Stochastic Modelling Exam Notes

Stochastic Modelling (University of Melbourne)

Discrete Time Markov Chains

Then $\mathcal{S}$ can be partitioned into a collection of disjoint subsets $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \dots, \mathcal{S}_M$ (where M might be infinite) such that $j, k \in \mathcal{S}_m$ if and only if $j \Leftrightarrow k$.

So the state space $\mathcal{S}$ of a DTMC is partitioned into communicating classes by the communication relation $\leftrightarrow$.

If a DTMC has only one communicating class (i.e., all states communicate) then it is called an irreducible DTMC. Otherwise it is called reducible.

Definition: An irreducible DTMC is recurrent if $h_{i,j} = 1$ for every $i, j \in \mathcal{S}$. Otherwise it is transient.

Definition: An irreducible DTMC is periodic with period d if any (hence every) state has period $d > 1$. Otherwise it is aperiodic.

Definition: A recurrent DTMC is *positive recurrent* if $\mu_j < \infty$ for every $j \in \mathcal{S}$. Otherwise it is *null recurrent*.

Lemma: Any irreducible DTMC with finite state space is positive recurrent.

Theorem: (∗∗) An irreducible (time-homogeneous) DTMC with countable state space $\mathcal{S}$ has a stationary measure. It has a unique stationary distribution if and only if the chain is positive recurrent, and in this case $\pi_i = 1/\mu_i$ for each $i \in \mathcal{S}$.

Combining Theorems (∗) and (∗∗) we see that for a positive recurrent irreducible DTMC, the long run proportion of time spent in state i is the stationary probability π_i.

We say that a DTMC is ergodic if the limiting distribution exists and does not depend on the starting distribution. This is equivalent to saying that $a_j = \lim_{n\to\infty} p_{i,j}^{(n)}$ exists for each $i, j \in \mathcal{S}$, does not depend on i, and $\sum_{j\in\mathcal{S}} a_j = 1$.

Theorem: An irreducible (time-homogeneous) DTMC is ergodic if and only if it is aperiodic, and positive recurrent. For an ergodic DTMC the limiting distribution is equal to the stationary distribution.

Theorem: For a (time-homogeneous) DTMC, if a limiting distribution exists then it is a stationary distribution.

Theorem:(∗∗∗) Let $(X_n)_{n\in\mathbb{Z}_+}$ be an irreducible, aperiodic (time-homogeneous) DTMC with countable state space $\mathcal{S}$. Then for all $i, j \in \mathcal{S}$,

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_0 = i) \to \frac{1}{\mu_j}, \quad \text{as } n \to \infty.$$

Let $T(j) = \inf\{n \geq 0 : X_n = j\}$ denote the first hitting time of j (it is a stopping time). Then

$$m_{i,j} := \mathbb{E}[T(j) | X_0 = i],$$

is the expected time to reach state j starting from state i. By definition

- $m_{j,j} = 0$,
- if $h_{i,j} < 1$ then $\mathbb{P}(T(j) = \infty | X_0 = i) > 0$ so $m_{i,j} = \infty$,
- if $h_{i,j} = 1$ then $m_{i,j}$ may or may not be finite.

Let $T^+(i) = \inf\{n > 0 : X_n = i\}$. Then

$$\mu_i = \mathbb{E}[T^+(i) | X_0 = i],$$

is the expected time to return to state i.

Theorem: The vector $(m_{i,A})_{i\in\mathcal{S}}$ of mean hitting times is the minimal non-negative solution to

$$m_{i,A} = \begin{cases} 0, & \text{if } i \in A \\ 1 + \sum_{j\in\mathcal{S}} p_{i,j} m_{j,A}, & \text{otherwise.} \end{cases}$$

$$\mu_i = 1 + \sum_{j\in\mathcal{S}} p_{i,j} m_{j,i}.$$

Let

$$h_{i,j} = \mathbb{P}(j \in \{X_0, X_1, \dots\} | X_0 = i),$$

which is the probability that we ever reach state j, starting from state i. The quantities $h_{i,j}$ are referred to as hitting probabilities. Note that

- $h_{i,i} = 1$
- $h_{i,j} > 0$ if and only if $i \to j$.

Let

$$f_i = \mathbb{P}(i \in \{X_1, X_2, \dots\} | X_0 =$$

which is the probability that we ever return to
from i. The quantities f_i are referred to as return probabilities.

