Markovské řetězce se spojitým parametrem

Přednášející:
Mgr. Rudolf B. Blažek, Ph.D.

Katedra aplikované matematiky, Fakulta informačních technologií
České vysoké učení technické v Praze

Statistika pro informatiku

MI-SPI, LS 2015/16, Přednáška 14

Continuous-time Markov Chains

Lecturer:
Mgr. Rudolf B. Blažek, Ph.D.

Department of Applied Mathematics, Faculty of Information Technology
Czech Technical University in Prague

Statistics for Informatics

MIE-SPI, LS 2015/16, Lecture 14

Continuous-time Markov Chains

Recall Discrete-time Markov Chains

The Markov Property for a discrete-time Markov chains

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) =$$

$$= P(X_{n+1} = j \mid X_n = i) = \boldsymbol{P}_{i,j} = p_{i,j} = p(i,j)$$

In continuous time

- it is difficult to define conditional probability given X_r for all $r < s$
- we instead work with all choices of $0 \leq s_0 < s_1, \dots, s_n < s$ (for all $n \geq 1$)

Continuous-time Markov Chains

Definition *Markovský řetezec se spojitým parametrem*

X_t, $t \geq 0$ is a **continuous-time Markov chain** if for any times $0 \leq s_0 < s_1, \ldots, s_n < s$ and any states $i_0, i_1, \ldots, i_n, i, j$ (for all $n \geq 1$) we have

$$P(X_{t+s} = j \mid X_s = i, X_{s_n} = i_n, \ldots, X_{s_0} = i_0) =$$
$$= P(X_{t+s} = j \mid X_s = i) = P(X_t = j \mid X_0 = i) = p_t(i,j)$$

Construction: Discrete-time Markov Chain with Poisson Process Timing

Example

Let $N(t)$, $t \geq 0$ be a Poisson Process with rate λ.
Let Y_n be a discrete-time Markov chain with transition probabilities $u(i,j)$. Assume $N(t)$ is independent of Y_n.

Then $X_t = Y_{N(t)}$ is a continuous-time Markov chain.

X_t jumps according to $u(i,j)$ at each arrival of $N(t)$.

Most continuous-time Markov chains can be constructed is a similar way. We will show how later today.

Construction: Discrete-time Markov Chain with Poisson Process Timing

Example

$N(t)$, $t \geq 0$ is a Poisson Process with rate λ, therefore the number of arrivals in $(0,t)$ is a Poisson random variable

$$N(t) \sim \text{Poisson}(\lambda t)$$

$$P(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

Therefore

$$p_t(i,j) = P(X_t = j \mid X_0 = i)$$

Make n=0, or 1, or 2,... steps on the way from i to j in t units of time

$$= \sum_{n=0}^{\infty} P(N(t) = n, X_t = j \mid X_0 = i)$$

Construction: Discrete-time Markov Chain with Poisson Process Timing

Example

$$p_t(i,j) = \sum_{n=0}^{\infty} P(N(t) = n, X_t = j \mid X_0 = i)$$

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

$$= \sum_{n=0}^{\infty} P(N(t) = n,\ Y_{N(t)} = j \mid Y_{N(0)} = i)$$

$N(t)$ is independent of the MC Y_n

$$= \sum_{n=0}^{\infty} P(N(t) = n)\, P(Y_n = j \mid Y_0 = i)$$

Construction: Discrete-time Markov Chain with Poisson Process Timing

Example

$$p_t(i,j) = \sum_{n=0}^{\infty} P(N(t)=n)\, P(Y_n = j \mid Y_0 = i)$$

$$= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n\,!}\, u^n(i,j)$$

$N(t) \sim \text{Poisson}(\lambda t)$

n-step transition probability for discrete-time Markov chain Y_n

Jump Rates

Chapman-Kolmogorov Equation

Theorem

$$\sum_k p_s(i,k)\, p_t(k,j) = p_{s+t}(i,j)$$

Idea of the Proof

- We want to go from i to j in time $s + t$
- First we must go to some state k in time s
- Then we must finish by going from k to j in time t
- We must consider all states k, thus we sum over k

Chapman-Kolmogorov Equation

Theorem

$$\sum_k p_s(i,k)\, p_t(k,j) = p_{s+t}(i,j)$$

Assume we know $p_t(i,j)$ for all $t \leq t_0$ (and for all i,j)

- The theorem allows us to calculate $p_t(i,j)$ for all $t > 0$!!!

Let $t_0 \to 0$, consider the derivative at 0 for $j \neq i$:

$$\lim_{t_0 \to 0^+} \left(p_{t_0}(i,j) - 0\right) \big/ t_0$$

- The derivative at 0 should be be enough to determine $p_t(i,j)$ for all $t > 0$. Well, it is !! (We will show this later today.)

