This Week's Agenda

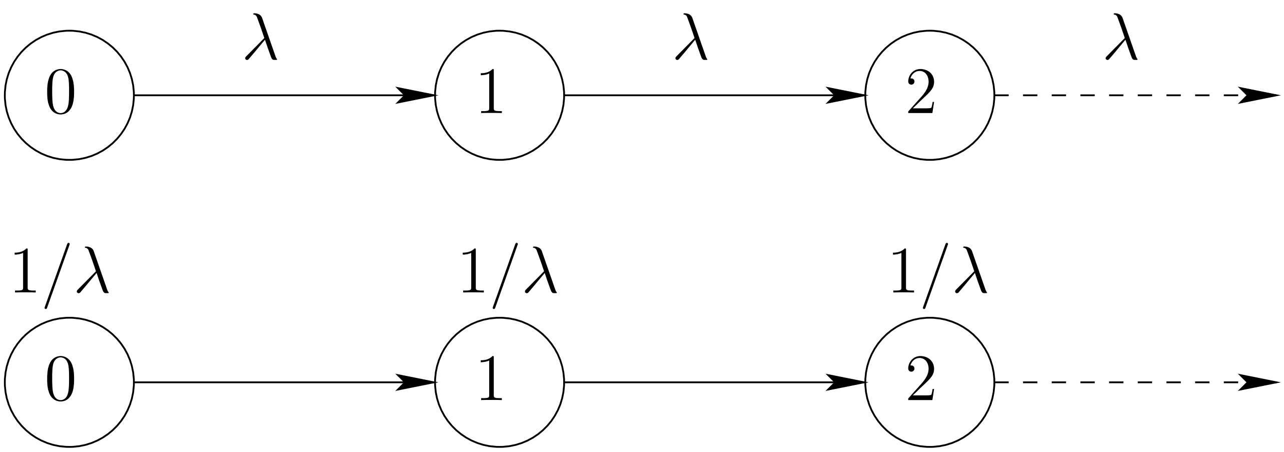
Last Time

- Exponential Distribution, Poisson Process, DTMC

Today

- Continuous Time Markov Chain

Another Look at the Poisson Process



There are two equivalent ways to look at this:

1. The process moves from state i to $i+1$ with average rate λ
2. The process stays at state i in an amount exponential(λ) (i.e. with mean $1/\lambda$), and then moves to $i+1$ with probability 1

A continuous time Markov chain (CTMC) generalizes this idea.

- The process moves around a directed graph (finite or infinite, but the number of states is countable)
- The process' staying times at different states have different exponential distributions (exponential(q_i) for state i, e.g.)
- The probability of moving from i to j is π_{ij}
- Equivalently, the "jumping rate" from i to j is $q_{ij} = q_i\pi_{ij}$

Right Continuous Continuous Time Processes

- A continuous time stochastic process $\{X_t \mid 0 \leq t < \infty\}$ is right continuous if

$$\forall t, \exists \epsilon > 0 \ : \ X_s = X_t, \ \forall s \in [t, t+\epsilon]$$

- A result from measure theory:

Theorem 1. *The probability of any event depending on a right continuous process can be determined from its* finite dimensional distributions*, i.e. from the probabilities*

$$\Pr[X_{t_1} = i_1, \dots, X_{t_n} = i_n],$$

for $n \geq 1, 0 \leq t_1 \leq \dots, \leq t_n$*, and* $i_1, \dots, i_n \in I$

- All continuous time processes we consider will be right continuous.

CTMC: First Definition

- A continuous time stochastic process $(X_t)_{t\geq 0}$ with a countable state space I is a continuous time Markov chain if there exists a given family of matrices $\{\mathbf{P}(t) = (p_{ij}(t))\}_{t\geq 0}$ such that

$$\Pr[X_{t_{n+1}} = j \mid X_{t_n} = i, X_{t_k} = i_k, 0 \leq k \leq n-1] = p_{ij}(t_{n+1} - t_n), \quad (1)$$

  for all $n \geq 0$, states $i_0, \dots, i_{n-1}, i, j$, and times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n \leq t_{n+1}$.

