STATS 217: Introduction to Stochastic Processes I

Lecture 25

Last time: Stationary distributions

- Let $(X_t)_{t \geq 0}$ be a CTMC on Ω. A probability distribution π on Ω is said to be a stationary distribution if

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

- This is equivalent to the condition that

$$\pi Q = 0.$$

- In terms of the matrix Q, the detailed balance conditions are given by

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

Convergence theorem

Let $(X_t)_{t \geq 0}$ be an irreducible CTMC on a finite state space Ω. Then, there exists a unique stationary distribution π, and

$$\max_{x \in \Omega} \mathrm{TV}(P^t(x, \cdot), \pi) \to 0 \quad \text{as } t \to \infty.$$

We have already done the work to prove this theorem.

Existence of stationary distribution

- The first point is the existence of the stationary distribution.
- Recall the notation $\lambda_i = \sum_{j \neq i} q_j$, $\Lambda = \max_{i \in \Omega} \lambda_i$.
- Since Ω is finite, $\Lambda < \infty$. In this case, recall that we have the representation $X_t = Y_{N(t)}$, where $N(t)$ is a PPP with rate λ and Y_n is a DTMC with the transition matrix

$$U_{ij} = \frac{q_{ij}}{\Lambda} \quad \forall i \neq j \qquad\qquad U_{ii} = 1 - \frac{\lambda_i}{\Lambda}$$

- Since $(X_t)_{t \geq 0}$ is irreducible, so is U, and hence, it has a unique stationary distribution π.

Existence of stationary distribution

- We can check that $\pi Q = 0$. Indeed,

$$
\begin{aligned}
\sum_{i \in \Omega} \pi_i q_{ij} &= \pi_j q_{jj} + \sum_{j \neq i} \pi_i q_{ij} \\
&= -\pi_j \lambda_j + \sum_{i \neq j} \pi_i U_{ij} \Lambda \\
&= -\pi_j \lambda_j + \Lambda \sum_{i \in \Omega} \pi_i U_{ij} - \Lambda \pi_j U_{jj} \\
&= -\pi_j \lambda_j + \Lambda \pi_j - \pi_j (\Lambda - \lambda_j) \\
&= 0.
\end{aligned}
$$

Convergence

- Note that

$$\begin{aligned} \mathrm{TV}(P^{t+s}(x,\cdot),\pi) &= \mathrm{TV}(\delta_x P^t P^s, \pi P^s) \\ &\leq TV(\delta_x P^t, \pi). \end{aligned}$$

- Hence, $\mathrm{TV}(P^t(x,\cdot),\pi)$ is non-increasing in t, so it suffices to show that it converges to 0 along (say) the natural numbers.
- But P^1 is an irreducible and aperiodic transition matrix with unique stationary distribution π, so that by looking at the corresponding DTMC, we have

$$\mathrm{TV}(P^n(x,\cdot),\pi) \to 0 \quad \text{as } n \to \infty.$$

Example: M/M/1 queues

- This is a popular queuing model in which the arrival of customers is modelled by a Poisson point process with rate λ. There is a single server, and service times are independent and exponentially distributed with parameter μ.
- Due to the memorylessness property of the exponential distribution, this can be modelled as a continuous time birth and death chain with jump rates

$$\begin{aligned} Q_{n,n+1} &= \lambda, \quad n = 0, 1, \ldots \\ Q_{n,n-1} &= \mu, \quad n = 1, 2, \ldots \end{aligned}$$

- Suppose instead that there are s servers, and customers are served if there is at least one server available. This is called the $M/M/s$ queueing model, and the jump rates are now

$$\begin{aligned} Q_{n,n+1} &= \lambda, \quad n = 0, 1, \ldots, \\ Q_{n,n-1} &= n\mu, \quad n = 1, \ldots, s, \\ Q_{n,n-1} &= s\mu, \quad n = s+1, s+2, \ldots, \end{aligned}$$

M/M/1 queues

- Suppose that $\lambda < \mu$ i.e., the rate of arrivals is smaller than the rate of service. Otherwise, the size of the queue explodes.
- When $\lambda < \mu$, we can use the detailed balance conditions

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

  to find the stationary distribution

$$\pi_n = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n, \quad n = 0, 1, \ldots$$

- Given this stationary distribution, one can compute many quantities of interest. For instance, the long-run fraction of time that the server is busy is

$$1 - \pi_0 = \frac{\lambda}{\mu}.$$

M/M/1 queues

- Moreover, the expected length of the queue under the equilibrium distribution is
$$L = \sum_{n=0}^{\infty} n\pi_n = \frac{\lambda}{\mu - \lambda}.$$
- Another important quantity is the total time T (waiting time + time with the server) spent by a customer in the system.
- If there are n customers already in the system when a new customer joins the queue, then since service times are i.i.d. exponentials with parameter μ, the total time spent by the customer is distributed as a sum of $n+1$ i.i.d. exponentials with parameter μ.

M/M/1 queues

- Then, using the law of total probability, we have

$$\begin{aligned}\mathbb{P}[T \leq t] &= \mathbb{P}[T \leq t \mid n \text{ customers already in the system}] \cdot \pi_n \\ &= 1 - \exp(-t(\mu - \lambda)),\end{aligned}$$

i.e. T has exponential distribution with mean

$$W = \frac{1}{\mu - \lambda} = \frac{L}{\lambda}.$$

Little's law

- The relationship
$$L = \lambda W$$
is called **Little's law** and is true even without the specific distributional assumptions (i.e. Poisson arrivals and exponential waiting times). Such queues are called $GI/G/1$ queues.
- Here's the intuition: Suppose each customer pays \$1 for each minute of time they spend in the system. When there are n customers in the system, the establishment is earning \$$n$ per minute, and hence, the establishment is earning an average of \$$L$ per minute.
- On the other hand, if each customer pays for their entire duration when they arrive, then the average rate of earning is $\lambda \times W$ per minute.