Jump Rates

Definition *Intenzita přechodu*

Denote the derivative of $p_t(i,j)$ with respect to time at 0 as

$$q(i,j) = \left.\frac{dp_t(i,j)}{dt}\right|_{t=0} = \lim_{h \to 0^+} \frac{p_h(i,j)}{h} \qquad \text{for } j \neq i$$

If the derivative exists, then we call $q(i,j)$ the **jump rate** from i to j.

Note that we do not calculate $q(i,i)$

Why is it called a jump rate?

Let's look at our construction of the Markov chain...

Why is $q(i,j)$ called a jump rate?

Previous Example

$X_t = Y_{N(t)}$; Y_n = discrete-time MC with transition prob. $u(i,j)$.
$N(t)$, $t \geq 0$ is a Poisson Process with rate λ, indep. of Y_n.

If X_t is at i, it makes jumps with rate λ (a Poisson process)
It goes to j with probability $u(i,j)$

This is a Poisson process thinning:
<u>X_t jumps from i to j as a Poisson Process with rate $\lambda\, u(i,j)$</u>

Next we will show that in this example $q(i,j) = \lambda\, u(i,j)$
That is why we call $q(i,j)$ the <u>jump rate</u> from i to j

Why is $q(i,j)$ called a jump rate?

Previous Example

$X_t = Y_{N(t)}$; Y_n = discrete-time MC with transition prob. $u(i,j)$.
$N(t)$, $t \geq 0$ is a Poisson Process with rate λ, indep. of Y_n.
We got before

$$p_t(i,j) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} u^n(i,j)$$

Therefore

$$\frac{p_h(i,j)}{h} = \frac{1}{h} \sum_{n=0}^{\infty} e^{-\lambda h} \frac{(\lambda h)^n}{n!} u^n(i,j)$$

Why is $q(i,j)$ called a jump rate?

Previous Example

$$\frac{p_h(i,j)}{h} = \frac{1}{h}\sum_{n=0}^{\infty} e^{-\lambda h}\frac{(\lambda h)^n}{n\,!}\, u^n(i,j)$$

Let $h \to 0$

$$\frac{p_h(i,j)}{h} = \frac{1}{h} e^{-\lambda h}\frac{(\lambda h)^0}{0\,!}\, \underbrace{u^0(i,j)}_{=0\ \forall j \neq i\ \text{(transition in 0 steps)}} + \boxed{\underbrace{e^{-\lambda h}}_{\to 1}\frac{(\lambda)^1}{1\,!}\, u^1(i,j)} \to \lambda\, u(i,j)$$

$$+ \underbrace{e^{-\lambda h}}_{\to 1}\sum_{n=2}^{\infty}\underbrace{\frac{\lambda^n h^{n-1}}{n\,!}}_{\to 0}\, u^n(i,j) \longrightarrow 0$$

Why is $q(i,j)$ called a jump rate?

Summary of Previous Example

$X_t = Y_{N(t)}$; Y_n = discrete-time MC with transition prob. $u(i,j)$.

$N(t)$, $t \geq 0$ is a Poisson Process with rate λ, indep. of Y_n.

X_t jumps from i to j as a Poisson Process w/ rate $\lambda\, u(i,j)$

The jump rates of the process X_t are

$$q(i,j) = \lim_{h \to 0} \frac{p_h(i,j)}{h} = \lim_{h \to 0} \lambda e^{-\lambda h} u(i,j) = \lambda u(i,j)$$

<u>Most Markov chains can be constructed in a similar way!!</u>

That is why for <u>any</u> continuous-time Markov chain we call $q(i,j)$ the <u>jump rates</u> from i to j.

Poisson Process as a Markov Chain

Simple Example — *Poisson Process*

Consider a Poisson process with rate λ.

