q(x,y) \;=\; \pi(y)Q(y,x) \quad \text{for all } x \neq y.$$

As in discrete time, a Markov chain with a reversible initial distribution behaves statistically the same when observed backwards in time.

Theorem

If Q is π-reversible, then π is invariant.

Proof.

Because $\pi(x)Q(x,y) = \pi(y)Q(y,x)$ also for $x = y$, and because Q has zero row sums,

$$\pi Q(y) = \sum_x \pi(x)Q(x,y) = \sum_x \pi(y)Q(y,x) = \pi(y)\sum_x Q(y,x) = 0.$$

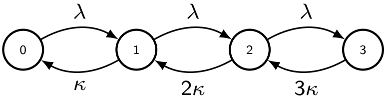
$\square$

Example: Taxi company

A taxi company has three cabs. Customers arrive with rate 2 per hour, and rides take on average 20 min. If all cabs are busy, a customer goes elsewhere. What is the (invariant) probability that all cabs are busy?

If customer interarrivals $=_{\mathrm{st}}$ Exp(λ) with rate $\lambda = 2$ and ride durations $=_{\mathrm{st}}$ Exp(κ) with rate $\kappa = 3$ (1/hour), all independent, then the number of busy cabs $(X_t)_{t \in \mathbb{R}_+}$ is a continuous-time Markov chain with

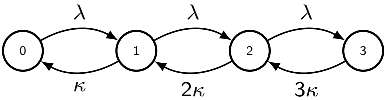
$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 \\ \kappa & -\lambda-\kappa & \lambda & 0 \\ 0 & 2\kappa & -\lambda-2\kappa & \lambda \\ 0 & 0 & 3\kappa & -3\kappa \end{bmatrix}$$



(We discuss more about this type of models during the next lecture.)

Example: Taxi company

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 \\ \kappa & -\lambda-\kappa & \lambda & 0 \\ 0 & 2\kappa & -\lambda-2\kappa & \lambda \\ 0 & 0 & 3\kappa & -3\kappa \end{bmatrix}$$



Detailed balance conditions: $\pi(k)Q(k,k+1) = \pi(k+1)Q(k+1,k)$

$$\begin{aligned} \lambda\pi(0) &= \kappa\pi(1), \\ \lambda\pi(1) &= 2\kappa\pi(2), \\ \lambda\pi(2) &= 3\kappa\pi(3) \end{aligned}$$

$$\pi(k) = \pi(0)\frac{(\lambda/\kappa)^k}{k!} = \dots = \frac{\frac{(\lambda/\kappa)^k}{k!}}{\sum_{j=0}^{3}\frac{(\lambda/\kappa)^j}{j!}}$$

$$\mathbb{P}(\text{all cabs busy}) = \pi(3) = \frac{\frac{1}{6}(2/3)^3}{1+2/3+\frac{1}{2}(2/3)^2+\frac{1}{6}(2/3)^3} \approx 0.0255$$

Kirjallisuusviitteet

Aineistolähteet

Esityksessä käytetyt kuvat (esiintymisjärjestyksessä)

1. Image courtesy of think4photop at FreeDigitalPhotos.net
2. Image courtesy of Hockeybroad/Cheryl Adams at Wikipedia.