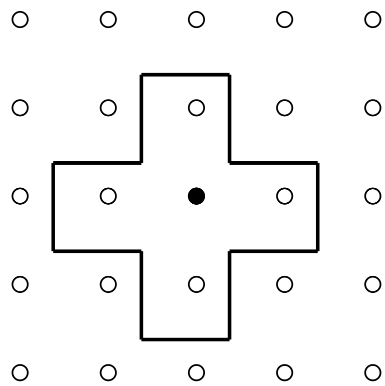
4. Markov Chains, Part (ii)

Example: (spatial models on a lattice)

Let $\{Y_{ij}, 1 \le i \le m, 1 \le j \le n\}$ be a spatial stochastic process. Let $\mathcal{N}(i,j)$ define a neighborhood , e.g. $\mathcal{N}(i,j) = \{(p,q) : |p-i| + |q-j| = 1\}$



- We may model the conditional distribution of Y_{ij} given the neighbors. E.g., if your neighbors vote Republican (R), you are more likely to do so. Say,

$$\mathbb{P}\big[Y_{ij} = R \mid \{Y_{pq}, (p,q) \neq (i,j)\}\big] = \tfrac{1}{2} + \alpha\Big[\Big(\textstyle\sum_{(p,q)\in\mathcal{N}(i,j)} I_{\{Y_{pq}=R\}}\Big) - 2\Big].$$

- The full distribution π of Y is, in principle, specified by the conditional distributions. In practice, one simulates from π using MCMC, where the proposal distribution is to pick a random (i,j) and swap it from R to D or vice versa.

Example: Sampling conditional distributions.

- If X and Θ have joint density $f_{X\Theta}(x, \theta)$ and we observe $X = x$, then one wants to sample from $f_{\Theta|X}(\theta \,|\, x) = \frac{f_{X\Theta}(x,\theta)}{f_X(x)} \propto f_\Theta(\theta) f_{X|\Theta}(x \,|\, \theta)$
- For **Bayesian inference**, $f_\Theta(\theta)$ is the **prior** distribution, and $f_{\Theta|X}(\theta \,|\, x)$ is the **posterior**. The model determines $f_\Theta(\theta)$ and $f_{X|\Theta}(x \,|\, \theta)$. The normalizing constant, $f_X(x) = \int f_{X\Theta}(x, \theta)\, d\theta$, is often unknown.
- Using MCMC to sample fro the posterior allows numerical evaluation of the **posterior mean** $\mathbb{E}\big[\Theta \,|\, X = x\big]$, **posterior variance** $\mathrm{Var}\big(\Theta \,|\, X = x\big)$, or a $1 - \alpha$ **credible region** defined to be a set A such that $\mathbb{P}\big[\Theta \in A \,|\, X = x\big] = 1 - \alpha$.

Galton-Watson Branching Process

- Let X_n be the size of a population at time n. Suppose each individual has an iid offspring distribution, so $X_{n+1} = \sum_{k=1}^{X_n} Z_{n+1}^{(k)}$ where $Z_{n+1}^{(k)} \sim Z$.
- We can think of $Z_{n+1}^{(k)} = 0$ being the death of the parent, or think of X_n as the size of the n^{th} generation.
- Suppose $X_0 = 1$. Notice that $X_n = 0$ is an absorbing state, termed **extinction**. Supposing $\mathbb{P}[Z = 0] > 0$ and $\mathbb{P}[Z > 1] > 0$, we can show that $1, 2, 3, \ldots$ are transient states. Why?

- Thus, the branching process either becomes extinct or tends to infinity.

- Set $\phi(s) = \mathbb{E}[s^Z]$, the **probability generating function** for Z. Set $\phi_n(s) = \mathbb{E}[s^{X_n}]$. Show that $\phi_n(s) = \phi^{(n)}(s) = \phi(\phi(\dots\phi(s)\dots))$, meaning ϕ applied n times.