The characteristic equation of

$$h_{i,0} = p h_{i+1,0} + (1-p) h_{i-1,0}$$

is

$$pz^2 - z + (1-p) = 0$$

which (as we have seen before) has roots $z = 1$ and $z = (1-p)/p$. If $(1-p)/p \neq 1$, the general solution for $j \geq 1$ is of the form

$$h_{i,0} = A + B\left(\frac{1-p}{p}\right)^j.$$

Simple Random Walk

Stirling's formula $n! \approx \sqrt{2\pi n}\, n^n e^{-n}$ gives us the fact that

$$p_{j,j}^{(2n)} \approx \frac{(4p(1-p))^n}{\sqrt{n\pi}},$$

and the series $\sum_{n=0}^{\infty} p_{j,j}^{(2n)}$

- diverges if $4p(1-p) = 1$ (i.e. $p = 1/2$), so the DTMC is recurrent
- converges if $4p(1-p) < 1$ (i.e. $p \neq 1/2$) (compare to geometric series), so the DTMC is transient.

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,m} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,m} \end{pmatrix}$$

Poisson Processes

Definition:

A nonnegative integer-valued process $(N_t)_{t\geq 0}$ is a Poisson process with a rate λ if

(i) it has independent increments on disjoint intervals: for $k \geq 2$ and $0 \leq s_1 < t_1 \leq s_2 < \cdots < t_k$,

$$N_{t_1} - N_{s_1}, \dots, N_{t_k} - N_{s_k}$$

are independent variables, and

(ii) For each $t > s \geq 0$, $N_t - N_s \sim \text{Poi}(\lambda(t-s))$.

Let $T_0 = 0$ and $T_j = \min\{t : N_t = j\}$ (the time of j^{th} jump) and define $\tau_j = T_j - T_{j-1}$ (time between the $(j-1)^{\text{st}}$ and j^{th} jumps).

Theorem: $(N_t)_{t\geq 0}$ is a Poisson process with rate λ if and only if $(\tau_j)_{j\in\mathbb{N}}$ are independent Exp(λ) random variables.

Theorem: Let $(N_t)_{t\geq 0}$ and $(M_t)_{t\geq 0}$ be two independent Poisson processes with rates λ_1 and λ_2 respectively and $L_t = N_t + M_t$. Then $(L_t)_{t\geq 0}$ is a Poisson process with rate $\lambda_1 + \lambda_2$.

Suppose in a Poisson process $(N_t)_{t\geq 0}$ each 'customer' is 'marked' independently with probability p. Let M_t count the number of 'marked customers' that arrive on $[0, t]$.

Theorem. The processes $(M_t)_{t\geq 0}$ and $(N_t - M_t)_{t\geq 0}$ are independent Poisson processes with rates λp and $\lambda(1-p)$ respectively.

Suppose that $(N_t)_{t\geq 0}$ is a Poisson process and $(X_i)_{i\in\mathbb{N}}$ are independent and identically-distributed random variables, which are also independent of $(N_t)_{t\geq 0}$.

For $t \geq 0$, define $Y_t = \sum_{j\leq N_t} X_j$. Then $(Y_t)_{t\geq 0}$ is called a compound Poisson process.

It can be shown that $(Y_t)_{t\geq 0}$ has independent increments and it is possible to compute the distribution of Y_t by conditioning on N_t.

Theorem: The conditional distribution of $(T_1, \cdots, T_k)$ given that $N_t = k$ is the same as the distribution of order statistics of a sample of k independent and identically-distributed random variables uniformly distributed on $[0, t]$.

I.e.

$$(T_1, \cdots, T_k) | \{N_t = k\} \stackrel{d}{=} (U_{(1)}, \cdots, U_{(k)})$$

where $U_1, \cdots, U_k$ are independent $\sim U(0, t)$.

The same representation holds for the conditional distribution of $(T_1, \cdots, T_k)$ given that $T_{k+1} = t$.

Continuous Time Markov Chains

For a CTMC, if $X_t = i$, we wait an exponential(λ_i) time and then jump to a new state. The probability of jumping to j is $b_{i,j}$.

$$\frac{q_{i,j}}{\sum_{\ell\in\mathcal{S}} q_{i,\ell}} = b_{i,j}. \quad \sum_{\ell\in\mathcal{S}} q_{i,\ell} = \lambda_i)$$

A similar argument to the one we saw in the DTMC setting shows that an irreducible finite-state CTMC is positive recurrent.