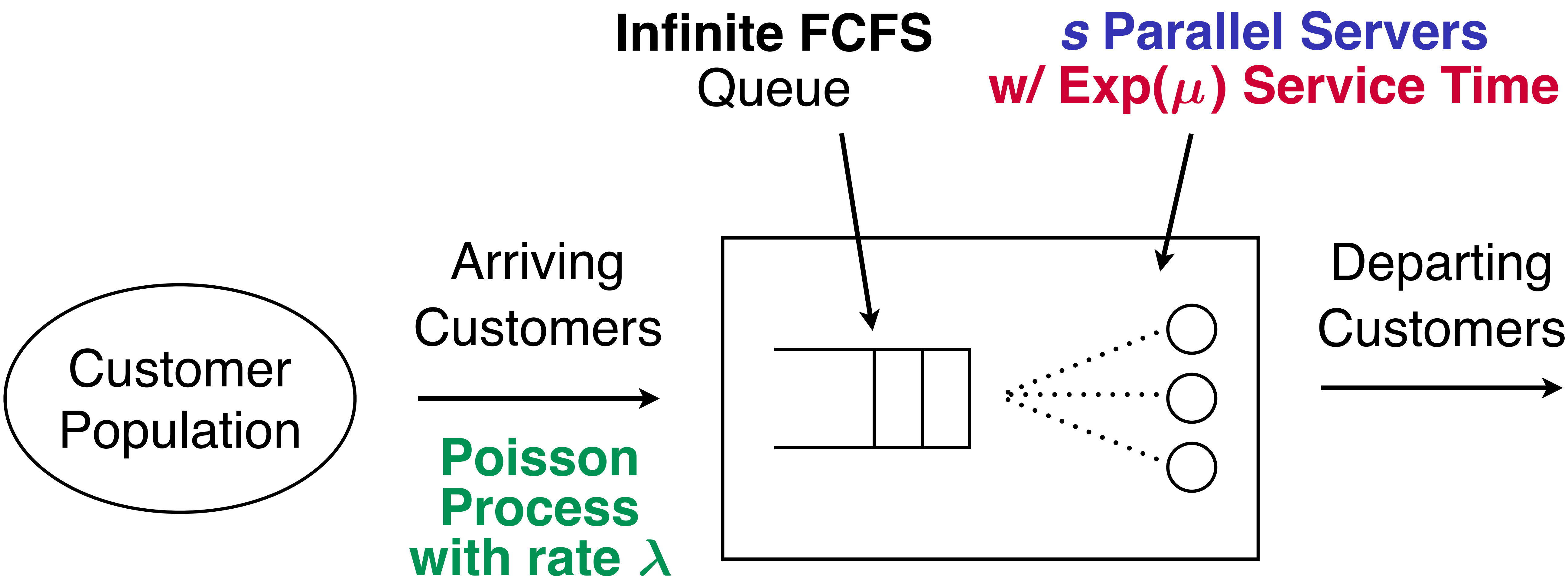
Let $X(t)$ be number of arrivals of the process in $(0,t)$.

$X(t)$ increases from n to $n+1$ at rate λ:

$$q(n, n+1) = \lambda \qquad \text{for all } n \geq 0$$

This simple example will allow us to create other more complicated examples...

Queueing System M / M / *s*



Queueing System M / M / *s*

Example — M / M / *s* Queue

Consider load-balancing s replicated database servers.
A request is routed to the next available server.
Requests line-up in a single queue if all servers are busy.
Requests arrive at times of a Poisson Process w/ rate λ:

$$q(n, n{+}1) = \lambda \qquad \text{for all } n \geq 0$$

Service times are random independent ~ Exp(μ):

$$q(n, n{-}1) = n\mu \qquad \text{if all } 0 \leq n \leq s$$

$$q(n, n{-}1) = s\mu \qquad \text{if all } n \geq s$$

Queueing System M / M / s with Balking

Example — M / M / s Queue with Balking

The same setup; requests arrive as Poisson Process (λ).

The load balancer forwards requests randomly to a different data center – with probability $(1-a_n)$ if there are n requests.

I.e., the new request stays with probability a_n:

$$q(n, n+1) = \lambda\, a_n \qquad \text{for all } n \geq 0$$

Service times are as before independent ~ Exp(μ):

$$q(n, n-1) = n\mu \qquad \text{if all } 0 \leq n \leq s$$

$$q(n, n-1) = s\mu \qquad \text{if all } n \geq s$$

From Jump Rates To a Discrete-time Markov Chain with Randomly Timed Jumps

Jump Rates vs. Transition Probabilities

Example

In our example the jump rates are

$$q(i,j) = \lambda u(i,j)$$

We leave i with rate λ (a Poisson process)...

... and go to j with probability $u(i,j)$

Note this can be reversed

$$u(i,j) = q(i,j)/\lambda$$

This is how we can construct the MC from its jump rates...

... but what is λ if we only know $q(i,j)$?

From Jump Rates To a Markov Chain

Constructing the MC from its jump rates

Let $\lambda_i = \sum_{j \neq i} q(i,j)$. This is the rate at which X_t leaves i.

$\lambda_i = \infty$ ... X_t leaves i immediately ... so we assume $\lambda_i < \infty$.