- Chapman-Kolmogorov equations:

$$p_{ij}(s+t) = \sum_{k\in I} p_{ik}(s)p_{kj}(t).$$

  In matrix terms, we have

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t), \ \forall s, t \geq 0. \quad (2)$$

- The set $\{P(t), t \geq 0\}$ satisfying (2) is called a semigroup

Example: Poisson Process

- Consider a Poisson process $\{X_t\}_{t\geq 0}$ with rate λ.
- For $j \leq i$, it's clear that $p_{ij}(t) = 0$
- For $j > i$, $p_{ij}(t)$ is the probability that there are $j - i$ arrivals within an amount t of time, thus

$$p_{ij}(t) = e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!}.$$

- Sanity check for Chapman-Kolmogorov

$$\begin{aligned}
\sum_{k=0}^{\infty} p_{ik}(s)p_{kj}(t) &= \sum_{k=i}^{j} p_{ik}(s)p_{kj}(t) \\
&= \sum_{k=i}^{j} e^{-\lambda s}\frac{(\lambda s)^{k-i}}{(k-i)!}e^{-\lambda t}\frac{(\lambda t)^{j-k}}{(j-k)!} \\
&= \frac{e^{-\lambda(s+t)}\lambda^{j-i}}{(j-i)!}\sum_{l=0}^{j-i}\binom{j-i}{l}s^l t^{j-i-l} \\
&= p_{ij}(s+t).
\end{aligned}$$

CTMC: towards the second definition

We will be needing the following concepts associated with a continuous time stochastic process $(X_t)_{t \geq 0}$:

- Jump times
- Holding times
- Explosion and explosion time
- Jump process and jump chain
- Minimal (and right-continuous) process
- Q-matrix
- Jump matrix

Why do we need the second definition? In many cases

- it is more intuitive
- it is easier to construct the CTMC using this definition
- it contains a discrete (and often finite) set of parameters specifying the CTMC
- the two definitions are equivalent

Jump, Holding, Explosion Times

Given $(X_t)_{t\geq 0}$

- The jump times $J_0, J_1, \dots$ are defined by

$$J_0 = 0,\ J_n = \inf\{t \geq J_{n-1} \mid X_t \neq X_{J_{n-1}}\},\ n \geq 1.$$

- The holding times $S_0, S_1, \dots$ are defined by

$$S_n = \begin{cases} J_{n+1} - J_n & J_n < \infty \\ \infty & \text{otherwise} \end{cases}$$

- The explosion time is

$$\xi = \sup_n J_n = \sum_{n=0}^{\infty} S_n.$$

  When $\xi < \infty$, the process makes infinitely many jumps in a finite amount of time: not desirable.

- The jump process $(Y_n)n \geq 0$ is defined by $Y_n = X_{J_n}$. If $(Y_n)_{n\geq 0}$ is a DTMC, then it is called the jump chain or embedded chain of $(X_t)_{t\geq 0}$.

- $(X_t)_{t\geq 0}$ is minimal when we require $X_t = \infty$ for $t \geq \xi$.

Q-matrices

- Let I be a countable set
- A Q-matrix on I is a matrix $\mathbf{Q} = (q_{ij} \ : \ i, j \in I)$ satisfying
  (a) $0 \leq -q_{ii} < \infty$, for all $i \in I$
  (b) $q_{ij} \geq 0$, for all $i, j \in I$
  (c) $\sum_{j \in I} q_{ij} = 0$
- For convenience, define $q_i = -q_{ii} \geq 0$
- Example

$$\mathbf{Q} = \begin{pmatrix} -3 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & -5 & 2 \\ 0 & 2 & 2 & -4 \end{pmatrix}$$

Later on,

- q_{ij} will be interpreted as the rate of jumping from i to j
- $\mathbf{Q}$ will be called the (infinitesimal) generator matrix for a CTMC

Jump Matrix

- Given a Q-matrix $\mathbf{Q} = (q_{ij})$, define the jump matrix $\mathbf{\Pi}$:

$$\pi_{ij} = \begin{cases} q_{ij}/q_i & j \neq i,\ q_i \neq 0 \\ 0 & j \neq i,\ q_i = 0 \end{cases}$$

$$\pi_{ii} = \begin{cases} 0 & q_i \neq 0 \\ 1 & q_i = 0 \end{cases}$$

- Example, jump matrix of previous Q-matrix

$$\mathbf{Q} = \begin{pmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ 0 & \frac{3}{5} & 0 & \frac{2}{5} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

CTMC: Second Definition

A (minimal right-continuous) stochastic process $(X_t)_{t \geq 0}$ is a continuous time Markov chain if there is a Q-matrix $\mathbf{Q} = (q_{ij})$ with corresponding jump matrix $\mathbf{\Pi}$ such that

(i) The jump process $(Y_n)_{n \geq 0}$ of $(X_t)_{t \geq 0}$ is a discrete-time Markov chain with transition probability matrix $\mathbf{\Pi}$

(ii) For any $n \geq 0$, conditional on $Y_0, \dots, Y_n$, the holding times $S_0, \dots, S_n$ are independent exponential random variables with parameters $q_{Y_0}, \dots, q_{Y_n}$, respectively