Theorem: An irreducible and positive recurrent CTMC has a unique stationary distribution π. For such a CTMC, the limiting proportion of time spent in state i is π_i and the limiting distribution is π (irrespective of the initial distribution).

For an irreducible transient CTMC, the limiting proportion of time spent in each state is 0, and there is no limiting *distribution*.

$$\pi_i = \frac{\mathbb{E}[T_1^{(i)} | X_0 = i]}{\mathbb{E}[T^{i,i} | X_0 = i]}.$$

The quantity $\frac{\mathbb{E}[T_1^{(i)}|X_0=i]}{\mathbb{E}[T^{i,i}|X_0=i]}$ is a bit like the proportion of time spent at i up to the first time that we *return* to i (start from i initially). The numerator of this quantity is $\frac{1}{\lambda_i} = -\frac{1}{q_{i,i}}$. By a first step analysis, the denominator is

$$\frac{1}{\lambda_i} + \sum_{j\in\mathcal{S}} b_{i,j} m_{j,i}.$$

This is because on average we take time $1/\lambda_i$ to escape from i, at which point we go to j with probability $b_{i,j}$ and then we have to get from j to i (which takes time $m_{j,i}$ on average).

Theorem: For a CTMC with state space $\mathcal{S}$, and $A \subset \mathcal{S}$, the vector of mean hitting times $(m_{i,A})_{i\in\mathcal{S}}$ is the minimal non-negative solution to

$$m_{i,A} = \begin{cases} 0, & \text{if } i \in A, \\ \frac{1}{\lambda_i} + \sum_{j\in\mathcal{S}} b_{i,j} m_{j,A}, & \text{if } i \notin A. \end{cases}$$

Observe that

$$\begin{aligned} p_{i,j}^{(s+t)} &= \sum_k \mathbb{P}(X_{s+t} = j | X_s = k, X_0 = i)\mathbb{P}(X_s = k | X_0 = i) \\ &= \sum_k p_{i,k}^{(s)} p_{k,j}^{(t)}. \end{aligned}$$

$$\frac{d}{dt} P^{(t)} = Q P^{(t)} \quad (\ddagger) \qquad \text{backward equations}$$

For solving $p_{i,j}(t)$

$$\frac{d}{dt} P^{(t)} = P^{(t)} Q. \quad (\dagger) \qquad \text{forward equations}$$

Differential Equation: $p_{i,j}'(t) = \alpha_1 + \alpha_2 p_{i,j}(t)$

$\Rightarrow p_{i,j}(t) = A + Be^{\alpha_2 t}$, $\alpha_1, \alpha_2, A, B \in \mathbb{R}$

The generator of such a birth and death process has the form

$$Q = \begin{pmatrix} -\nu_0 & \nu_0 & 0 & 0 & 0 & \ddots \\ \mu_1 & -(\mu_1+\nu_1) & \nu_1 & 0 & 0 & \ddots \\ 0 & \mu_2 & -(\mu_2+\nu_2) & \nu_2 & 0 & \ddots \\ 0 & 0 & \mu_3 & -(\mu_3-\nu_3) & \nu_3 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

The CTMC evolves by remaining in state k for an exponentially-distributed time with rate $\nu_k + \mu_k$, then it moves to state $k+1$ with probability $b_{k,k+1} = \nu_k/(\nu_k + \mu_k)$ and state $k-1$ with probability $b_{k,k-1} = \mu_k/(\nu_k + \mu_k)$, and so on.

Assume $\nu_i > 0$ and $\mu_i > 0$ for all i. Assuming non-explosivity, we derive the stationary distribution (if it exists) by solving $\pi Q = 0$. In fact we can solve the detailed balance equations:

$$\nu_k \pi_k = \mu_{k+1}\pi_{k+1}, \qquad k \in \mathbb{Z}_+$$

This has solution

$$\pi_k = \pi_0 \prod_{\ell=1}^{k} \frac{\nu_{\ell-1}}{\mu_\ell}.$$

Solution : $\underset{\sim}{\pi}\, Q = \underset{\sim}{0}$

Queueing Systems

- Arrival times $T_1, T_2, T_3, \cdots$. The inter-arrival times are $\tau_1 = T_1 - T_0, \tau_2 = T_2 - T_1, \tau_3 = T_3 - T_2 \ldots$
- The inter-arrival times are assumed to be i.i.d.
- Alternatively, we could use a counting process N_t giving the number of arrivals in $[0, t]$, $t \geq 0$.