For $\lambda_i > 0$ we define (a **r**outing matrix)

$$r(i,j) = q(i,j)/\lambda_i$$

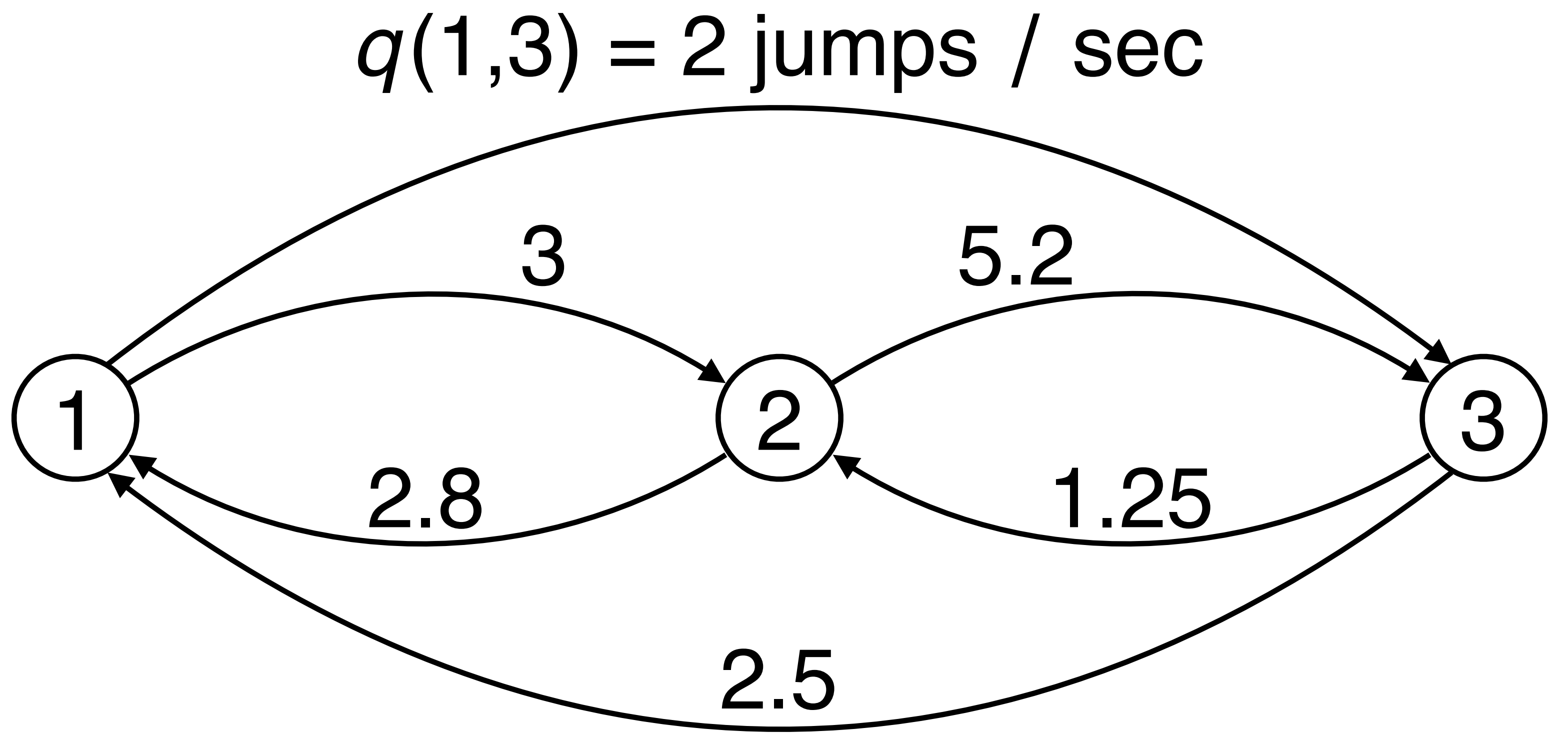
If X_t is in state i with $\lambda_i = 0$, then X_t stays in i forever.

If $\lambda_i > 0$, then X_t waits in i for a random time $\sim \mathrm{Exp}(\lambda_i)$.

Then X_t jumps to j with probability $r(i,j)$.

We have many Poisson processes – with λ_i for each state i

From Jump Rates To a Markov Chain



$\lambda_1 = 2 + 3 = 5$ jumps away from ① / second

$r(1,2) = q(1,2)/\lambda_1 = 3/5 = P(Y_{n+1} = ② | Y_n = ①)$

$r(1,3) = q(1,3)/\lambda_1 = 2/5$

$\lambda_2 = 2.8 + 5.2 = 8$ $\quad r(2,1) = 2.8/8$ $\quad r(2,3) = 5.2/8$

$\lambda_3 = 1.25 + 2.5 = 3.75$ $\quad r(3,1) = 2.5/3.75$ $\quad r(3,2) = 1.25/3.75$

A Single Poisson Process for Markov Chains with Bounded Rates

Constructing the MC from bounded jump rates

Let $\lambda_i = \sum_{j \neq i} q(i,j)$ and $\lambda_{max} = \max_i \lambda_i$.

