Plan Group Work

0. The title says it all

   Next Time: MCMC and General-state Markov Chains

   Midterm Exam: Tuesday 28 March in class
   Homework 4 due Thursday

Unless otherwise noted, let X be an irreducible, aperiodic Markov chain on countable state-space $\mathcal{S}$ with initial distribution μ and transition probabilities P.

You are free to use the Reference items at the end. Note in particular Reference 11 which strengthens the condition we saw last time.

Example 1. Define a Markov chain on $\{1, \ldots, 6\}$ with transition probability matrix

$$P = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0.25 & 0.5 & 0.25 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0.75 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.25 & 0.25 & 0 & 0.5 \end{bmatrix} \tag{1}$$

Tell me everything you can about this chain.

Example 2. Define a Markov chain on $\{1, \ldots, 6\}$ with transition probability matrix

$$P = \begin{bmatrix} 0 & 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.75 \\ 0 & 0 & 0 & 0 & 0.25 & 0.75 \\ 0.8 & 0.2 & 0 & 0 & 0 & 0 \\ 0.8 & 0.2 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2}$$

Tell me everything you can about this chain.

Example 3. A Queueing Example

Let $(u_k)_{k \geq 0}$ be a non-negative sequence with $\sum_k u_k = 1$. Let Z_n, for $n \geq 0$, denote the length of a queue at time n, where at each time, one customer arrives *and* k customers are served with probability u_k, if there are at least k in the queue.

Then Z is a Markov chain on $\mathbb{Z}_{\oplus}$.

(a) Find its transition probabilities.
(b) Decompose the state space
(c) Find the period of each communicating class.
(d) Find conditions for recurrence or transience of the chain.

Example 4. Consider a Markov chain with two states and transition probabilities

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix} \tag{3}$$

Show by induction that

$$P^n = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}(2p-1)^n & \frac{1}{2} - \frac{1}{2}(2p-1)^n \\ \frac{1}{2} - \frac{1}{2}(2p-1)^n & \frac{1}{2} + \frac{1}{2}(2p-1)^n \end{bmatrix} \tag{4}$$

What's the limiting distribution of this chain?

Example 5.

Consider the following model for the diffusion of gas. Suppose that M molecules are distributed between two chambers that are separated by a permeable boundary. At each time, one of the molecules – with all equally likely to be chosen – crosses the boundary from one chamber to the other.

Tell me everything you can about this process.

Example 6.

Consider non-negative sequences p_i, q_i, and r_i for $i \ge 0$ satisfying

$$q_0 = 0 \tag{5}$$

$$r_0 + p_0 = 1 \tag{6}$$

$$q_i + r_i + p_i = 1 \quad \text{for } i > 0 \tag{7}$$

Define a random walk on the integers with transition matrix

$$P(i, i-1) = q_i \qquad P(i,i) = r_i \qquad P(i, i+1) = p_i, \tag{8}$$

and all other entries zero.

Let $\alpha_0 = 1$ and $\alpha_n = \dfrac{q_1 \cdots q_n}{p_1 \cdots p_n}$. Show that the chain is transient if $\sum_n \alpha_n < \infty$ and recurrent if $\sum_n \alpha_n = \infty$.

As an optional addition, show that the chain is positive recurrent if $\sum_n p_0/(\alpha_n p_n) < \infty$.

Example 7. A spider is hunting a fly, and the fly is trying to survive. The spider starts in location 1 and moves between locations 1 and 2 according to the Markov transitions

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}. \tag{9}$$

The fly starts in location 2 and moves between the locations with transitions

$$P = \begin{bmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{bmatrix}. \tag{10}$$

The hunt ends if the two ever land on the same location, in which case the fly is eaten.

Show that this progress of the hunt can be described (except for knowing at which location the hunt ends) by a three-state Markov chain. Find the transition probabilities for this chain. What is the expected duration of the hunt?

Example 8.

Consider a symmetric random walk in d dimensions.

The drift operator $\Delta = PV - V$ for this chain is a normalized (by $1/d$) version of an operator called the *discrete Laplacian* because on finer and finer grids, it approximates the standard Laplacian differential operator $\Delta = \nabla \cdot \nabla = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$. Thus, the behavior of random walks on Euclidean space connects quite deeply to the solutions of differential equations.

