Lecture 14: Queueing theory

Last time

Time-homogenous Markov process with discrete space of states can be described as a combination of two sets of parameters.

(1) $\Lambda_1, \Lambda_2, \ldots$ where $(\forall_i)(\Lambda_i > 0)$ and time spent is state (i) has distr. $T_i \sim Exp(\Lambda_i)$

(2) Transition matrix $P = [p_{ij}]_{\substack{n \times n \\ \infty \times \infty}}$ describing probability of transition between any two states i and j.

Stationary distribution for THMPWDS $\mathcal{X}(t)$

- Transition function in time t: $P_{i,j}(t) = Pr[\mathcal{X}(t) = j \mid \mathcal{X}(0) = i]$
- Transition matrix in time t: $\mathbb{P}(t) = [P_{i,j}(t)]_{\substack{n \times n \\ \infty \times \infty}}$

example: Poisson Process $N(t)$ with int. λ

$$P_{i,j}(t) = Pr[\underbrace{N(t)}_{Poiss(\lambda t)} = j - i] = \frac{e^{-\lambda t} \cdot (\lambda t)^{j-i}}{(j-i)!}$$

- Stationary distribution: $\overline{\pi} = (\pi_1, \pi_2, \ldots)$ such that $(\forall_{t \geq 0})(\overline{\pi} = \overline{\pi} \cdot \mathbb{P}(t))$ and $\pi_1 + \pi_2 + \ldots = 1$

Theorem (global equilibrium equations)

Let $\bar{\pi}$ be a stationary distribution for THMRVDS described by $P=[p_{i,j}]$ and $\Lambda_1, \Lambda_2, \dots \Lambda_i > 0$. Then $(\forall i \in S)(\pi_i \Lambda_i = \sum_{k \in S} \pi_k \cdot \Lambda_k \cdot p_{k,i})$.

How often we are leaving state (i) in the long run.

how often we leave some (k) and go to (i)

Remark: What if all $\Lambda_1 = \Lambda_2 = \dots = \Lambda_n$ then $(\forall i \in S)(\pi_i = \sum_{k \in S} \pi_k p_{k,i})$

→ in matrix form $\bar{\pi} = \bar{\pi} \cdot P$

Local equilibrium equations $(\forall i,j \in S)(\pi_i \lambda_{i,j} = \pi_j \lambda_{j,i})$

(i) ⇝ (j) (j) ⇝ (i)

Remark: do not need to hold even if $\bar{\pi}_i$ exists.

Lemma if we find $\bar{\pi}$ for which LEE hold, then also GEE hold for $\bar{\pi}$ (and $\bar{\pi}$ is stationary distribution)

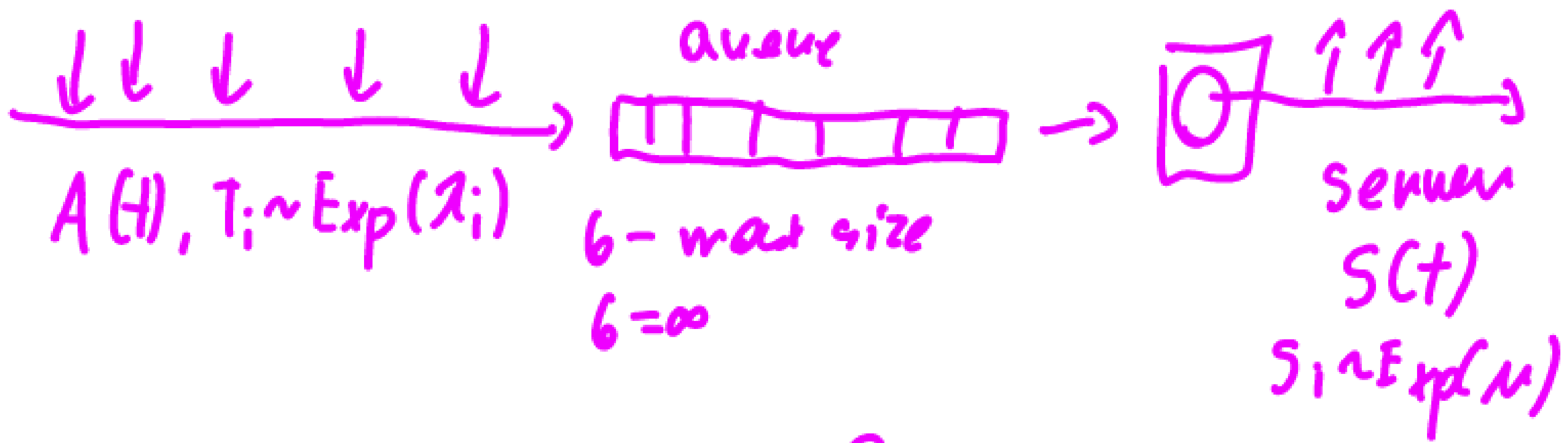
Pf We sum up both sides of LEE over all states.

- Left-side: $\sum_{i \in S} \pi_i \lambda_{i,j} = \sum_{i \in S} \pi_i \Lambda_i \cdot p_{i,j}$ ← frequency of entering (j) in GEE.