Example: the Poisson process with rate λ can be defined with the following Q-matrix

$$\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda & 0 & \dots \\ 0 & -\lambda & \lambda & \dots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Relationship Between Two Definitions

- Think of them as two different *constructions* of a CTMC
- Relationships between $\{\mathbf{P}(t)\}_{t\geq 0}$ and $\mathbf{Q}$:

  (a) $\{\mathbf{P}(t)\}_{t\geq 0}$ is the minimal non-negative solution to the forward equation

  $$\mathbf{P}'(t) = \mathbf{QP}(t), \;\; \mathbf{P}(0) = \mathbf{I},$$

  namely,

  $$p'_{ij}(t) = \sum_{k\in I} q_{ik}p_{kj}(t), \; p_{ij}(0) = \delta_{ij}.$$

  (b) $\{\mathbf{P}(t)\}_{t\geq 0}$ is the minimal non-negative solution to the backward equation

  $$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}, \;\; \mathbf{P}(0) = \mathbf{I},$$

  namely,

  $$p'_{ij}(t) = \sum_{k\in I} p_{ik}(t)q_{kj}, \; p_{ij}(0) = \delta_{ij}.$$

The *Kronecker delta* $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$

Getting the $\mathrm{P}(t)$ from Q

- To get the $\mathbf{P}(t)$ from $\mathbf{Q}$, solve either the forward equation or the backward equation
- If I is finite, the following solution is the **unique** solution to the forward and backward equations (i.e., even among the non-minimal, negative ones)

$$\mathbf{P}(t) = e^{t\mathbf{Q}} = \sum_{n=0}^{\infty} \mathbf{Q}^n \frac{t^n}{n!}.$$

Getting Q from the $\mathrm{P}(t)$

$$\mathbf{P}'(0) = \mathbf{P}(0)\mathbf{Q} = \mathbf{Q}.$$

In other words,

$$q_{ij} = p'_{ij}(0), \text{ for all } i, j.$$

Class Structure

- $i \rightsquigarrow j$ if $\Pr_i(\exists t \geq 0, X_t = j) > 0$
- i communicate with j, i.e. $i \leftrightarrow j$, if $i \rightsquigarrow j$ and $j \rightsquigarrow i$
- communicating class, closed class, irreducibility are defined similarly

Theorem 2. *Given states $i \neq j$, the following are equivalent*

(1) $i \rightsquigarrow j$

(2) $i \rightsquigarrow j$ *in the jump chain* $\mathbf{\Pi}$

(3) $q_{ii_1} q_{i_1 i_2} \cdots q_{i_n j} > 0$ *for some states* $i_1, \dots, i_n$, $n \geq 0$

(4) $p_{ij}(t) > 0$ *for all* $t > 0$

(5) $p_{ij}(t) > 0$ *for some* $t > 0$

Recurrence and Transience

- i is recurrent if

$$\Pr[\{t : X_t = i\}\text{is unbounded}] = 1$$

- i is positive recurrent iff the expected return time to i is finite ($\mu_{ii} < \infty$), and i is null recurrent otherwise
- i is transient if

$$\Pr[\{t : X_t = i\}\text{is unbounded}] = 0$$

Theorem 3. *We have*

(1) i is recurrent if and only if i is recurrent for $\mathbf{\Pi}$

(2) every state is either recurrent or transient

(3) (positive/null) recurrence and transience are class properties

Stationary Distributions

A distribution (resp., measure) λ on I is a stationary (or invariant) distribution (resp., measure) if $\lambda \mathbf{P}(t) = \lambda,\ \forall t > 0$

Theorem 4. *λ is stationary iff $\lambda \mathbf{Q} = 0$ (this holds for both measure and distribution cases). Specifically, λ is stationary iff*

$$\sum_{i \neq j} \lambda_i q_{ij} = \lambda_j q_j = \lambda_j \sum_{k \neq j} q_{jk} \ \ \forall j \in I. \tag{3}$$

This is the balanced equation for j. Moreover, λ is a stationary measure iff $\tau \mathbf{\Pi} = \tau$, where $\tau_i = \lambda_i q_i, \forall i \in I$.

Theorem 5. *Let $\mathbf{Q}$ be irreducible, then $\mathbf{Q}$ is non-explosive and has an invariant distribution if and only if $\mathbf{Q}$ is positive recurrent. Moreover, when $\mathbf{Q}$ is positive recurrent with invariant distribution λ, we have $\mu_{ii} = \frac{1}{\lambda_i q_i}, \forall i \in I$.*

Detailed Balance Condition

Theorem 6. *If λ satisfies the detailed balance condition*

$$\lambda_k q_{kj} = \lambda_j q_{jk},\ \forall j, k \in I, \tag{4}$$

then λ is invariant.