Assume that $\lambda_{max} < \infty$ (bounded rates) and define

$$u(i,j) = q(i,j)/\lambda_{\max} \qquad \text{for } j \neq i$$

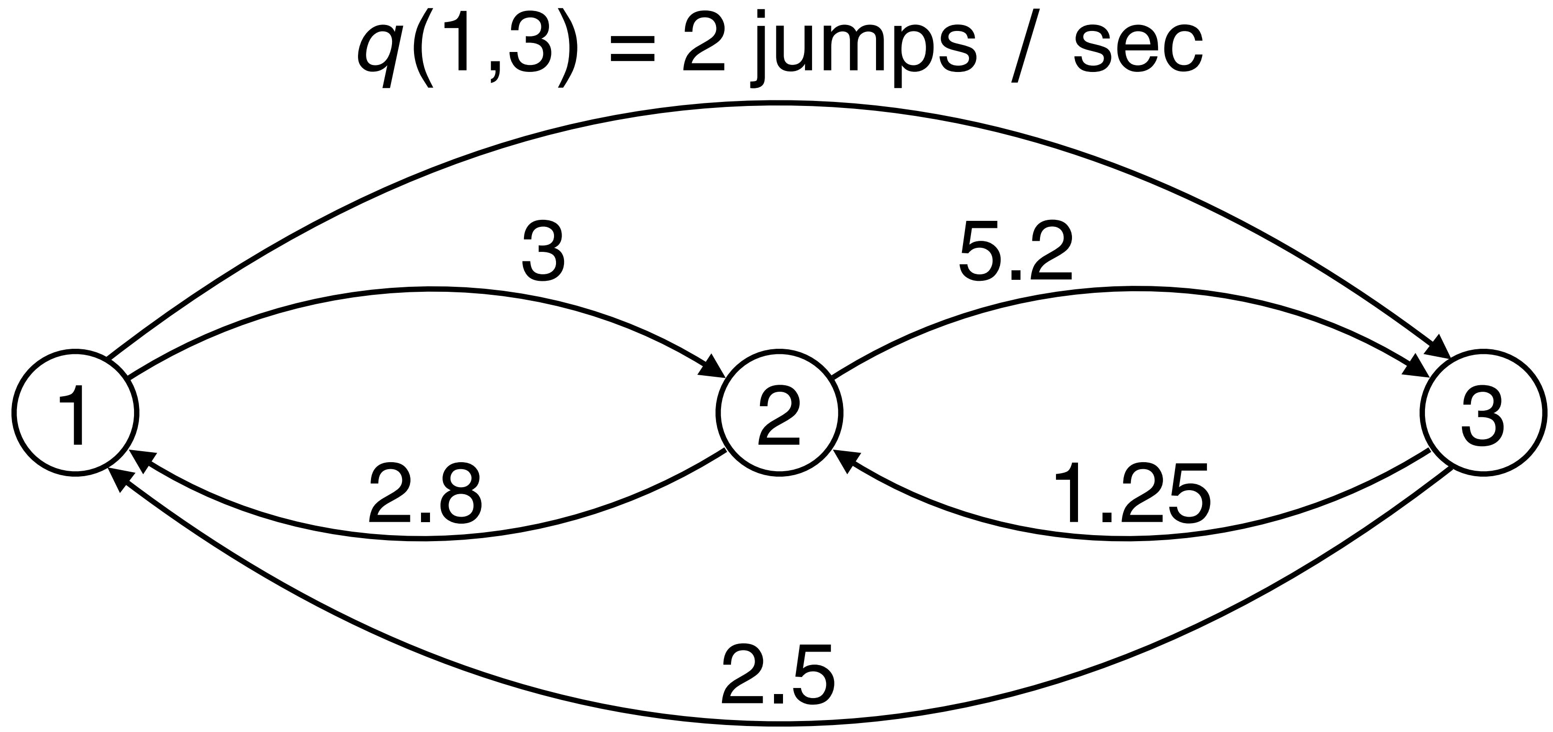
$$u(i,i) = 1 - \sum_{j \neq i} u(i,j)$$

X_t tries to make a transition w/ rate λ_{max} (Poisson Process).

X_t leaves i for j with probability $u(i,j)$, but stays in i with $u(i,i)$.

Any MC with bounded rates can be constructed as our initial example!

From Jump Rates To a Markov Chain



$\lambda_1 = 2 + 3 = 5$ $\quad$ $\lambda_2 = 2.8 + 5.2 = 8$ $\quad$ $\lambda_3 = 1.25 + 2.5 = 3.75$

$\lambda_{max} = \max\{5, 8, 3.75\} = 8$ jumps / sec (on average, random times)