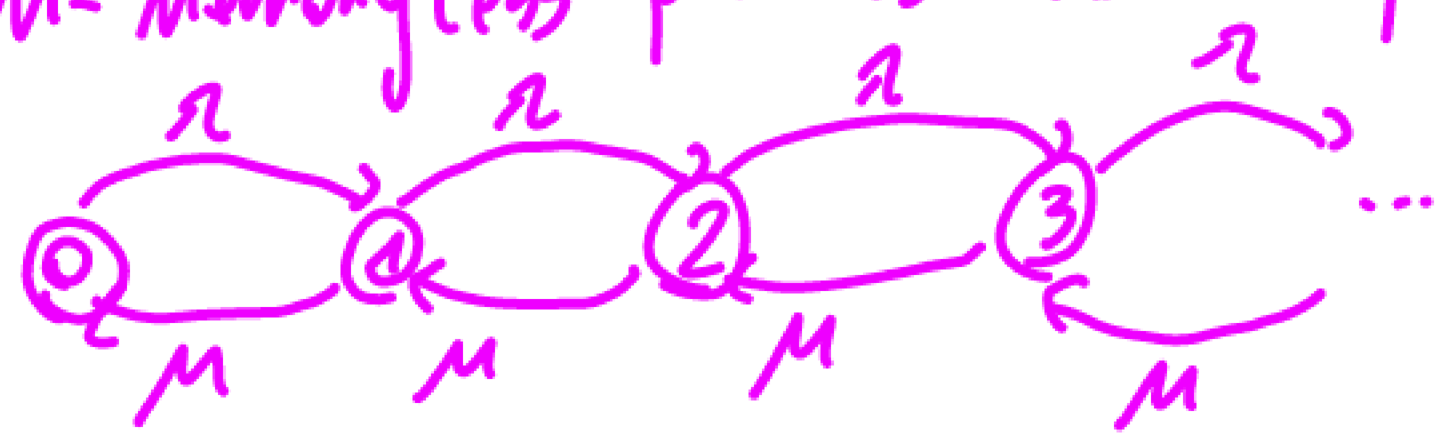
  def: $\lambda_{i,j} = \Lambda_i \cdot p_{i,j}$

- Right-side: $\sum_{i \in S} \pi_j \cdot \lambda_{j,i} = \pi_j \sum_{i \in S} \lambda_{j,i} = \pi_j \cdot \Lambda_j$ ← frequency of leaving (j) in GEE.

Example $M/M/1/\infty$



M = Memoryless process = Poisson process



N - the number of clients in the system.

- LEE : $(\forall i \in S)\ (\pi_i \lambda = \pi_{i+1} \mu)$

$$\pi_{i+1} = \frac{\lambda}{\mu} \cdot \pi_i = \left(\frac{\lambda}{\mu}\right)^2 \cdot \pi_{i-1} = \ldots = \left(\frac{\lambda}{\mu}\right)^{i+1} \cdot \pi_0$$

- $\sum_{i \in S} \pi_i = 1 \to \sum_{i \in S} \left(\frac{\lambda}{\mu}\right)^i \cdot \pi_0 = 1 \to \pi_0 = \frac{1}{\sum_{i \in S} \left(\frac{\lambda}{\mu}\right)^i} \overset{\lambda < \mu}{=\!=} \frac{1}{\frac{1}{1 - \frac{\lambda}{\mu}}} = 1 - \frac{\lambda}{\mu}$

combining two above we get that:

$$\pi_i = \left(\frac{\lambda}{\mu}\right)^i \left(1 - \frac{\lambda}{\mu}\right)$$

---

① $E[N] \overset{+\infty}{=} \sum_{i \in S} i \cdot \pi_i = \sum_{i \in S} i \left(\frac{\lambda}{\mu}\right)^i \cdot \left(1 - \frac{\lambda}{\mu}\right) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right) \sum_{i \in S} i \left(\frac{\lambda}{\mu}\right)^{i-1}$

$|x| < 1$

$\frac{1}{(1-x)^2} = \sum_{i=0}^{n} i x^{i-1}$

$$= \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}} = \frac{\lambda}{\mu - \lambda}$$

② U - time when the server is in use (in %) in long term perspective.

$$U = 1 - \pi_0 = 1 - \left(1 - \frac{\lambda}{\mu}\right) = \frac{\lambda}{\mu}$$

③ How long a client needs to wait (to be served) if there are n clients in the system when he appears.

$$S = S_1 + S_2 + \ldots + S_n, \quad S_i \sim Exp(\mu) \rightarrow S \sim \Gamma(n, \mu), \quad g_n(x) = \mu \frac{(\mu x)^{n-1}}{\Gamma(n)} \cdot e^{-\mu x}$$

④ average time from entering to leaving a system for a client.

Law of total probability for continuous distributions:

$$f_V(x) = \sum_{n=0}^{\infty} g_{n+1}(x) \cdot Pr[n \text{ clients in the system}] = \ldots =$$

$g_{n+1}(x)$: Pr[wait x | n clients in the system]

$$= (\mu - \lambda) e^{-(\mu - \lambda)x}, \quad V \sim Exp(\mu - \lambda)$$

Example 2 $M/M/\infty \equiv P2P$, N - number of clients = number of servers