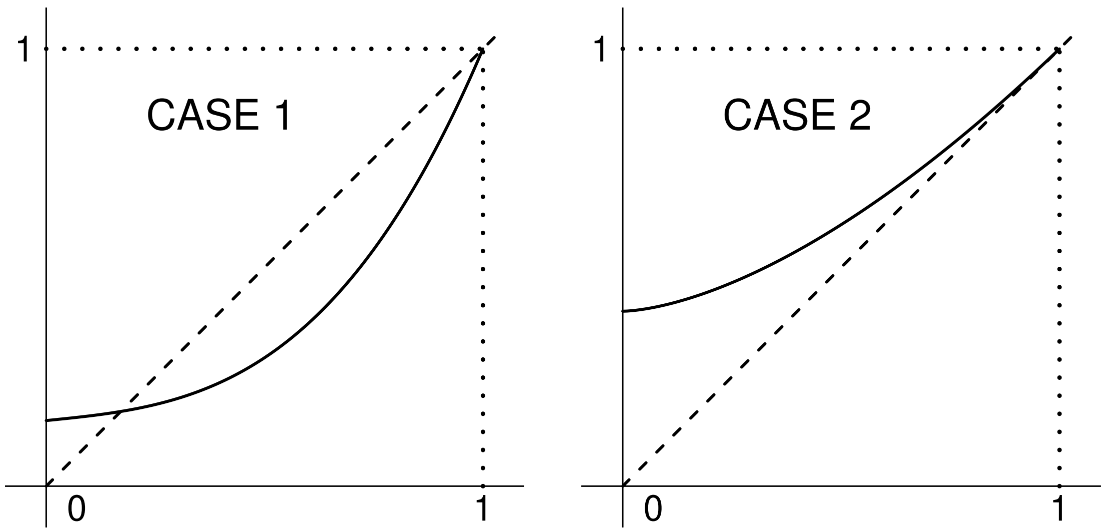
- If we can find a solution to $\phi(s) = s$, we will have $\phi_n(s) = s$ for all n. This suggests plotting $\phi(s)$ vs s, noting that

(i) $\phi(1) = 1$. Why?

(ii) $\phi(0) = \mathbb{P}[Z=0]$. Why?

(iii) $\phi(s)$ is **increasing**, i.e., $\frac{d\phi}{ds} > 0$. Why?

(iv) $\phi(s)$ is **convex**, i.e., $\frac{d^2\phi}{ds^2} > 0$. Why?



- Now, notice that, for $0 < s < 1$,

$$\mathbb{P}\big[\text{extinction}\big] = \lim_{n\to\infty} \mathbb{P}\big[X_n = 0\big] = \lim_{n\to\infty} \phi_n(s).$$

- Argue that, in CASE 1, $\lim_{n\to\infty} \phi_n(s)$ is the unique fixed point $\phi(s) = s$ for $0 < s < 1$. In CASE 2, $\lim_{n\to\infty} \phi_n(s) = 1$.

- Conclude that in CASE 1 (with $\frac{d\phi}{ds}\mid_{s=1} > 1$, i.e. $\mathbb{E}[Z] > 1$) there is some possibility of infinite growth. In CASE 2 (with $\frac{d\phi}{ds}\mid_{s=1} \leq 1$, i.e. $\mathbb{E}[Z] \leq 1$) extinction is assured.

- Example: take

$$Z = \begin{cases} 0 & w.p. & 1/4 \\ 1 & w.p. & 1/4 \\ 2 & w.p. & 1/2 \end{cases} .$$

Find the chance of extinction.

- Now suppose the founding population has size k (i.e., $X_0 = k$). Find the chance of extinction.

Time Reversal

- Thinking backwards in time can be a powerful strategy. For any Markov chain $\{X_k, k = 0, 1, \ldots, n\}$, we can set $Y_k = X_{n-k}$. Then Y_k is a Markov chain. Suppose X_k has transition matrix p_{ij}, then Y_k has inhomogeneous transition matrix $P^{*[k]}_{ij}$, where

$$\begin{aligned} P^{*[k]}_{ij} &= \mathbb{P}\big[Y_{k+1}{=}j \,|\, Y_k{=}i\big] \\ &= \mathbb{P}\big[X_{n-k-1}{=}j \,|\, X_{n-k}{=}i\big] \\ &= \frac{\mathbb{P}\big[X_{n-k}{=}i \,|\, X_{n-k-1}{=}j\big]\,\mathbb{P}\big[X_{n-k-1}{=}j\big]}{\mathbb{P}\big[X_{n-k}{=}i\big]} \\ &= P_{ji}\,\mathbb{P}\big[X_{n-k-1}{=}j\big] \,/\, \mathbb{P}\big[X_{n-k}{=}i\big]. \end{aligned}$$

- If $\{X_k\}$ is **stationary**, with stationary distribution π, then $\{Y_k\}$ is homogeneous, with

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

- Surprisingly many important chains have the **time-reversibility** property $P^*_{ij} = P_{ij}$, for which the chain looks the same forwards as backwards.

Theorem. (detailed balance equations)
If π is a probability distribution with $\pi_i P_{ij} = \pi_j P_{ji}$ and P is irreducible, then P is time-reversible and has stationary distribution π.

