5. Continuous-time Markov Chains

- Many processes one may wish to model occur in continuous time (e.g. disease transmission events, cell phone calls, mechanical component failure times, $\ldots$). A discrete-time approximation may or may not be adequate.

- $\{X(t), t \geq 0\}$ is a **continuous-time Markov Chain** if it is a stochastic process taking values on a finite or countable set, say $0, 1, 2, \ldots$, with the **Markov property** that

$$\mathbb{P}\big[X(t+s)\!=\!j \,|\, X(s)\!=\!i, X(u)\!=\!x(u) \text{ for } 0 \leq u \leq s\big] \\ = \mathbb{P}\big[X(t+s)\!=\!j \,|\, X(s)\!=\!i\big].$$

- Here we consider **homogeneous** chains, meaning $\mathbb{P}[X(t+s)\!=\!j \,|\, X(s)\!=\!i] = \mathbb{P}[X(t)\!=\!j \,|\, X(0)\!=\!i]$

- Write $\{X_n, n \geq 0\}$ for the sequence of states that $\{X(t)\}$ arrives in, and let S_n be the corresponding arrival times. Set $X_n^A = S_n - S_{n-1}$.
- The Markov property for $\{X(t)\}$ implies the (discrete-time) Markov property for $\{X_n\}$, thus $\{X_n\}$ is an **embedded Markov chain**, with transition matrix $P = [P_{ij}]$.
- Similarly, the inter-arrival times $\{X_n^A\}$ must be conditionally independent given $\{X_n\}$. Why?

- Show that X_n^A has a memoryless property conditional on X_{n-1}, $\mathbb{P}\big[X_n^A > t + s \,|\, X_n^A > s, X_{n-1}{=}x\big] = \mathbb{P}\big[X_n^A > t \,|\, X_{n-1}{=}x\big]$ i.e., X_n^A is conditionally exponentially distributed given X_{n-1}.

- We conclude that a continuous-time Markov chain is a special case of a semi-Markov process:

  Construction 1. $\{X(t), t \geq 0\}$ is a continuous-time homogeneous Markov chain if it can be constructed from an embedded chain $\{X_n\}$ with transition matrix P_{ij}, with the duration of a **visit** to i having Exponential (ν_i) distribution.

- We assume $0 \leq \nu_i < \infty$ in order to rule out trivial situations with instantaneous visits.

- An alternative to Construction 1 is as follows:

<u>Construction 2</u>

When $X(t)$ arrives in state i, generate random variables having independent exponential distributions, $Y_j \sim \text{Exponential}\,(q_{ij})$ where $q_{ij} = \nu_i P_{ij}$ for $j \neq i$. Choose the next state to be $k = \arg\min_j Y_j$, and the time until the transition (i.e. the visit time in i) to be $\min_j Y_j$.

- Why is this equivalent to Construction 1?

(i) check that $\mathbb{P}\big[\text{next state is } k\big] = P_{ik}$

(ii) Check that $\min_j Y_j \sim \text{Exponential}\,(\nu_i)$.

- We assume that Markov chains of interest are **regular**, meaning that the # of transitions in any finite length of time is finite with probability 1. A non-regular process is **explosive**. E.g., if an increasing chain takes time α^n to jump from n to $n+1$, then the chain will reach infinity in a finite time, $1/(1-\alpha)$ for $0 < \alpha < 1$.

- We define $P_{ij}(t) = \mathbb{P}[X(t+s) = j \,|\, X(s) = i]$

Lemma 1 (see Ross, Problem 5.8 with solution in the back)

(i) $\lim_{t\to 0} \frac{1-P_{ii}(t)}{t} = \nu_i$

(ii) $\lim_{t\to 0} \frac{P_{ij}(t)}{t} = q_{ij}$ for $j \neq i$

- This leads to another characterization of continuous Markov chains$\dots$

Construction 3. A continuous-time homogeneous Markov chain is determined by its infinitesimal transition probabilities:

$$
\begin{aligned}
P_{ij}(h) &= hq_{ij} + o(h) \text{ for } j \neq 0 \\
P_{ii}(h) &= 1 - h\nu_i + o(h)
\end{aligned}
$$

- This can be used to simulate approximate sample paths by discretizing time into small intervals (the Euler method).