Standard setup for service

- There is a total of m spaces for both receiving service and waiting for it.
- If there is an idle server, service commences for an arriving customer immediately.
- The service time $S_i^{(j)}$ of the ith customer at the jth server is a random variable.
- The service times are assumed to be independent (and also identically distributed for each fixed j
- When a server is serving a customer, it cannot provide any service to other customers.
- If all servers are busy, then the arriving customers join a queue if there is enough space, otherwise, the customer is rejected.

$M/M/1$ queue

- Arrival stream: Poisson process with intensity λ
- Service: $n = 1$ server, service times $\sim \exp(\mu)$
- Infinite space for waiting: $m = \infty$
- The state X_t gives the number of customers at time t:
  - If $X_t = 0$ the server is idle.
  - If $X_t = k \geq 1$ one customer is being served and $k - 1$ customers are waiting in the queue.

This is a CTMC (in fact a birth and death process) with non-zero transition rates $q_{i,i+1} = \lambda$ and $q_{i+1,i} = \mu$ for all $i \in \mathbb{Z}_+$.

$M/M/1$ stationary distribution

Using our results from CTMCs, we see that a stationary distribution for $(X_t)_{t\geq 0}$ exists if (and only if) the chain is positive recurrent. This is equivalent to $\rho = \lambda/\mu < 1$, in which case, for $n \in \mathbb{Z}_+$,

$$\pi_n = (\lambda/\mu)^n \pi_0.$$

Using the normalisation condition $\sum_{i=0}^{\infty} \pi_n = 1$, we see that

$$\pi_0 \sum_{i=0}^{\infty} (\lambda/\mu)^i = 1$$

which tells us that

$$\pi_0 = 1 - \rho$$

and, for $n \geq 0$,

$$\pi_n = (1 - \rho)\rho^n.$$

So the stationary distribution for the number of customers in the system is geometric*$(1 - \rho)$. (Note that this geometric takes values in $\mathbb{Z}_+$).

Recall that ℓ is the expected number of customers in the system, while ℓ_q is the expected number of customers waiting for service (both at stationarity).

Little's law says that

$$\ell = \lambda d,$$

and

$$\ell_q = \lambda \mathbb{E}[W].$$

It follows that the expected waiting time is

$$\mathbb{E}[W] = \frac{\rho}{\mu - \lambda}.$$

Once we have the expected waiting time, we can calculate the expected total time d in the system via the formula

$$d = \mathbb{E}[W] + \frac{1}{\mu} = \frac{1}{\mu - \lambda}.$$

$M/M/a$ Queue

This system has the following properties

- $a \geq 1$ servers,
- a Poisson arrival process with rate λ,
- a FIFO service discipline,
- independent $\exp(\mu)$ service times,
- when an arrival finds more than one idle server, it chooses one at random,
- when k servers are working, the total service rate is $k\mu$.

The transition rates are $q_{i,i+1} = \lambda$, for $i \geq 0$ and $q_{i,i-1} = \mu \times \min(a, i)$ for $i \geq 1$.

Exercise: draw the transition diagram

This is a birth-and-death process with $\nu_i = \lambda$ for $i = 0, 1, 2, \ldots$ and $\mu_i = i\mu$ for $i = 1, 2, \ldots, a$ and $\mu_i = a\mu$ for $i > a$.

Renewal Theory

$N_t = \frac{t}{\mu}$, where $\mu = E(\tau_i)$ renewal time

$\sigma^2 = V(\tau_i)$

$N_t \sim N\left(\frac{t}{\mu}, \frac{t\sigma^2}{\mu^3}\right)$

The residual lifetime R_t at time t is the amount of time until the next renewal time.

Since $T_{N_t} \leq t < T_{N_t+1}$, the residual lifetime at time t is $R_t = T_{N_t+1} - t > 0$.