Convergence to Equilibrium

Theorem 7. *Let* $\mathbf{Q}$ *be irreducible, non-explosive with an invariant distribution* λ*. Then,*

$$\lim_{t\to\infty} p_{ij}(t) = \lambda_j, \ \forall j \in I.$$

In particular, the invariant distribution is unique.

Ergodic Theorem

Theorem 8. *Let* $\mathbf{Q}$ *be irreducible, then*

$$\Pr\left[\lim_{t\to\infty} \frac{1}{t}\int_0^t 1_{\{X_s=i\}} ds = \frac{1}{\mu_{ii} q_i}\right] = 1$$

Moreover, if $\mathbf{Q}$ *is positive recurrent, implying* $\mathbf{Q}$ *has a unique invariant distribution* λ*, then for any bounded function* $f : I \to \mathbb{R}$*,*

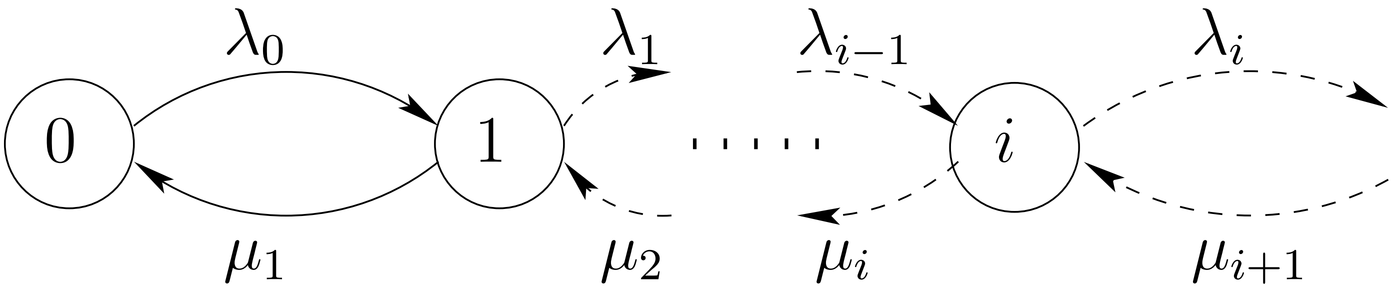
$$\Pr\left[\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X_s) ds = \bar{f}\right] = 1,$$

where

$$\bar{f} = \sum_{i\in I} \pi_i f_i.$$

Birth and Death Process (BDP)

A birth and death process can be illustrated as follows.



- The $\{\lambda_i\}_{i\geq 0}$ are called the birth rates
- The $\{\mu_i\}_{i\geq 0}$ are called the death rates
- The Poisson process is a special case of this process
- We often think of a state i as the number of "items" in a system
  - Items entering the system having i items according to a Poisson process with rate λ_i
  - Items leaving the system having i items according to a Poisson process with rate μ_i, independent from the entering items

BDP: Condition for Positive Recurrence

- The jump chain is the birth and death chain we've discussed, where

$$\begin{aligned} a_0 &= 1 \\ a_i &= \frac{\lambda_i}{\lambda_i + \mu_i}, \; i \geq 1 \\ b_i &= \frac{\mu_i}{\lambda_i + \mu_i}, \; i \geq 1 \end{aligned}$$

- The chain is current iff

$$\sum_{n=1}^{\infty} \frac{\mu_1 \cdots \mu_n}{\lambda_1 \cdots \lambda_n} = \infty$$

- The chain is also positive recurrent if, additionally,

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < \infty$$

BDP: Stationary Distribution

The *balance equation* for a measure τ to be stationary is

$$\begin{aligned} \tau_1\mu_1 &= \tau_0\lambda_0 \\ \tau_{i+1}\mu_{i+1} + \tau_{i-1}\lambda_{i-1} &= \tau_i(\lambda_i + \mu_i),\ i \geq 1 \end{aligned}$$

Solving this, we get

$$\tau_n = \tau_0 \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}$$

For τ to be an invariant distribution, we further need $\sum_{n=0}^{\infty} \tau_n = 1$, which is solvable if

$$C = 1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < \infty.$$

This is true when the chain is positive recurrent! In conclusion,

$$\begin{aligned} \tau_0 &= \frac{1}{C} \\ \tau_n &= \frac{1}{C} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}, \quad n \geq 1 \end{aligned}$$

Ergodic theorem: as time tends to ∞, fraction of time spent in state i is $\tau_i(\lambda_i + \mu_i)$.