- The Markov property is equivalent to independent increments for a Poisson counting process (which is a continuous Markov chain).

- Lemma 1 can be rewritten as
  $\frac{d}{dt}\gamma(t)\,|_{\,t=0} = \gamma(0)\,Q$
  with $\gamma(t)$ a row vector, $\gamma_i(t) = \mathbb{P}[X(t) = i]$, and

$$
\begin{aligned}
Q_{ij} &= q_{ij} && \text{for } i \neq j \\
Q_{ii} &= -\nu_i = -\sum_{j \neq i} q_{ij}
\end{aligned}
$$

- this identity follows from definitions of $\gamma(t)$ and $P_{ij}(t)$, **noting the necessary interchange of sum & limit.**

Example. A population of size N has I_t infected individuals, S_t susceptible individuals and R_t recovered/removed individuals. New infections occur at rate $\beta I_t S_t$ and infected individuals become removed/recovered at rate γ, i.e. the overall rate of leaving the infected state is γI_t. Supposing the system is Markovian, what are the infinitesimal transition probabilities?

<u>Theorem</u> (Kolmogorov's Backward Equation)

$\frac{d}{dt} P_{ij}(t) = \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t)$.

Or, in matrix notation, with $P(t) = [P_{ij}(t)]$,

$$\boxed{\frac{d}{dt} P(t) = Q\, P(t)}$$

- The backward equation can be used to find transition probabilities, since it has solution $P(t) = e^{Qt}$ [when this is well defined] where $e^{Qt} = \sum_{k=0}^{\infty} Q^k t^k / k!$

<u>Example</u>: For the two-state Markov chain, with rates α and β as shown, find $\mathbb{P}[X(t) = 0 \,|\, X(0) = 0]$.

$$\begin{tikzcd} \boxed{1} \arrow[d, bend left, "\beta"] \\ \boxed{0} \arrow[u, bend left, "\alpha"] \end{tikzcd}$$

Example continued

- To sketch a proof of the backward equation, we first show

Lemma 2. $P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t)P_{kj}(s)$.

Why is this true?

- Then take limits, identifying an issue of exchanging limits and summation but referring to Ross for the details.

- A rather subtly different result is

Kolmogorov's Forward Equation

$\frac{d}{dt}P_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t)$

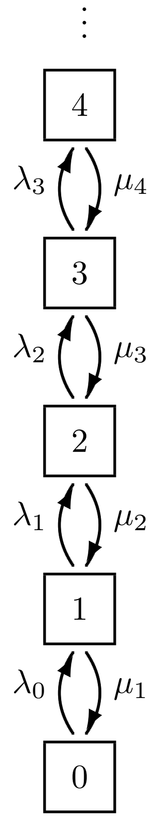
Or, in matrix notation,

$\boxed{\frac{dP}{dt} = P(t)\,Q}$

- This can be written as $\frac{d}{dt}\gamma(t) = \gamma(t)Q$ (Compare with comment on Lemma 1).

- Unfortunately, the forward equation requires regularity conditions to be true (the backward equation is generally true).

- For finite state chains, the forward equation always holds. It can be shown that the forward equation holds whenever $\sum_k P_{ik}(t)\nu_k < \infty$ for any i and t, Example: The continuous-time birth and death process is as shown. For this model, the forward equation has a unique solution which also solves the backward equation (e.g., Grimmett & Stirzaker, *Probability and Random Processes*). We show this for the pure birth process, with $\mu_i = 0$ for all i.



Example continued

Derivation of the Forward Equation

(identifying issues of exchanging summation & limits, but not attempting to fully resolve them).

- Perhaps the main use of the forward/backward equations is to show $P(t) = e^{Qt}$, assuming the (possibly infinite-dimensional) matrix exponential exists.
- The general method of deriving a differential equation can be used to find other quantities...

Example. Let $X(t)$ count individuals in a population. Suppose each individual reporduces at rate λ, dividing into two individuals (think of bacteria). Each individual dies at rate μ. Construct an appropriate Markov model, and hence find $\mathbb{E}[X(t)]$.