Proof

- The detailed balance equations are much simpler than the general equations for finding π. It is sometimes worth trying to solve them when looking for π, in case you get lucky!
- Intuition: $\pi_i P_{ij}$ is the "rate of going directly from i to j." A chain is reversible if this equals "rate of going directly from j to i."

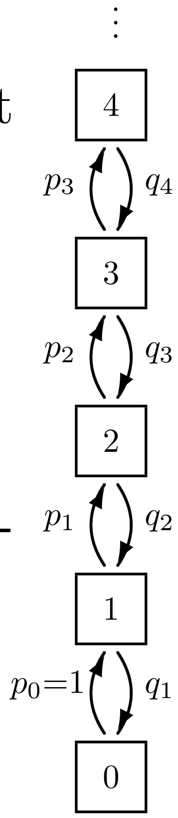
- An equivalent condition for reversibility is "every loop of states starting and ending in i has the same probability as the reverse loop," i.e., for all $i_1, \ldots, i_k$

$$P_{ii_1} P_{i,i_2} \ldots P_{i_{k-1} i_k} P_{i_k i} = P_{ii_k} P_{i_k i_{k-1}} \ldots P_{i_2 i_1} P_{i_1 i}$$

<u>Proof</u>

Example: Birth-Death Process.

Let $\mathbb{P}\big[X_n = X_n + 1 \,|\, X_n = i\big] = p_i$ and $\mathbb{P}\big[X_n = X_n - 1 \,|\, X_n = i\big] = q_i = 1 - p_i$.



- If $\{X_n\}$ is positive recurrent, it must be reversible. Heuristic reason:

- Find the stationary distribution giving conditions for its existence.

- Note that this chain is periodic ($d = 2$). There is still a unique stationary distribution which solves the detailed balance equations.

Example: Metropolis Algorithm. The Metropolis-Hastings algorithm with symmetric proposals ($q_{ij} = q_{ji}$) is the **Metropolis algorithm**. Here,

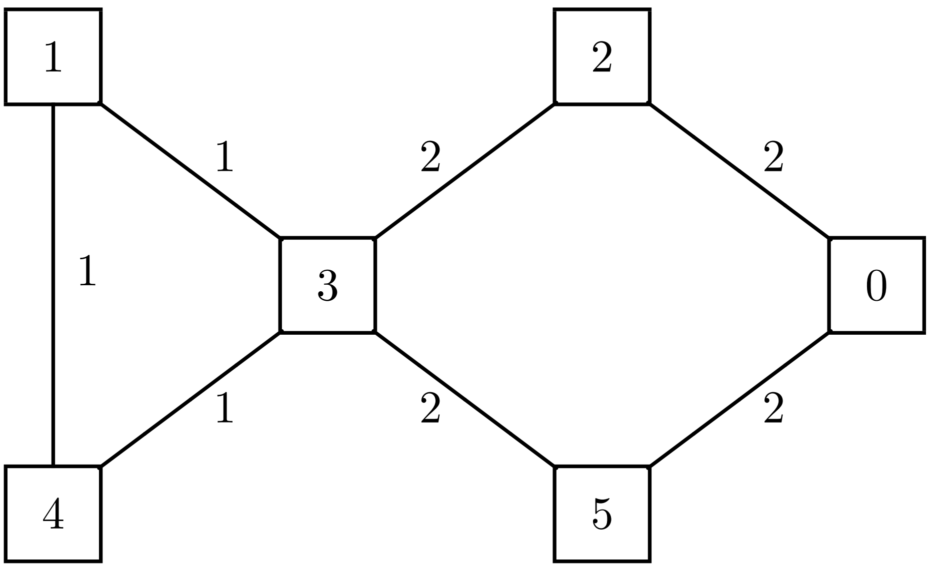
$$P_{ij} = \begin{cases} q_{ij} \min(1, \pi_j/\pi_i) & j \neq i \\ q_{ii} + \sum_{k \neq i} q_{ik} \{1 - \min(1, \pi_k/\pi_i)\} & j = i \end{cases}$$

- Show that P_{ij} is reversible, with stationary distribution π.

  Solution

Random Walk on a Graph

- Consider **undirected**, **connected** graph with a positive **weight** w_{ij} on each **edge** ij. Set $w_{ij} = 0$ if i & j are not connected by an edge.
- Set $P_{ij} = \dfrac{w_{ij}}{\sum_k w_{ik}}$
- The Markov chain $\{X_n\}$ with transition matrix P is called a random walk on a graph. Think of a driver who is lost: each vertex is an intersection, the weight is the width (in lanes) of a road leaving an intersection, the driver picks a random exit when he/she arrives at an intersection but has a preference for bigger streets.