R_t has density, $\frac{1 - F(y)}{\mu}$

where $F(y)$ is the cdf of τ_i

The age of the renewal process at time t is the time since the most recent renewal, i.e. $A_t = t - T_{N_t}$.

A_t has density, $\frac{1 - F(y)}{\mu}$

where $F(y)$ is the cdf of τ_i

For large t, find the joint probability density function of (R_t, A_t) in the computer packets example.

- First,

$$\mathbb{P}(A_t \leq x, R_t \leq y) = \mathbb{P}(A_t \leq x) - \mathbb{P}(R_t > y) + \mathbb{P}(A_t > x, R_t > y),$$

so

$$\frac{\partial^2 \mathbb{P}(A_t \leq x, R_t \leq y)}{\partial x \partial y} = \frac{\partial^2 \mathbb{P}(A_t > x, R_t > y)}{\partial x \partial y}.$$

- When t is large, $\mathbb{P}(A_t > y, R_t > x) \approx \int_{x+y}^{\infty} \frac{1 - F(z)}{\mu} dz$.
- Hence, the joint pdf is 1/12 if $1 < x + y < 5$ and 0 otherwise.

Probability Distributions

$\rho = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$

Exponential $(X \overset{d}{=} \exp(\lambda))$

$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$

$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$

$E(X) = \frac{1}{\lambda}, \quad V(X) = \frac{1}{\lambda^2}$

Gamma $(X \overset{d}{=} \gamma(r, \alpha))$

$f_Z(z) = \frac{\alpha^r}{(r-1)!} z^{r-1} \cdot e^{-\alpha z}, \ z \geq 0$

$= \frac{\alpha^r}{\Gamma(r)} \cdot z^{r-1} \cdot e^{-\alpha z}, \ z \geq 0$

$\Gamma(r) = \int_0^\infty e^{-x} \cdot x^{r-1} \, dx = (r-1)!$

$E(X) = \frac{r}{\alpha}, \quad V(X) = \frac{r}{\alpha^2}$

$$F_Z(z) = 1 - \sum_{k=0}^{r-1} \frac{(\alpha z)^k}{k!} e^{-\alpha z}$$

Bivariate Normal $((X, Y) \sim N_2(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho))$

$\Rightarrow X \overset{d}{=} N(\mu_X, \sigma_X^2)$ and $Y \overset{d}{=} N(\mu_Y, \sigma_Y^2)$

$(X \mid Y = y) \sim N\left(\mu_X + \rho\sigma_X \frac{(y - \mu_Y)}{\sigma_Y}, \sigma_X^2 (1 - \rho^2)\right)$

Discrete uniform $(X \overset{d}{=} U(a, b))$

$p_X(x) = \frac{1}{b - a + 1}, \quad x = a, a+1, \ldots, b$

$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{1}{12}(b-a) \cdot (b - a + 2)$

Geometric $(X \overset{d}{=} G(p))$

$P(X \geq x) = (1-p)^x$

$P_X(x) = (1-p)^{x-1} \cdot p, \quad x = 1, 2, 3, \ldots$

$E(X) = \frac{1}{p}$

$V(X) = \frac{1-p}{p^2}$

Normal $(X \sim N(\mu, \sigma^2))$

$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x - \mu)^2}, \quad x \in \mathbb{R}$

$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad x \in \mathbb{R}$

$Z = \frac{X - \mu}{\sigma} \iff X = \sigma Z + \mu$

$F_Z(z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$

$\Phi(-z) = 1 - \Phi(z)$

Continuous uniform $(X \overset{d}{=} R(a, b))$

$f_X(x) = \frac{1}{b - a}, \quad a \leq x \leq b$

$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x - a}{b - a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$

$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}$

Brownian

(Sum of continuous functions $\Rightarrow$ continuous) and (sum of independent RV's $\Rightarrow$ independent)

To show $(W_t)_{t \geq 0}$ Brownian Bridge or BM:

- Show for $0 \leq s_1 < t_1 \leq \cdots \leq s_k < t_k$, $W_{t_i} - W_{s_i}$ are independent and continuous
- $E(W_t) = 0$
- $Var(W_t) = t$ for BM or $Cov(W_t, W_s) = s(1 - t)$ for BB for $0 \leq s < t < 1$

$B_t \sim N(0, t)$ and $B_t - B_s \sim N(0, t - s)$