Limiting probabilities, irreducibility, stationary distributions and ergodicity

- If the embedded chain $\{X_n\}$ is ergodic with transition matrix $P = [P_{ij}]$ and $\pi_i = \sum_j \pi_j P_{ji} = \lim_{n\to\infty} P^n_{ji}$ then results for semi-Markov models give

$$\boxed{P_j \stackrel{\text{def}}{=} \lim_{t\to\infty} P_{ij}(t) = \frac{\pi_j/\nu_j}{\sum_i \pi_i/\nu_i}}$$

- In this case, if $\sum_i \pi_i/\nu_i < \infty$ then $\{X(t)\}$ is said to be **ergodic**.

- $\{X(t)\}$ is **irreducible** when $\{X_n\}$ is.

- A continuous time Markov chain is a non-lattice semi-Markov model, so it has not concept of periodicity. Thus $\{X(t)\}$ can be ergodic even if $\{X_n\}$ is periodic. Dealing with this situation requires careful definition of π so we ignore it here.

- Setting $\frac{d}{dt}P(t) = 0$ in the forward equation suggests another way to calculate the stationary distribution: P_i is the unique solution to
  $\boxed{\sum_i P_i Q_{ij} = 0, \qquad \sum_i P_i = 1}$
  Writing this out in full gives
  $\nu_j P_j = \sum_{j \neq i} q_{ij} P_i$,
  which can be interpreted as "rate of leaving j" = "rate of entering j."

- If $\mathbb{P}[X(0) = j] = P_j$, i.e. the chain is started in it stationary distribution, then
  $$\frac{d}{dt}\mathbb{P}[X(t) = j] = \frac{d}{dt}\sum_i P_i P_{ij}(t) = \sum_i P_i \frac{d}{dt} P_{ij}(t)$$
  $$= \sum_{i,k} P_i Q_{ik} P_{kj}(t) = 0,$$
  i.e., $\{X(t)\}$ is then stationary.

- Note that (as for semi-Markov processes) long run time averages equal limiting probabilities.

Example: A small barbershop, operated by a single barber, has waiting room for only one customer. Potential customers arrive at a Poisson rate of 3 per hr, and each service time is independent, exponentially distributed with mean 1/4 per hr. Find

(a) the average # of customers in the shop (including customers currently being cut).

(b) the proportion of potential customers entering the shop.

Example continued

Time Reversibility in Continuous Time

- Just as for discrete time, the reversed chain (looking backwards) is a Markov chain.
- It is intuitively clear that the time spent in a visit to state i is the same looking forwards as backwards, i.e. Exponential (ν_i).
- Thus, to find the reverse chain we must only find the transition probabilities of the reversed embedded chain. If $\{X_n\}$ is stationary and ergodic, with transition matrix $P = [P_{ij}]$ and stationary distribution π, then the reverse chain has transition matrix given by

$$\boxed{P^*_{ij} = \pi_j P_{ji}/\pi_i} \tag{1}$$

This implies that the Q matrix satisfies

$$\boxed{P_i q^*_{ij} = P_i q_{ji}} \tag{2}$$

where q^*_{ij} give the infinitesimal transition probabilities for the reversed chain, and P_i is the stationary $dist^n$ of $\{X(t)\}$.

- Why are (1) and (2) equivalent?

- A stationary, ergodic Markov chain is **time reversible** if $P_i q_{ij} = P_j q_{ji}$ (3)
- Similar to the discrete time case, this means "rate of going directly from i to j"
  $=$ "rate of going direclty from j to i"
- If $\{P_i\}$ is a probability distribution satisfying (3), then $\{X(t)\}$ is reversible, with stationary distribution $\{P_i\}$.

Example (A Stochastic Network). N customers move among r servers. The service time at server i is Exponential (μ_i). Following service, a customer moves on to server $j \neq i$ with equal probability $1/(r-1)$. Let $X(t) = (X_1(t), ..., X_r(t))$ where $X_k(t)$ counts customers at server k. Customers wait in line until being served. Find the limiting distribution of $X(t)$. Hint: employ reversibility.

Solution

Solution continued