- Show that the random walk on a graph is reversible, and has stationary distribution $\pi_i = \sum_j w_{ij} \big/ \sum_{jk} w_{jk}$
  Note: the double sum counts each weight twice.

- Hence, find the limiting probability of being at vertex 3 on the graph shown above.

Ergodicity

- A stochastic process $\{X_n\}$ is **ergodic** if limiting time averages equal limiting probabilities, i.e.,

$$\lim_{n\to\infty} \frac{\text{time in state } i \text{ up to time } n}{n} = \lim_{n\to\infty} \mathbb{P}[X_n = i].$$

- Show that an irreducible, aperiodic, positive recurrent Markov chain is ergodic (i.e., this general idea of ergodicity matches our definition for Markov chains).

Semi-Markov Processes

- A **Semi-Markov process** is a discrete state, continuous time process $\{Z(t), t \geq 0\}$ for which the sequence of states visited is a Markov chain $\{X_n, n = 0, 1, \ldots\}$ and, conditional on $\{X_n\}$, the times between transitions are independent. Specifically, define transition times $S_0, S_1, \ldots$ by $S_0 = 0$ and $S_n - S_{n-1} \sim F_{ij}$ conditional on $X_{n-1} = i$ and $X_n = j$. Then, $Z(t) = X_{N(t)}$ where $N(t) = \sup\{n : S_n \leq t\}$.

- $\{X_n\}$ is the **embedded Markov chain**.

- The special case where $F_{ij} \sim \text{Exponential}\,(\nu_i)$ is called a **continuous time Markov chain**

- The special case when

$$F_{ij}(t) = \begin{cases} 0, & t < 1 \\ 1, & t \geq 1 \end{cases}$$

results in $Z(n) = X_n$, so we retrieve the discrete time Markov chain.

Example: A taxi serves the airport, the hotel district and the theater district. Let $Z(t) = 0, 1, 2$ if the taxi's current destination is each of these locations respectively. The time to travel from i to j has distribution F_{ij}. Upon arrival at i, the taxi immediately picks up a new customer whose destination has distribution given by P_{ij}. Then, $Z(t)$ is a semi-Markov process.

- To study the limiting behavior of $Z(t)$ we make some definitions...
- Say $\{Z(t)\}$ is **irreducible** if $\{X_n\}$ is irreducible
- Say $\{Z(t)\}$ is **non-lattice** if $\{N(t)\}$ is non-lattice
- Set H_i to be the c.d.f. of the time in state i before a transition, so $H_i(t) = \sum_j P_{ij} F_{ij}(t)$ and the expected time of a visit to i is $\mu_i = \int_0^\infty x dH_i(x) = \sum_j P_{ij} \int_0^\infty x\, dF_{ij}(x)$.
- Let T_{ii} be the time between successive visits to i and define $\mu_{ii} = \mathbb{E}\big[T_{ii}\big]$.

Proposition: If $\{Z(t)\}$ is an irreducible, non-lattice semi-Markov process with $\mu_{ii} < \infty$, then

$$\lim_{t\to\infty} \mathbb{P}\big[Z(t)=i \mid Z(0)=j\big] = \frac{\mu_i}{\mu_{ii}}$$

$$= \lim_{t\to\infty} \frac{\text{time in } i \text{ before } t}{t}$$

Proof: identify a relevant alternating renewal process.

- By counting up time spent in i differently, we get another identity

$$\lim_{t\to\infty} \frac{\text{time in } i \text{ during } [0,t]}{t} = \frac{\pi_i \mu_i}{\sum_j \pi_j \mu_j}$$

supposing that $\{X_n\}$ is ergodic (irreducible, aperiodic, positive recurrent) with stationary distribution π.

Proof

- Calculations with semi-Markov models typically involve identifying a suitable renewal, alternating renewal, regenerative or reward process.

<u>Example</u> Let $Y(t) = S_{N(t)+1} - t$ be the residual life process for an ergodic, non-lattice semi-Markov model. Find an expression for $\lim_{t\to\infty} \mathbb{P}\big[Y(t) > x\big]$.

Hint: Consider an alternating renewal process which switches "on" when $Z(t)$ enters i, and switches "off" immediately if the next transition is not j or switches "off" once the time until the transition to j is less than x.

Proof (continued)