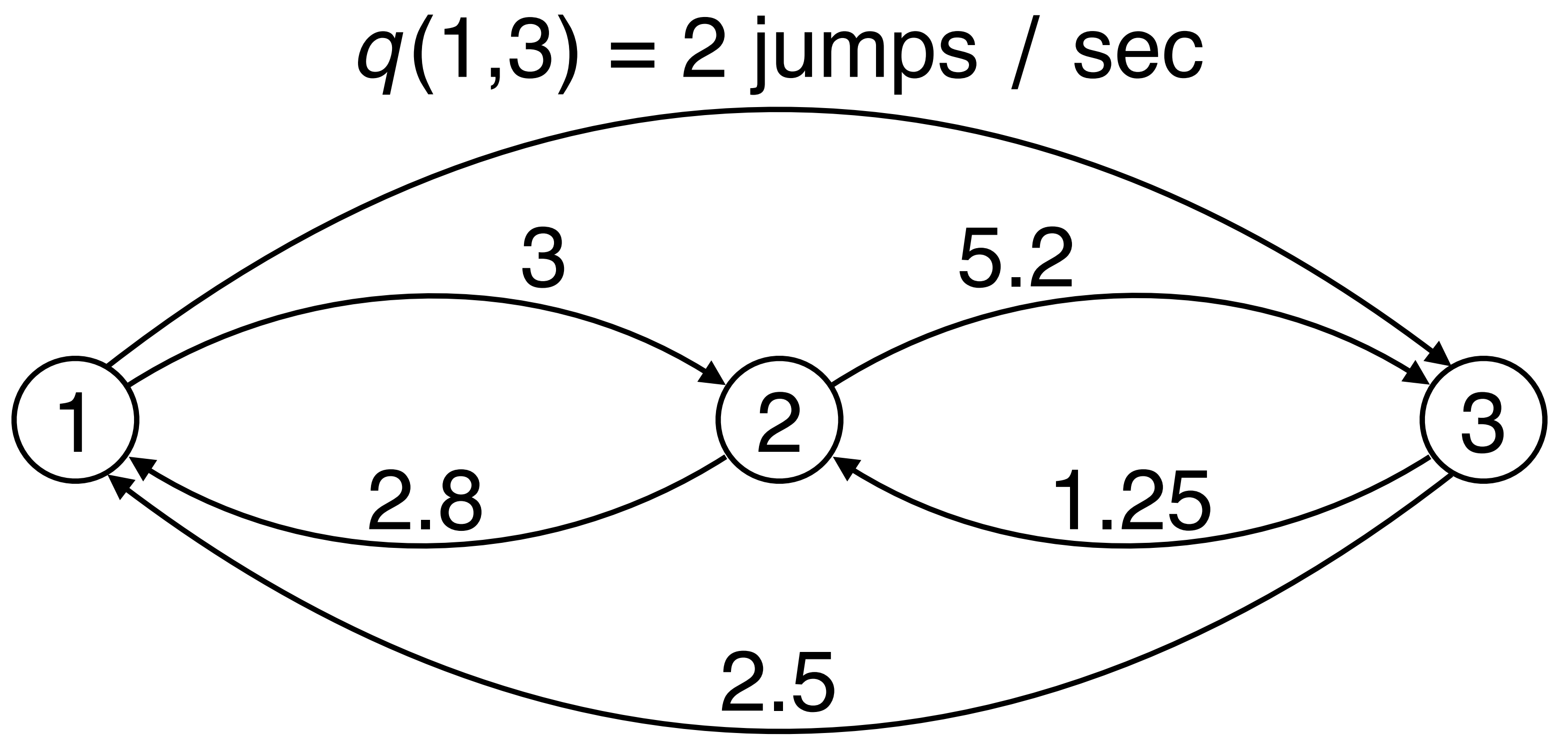
$u(1,2) = q(1,2)/\lambda_{max} = 3/8 = P(Y_{n+1} = ② \mid Y_n = ①)$

$u(1,3) = q(1,3)/\lambda_{max} = 2/8$ $\quad$ $u(1,1) = 1 - 5/8 = 3/8$

$u(2,1) = 2.8/8$ $\quad$ $u(2,3) = 5.2/8$ $\quad$ $u(2,2) = 0$

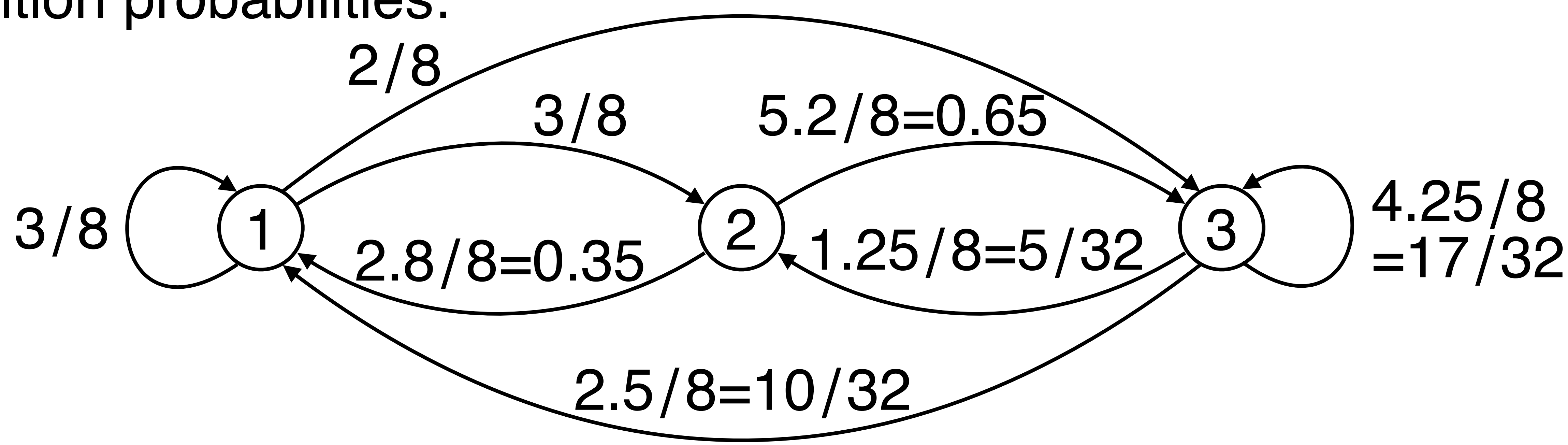
$u(3,1) = 2.5/8$ $\quad$ $u(3,2) = 1.25/8$ $\quad$ $u(3,3) = 1 - 3.75/8 = 4.25/8$

From Jump Rates To a Markov Chain

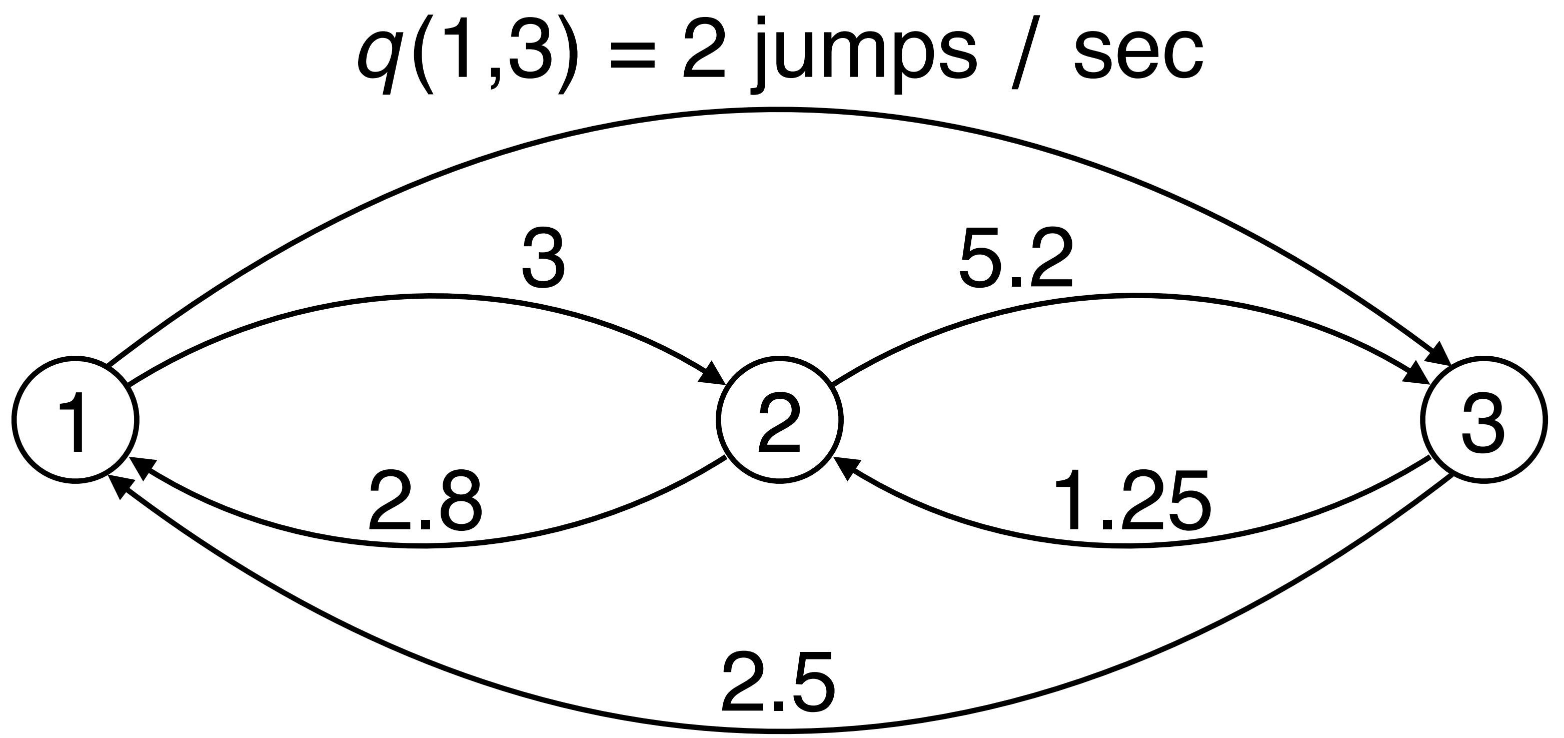


Poisson process: $\lambda_{max} = \lambda_2 = 8$ jumps/sec (on average, random times)

Transition probabilities:

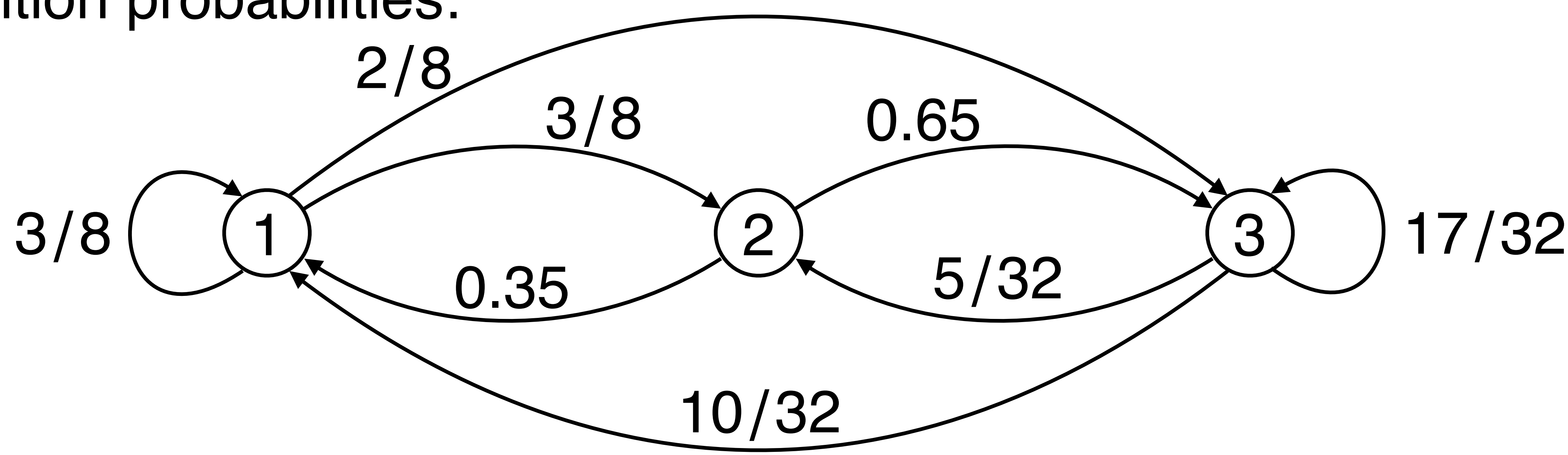


From Jump Rates To a Markov Chain



Poisson process: $\lambda_{\max} = \lambda_2 = 8$ jumps/sec (on average, random times)

Transition probabilities:



Computing Transition Probabilities from Jump Rates

Computing Transition Probabilities

Our goal is to use the jump rates $q(i,j)$ to find the transition probabilities

$$p_t(i,j) = P(X_t = j \mid X_0 = i)$$

Recall the Chapman-Kolmogorov equation:

$$\sum_k p_s(i,k)\, p_t(k,j) = p_{s+t}(i,j)$$

It can be differentiated to show that

$$p_t'(i,j) = \sum_{k \neq i} q(i,k) p_t(k,j) - \lambda_i p_t(i,j)$$

Computing Transition Probabilities

Kolmogorov's Backward Equation

Define a matrix

$$Q(i,j) = \begin{cases} q(i,j) & \text{if } j \neq i \\ -\lambda_i & \text{if } j = i. \end{cases}$$

The equation

$$p_t'(i,j) = \sum_{k \neq i} q(i,k) p_t(k,j) - \lambda_i p_t(i,j)$$

can be rewritten in matrix notation as $p_t' = Q\, p_t$.

This is the Kolmogorov's Backward Equation.

Computing Transition Probabilities

Kolmogorov's Backward Equation

The **Kolmogorov's Backward Equation** is

$$p_t' = Q\, p_t$$

Kolmogorov's Forward Equation

Using similar techniques, one can obtain **Kolmogorov's Forward Equation**

$$p_t' = p_t\, Q$$

Solving Kolmogorov Backward Equation

Theorem

The solution of the Kolmogorov's backward equation

$$p_t' = Q\,p_t$$

can be obtained as e^{Qt} (similarly as if Q were a number).

The exponential function for matrix Q is defined as

$$e^{Qt} = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n\,!} = \sum_{n=0}^{\infty} Q^n \cdot \frac{t^n}{n\,!}$$

Proof: By direct differentiation.