For example, we've seen that the existence of bounded functions V with ΔV satisfying certain inequalities corresponds to recurrence or transience of the chain. As another example, consider a region $S_0 \subset \mathcal{S}$ with boundary points ∂S_0. If f is a function on S_0 such that $\Delta f = 0$ on $S_0 - \partial S_0$ and $f = g$ on ∂S_0, then $f(s) = \mathsf{E}_s g(X_{T_{\partial S_0}})$.

So for example, in the one-dimensional symmetric random walk. Let $S_0 = \{0, \ldots, m\}$ and let $g(0) = 1$ and $g(m) = 0$. Then, the solution f gives the Gambler's ruin probabilities for every initial wealth. But solutions are just of the form

$$f(s) = 1 - \frac{s}{m} \tag{11}$$

by direct calculation.

(a) What is Δ for the 1 dimensional random walk? Express it in terms that are reminiscent of the continuous Laplacian describd above. See Reference 13 to check your answer, but don't look before trying it.

Note that we can think of Δ as an operator or as a matrix. What does the matrix look like in the one-dimensional case?

(b) In the Gambler's ruin example above, we computed the solution by inspection because a linear function has zero second differences. But we can use the theorem in Reference 14 as well.

Show that

$$\phi_k(j) = \sqrt{\frac{2}{m}} \sin \frac{\pi j k}{m} \tag{12}$$

is an eigenfunction (of the operator, or equivalently eigenvector of the matrix form) with eigenvalue $\lambda_k = 1 - \cos \pi k/m$.

Then, using the equation in the theorem, we get we get

$$f(j) = \frac{1}{m} \sum_{k=1}^{m-1} \frac{\sin \frac{\pi k}{m} \sin \frac{\pi k j}{m}}{1 - \cos \frac{\pi k}{m}} = 1 - \frac{j}{m}, \tag{13}$$

as we found above.

(c) Now, let $d = 2$ and let $S_0 \subset \mathcal{S}$ be a finite set with boundary points ∂S_0. We want to show here that if f is a function on S_0 such that $\Delta f = 0$ on $\text{Interior}(S_0) \equiv S_0 - \partial S_0$ and $f = g$ on ∂S_0, then $f(s) = \mathsf{E}_s g(X_{T_{\partial S_0}})$. We'll do this in several steps.

i. Draw a picture of a region S_0 satisfying the conditions above.

ii. If $s \in S_0$, then there are two possibilities, either the chain stays in $\text{Interior}(S_0)$ forever or it hits the boundary ∂S_0. Show that the latter has probability 1.

iii. Let $U = g(X_{T_{\partial S_0}})$. For any $s \in \mathcal{S}$, let s_N, s_E, s_W, s_S denote the neighbors of s (to the "north", "east", "west", and "south", respectively). By conditioning on the first step, show that

$$\mathsf{E}_s U = \frac{1}{4}\mathsf{E}_{s_N} U + \frac{1}{4}\mathsf{E}_{s_E} U + \frac{1}{4}\mathsf{E}_{s_S} U + \frac{1}{4}\mathsf{E}_{s_W} U. \tag{14}$$

iv. Show f, defined above, satisfies $\Delta f = 0$ on $\text{Interior}(S_0)$ and $f = g$ on ∂S_0.

v. What can you say about the uniqueness of this solution?

Example 9. Storage Model

Consider a storage system (dam, warehouse, insurance policy) that receives inputs at random times but otherwise drains at a regular rate.

Let $T_0 = 0$ and let the rest of the T_is be IID $\mathbb{Z}_{\oplus}$-valued with CDF G. These are inter-arrival times for the inputs to our storage system. Let the S_ns be IID $\mathbb{Z}_{\oplus}$ with CDF H. These are the amounts input at the time $Z_n = T_0 + \cdots + T_n$. Assume that the S_ns and T_ns are independent of each other as well. Suppose also that the storage system "drains" or outputs at rate r between inputs.

Define a process $(V_n)_{n\geq 0}$ by

$$V_{n+1} = (V_n + S_n - rT_{n+1})_+. \tag{15}$$

Here, V_n represents the contents of the storage system just before the nth input (that is, at time Z_n-).

Is V a Markov chain?

What is the structure of the transition probabilities?

What can we say about the long-run behavior of the chain?

What is special about the state $\{0\}$?

How might we generalize this model to make it more realistic?

Reference 10.
Let D denote the set of states in non-absorbing communicating classes and $C_1, \ldots, C_m$ be the absorbing communicating classes in the usual state decomposition.

Let $h_r(s) = H(s, C_r)$ for $i \in \mathcal{S}$. By conditioning on the first step, we get for $i \in D$ that

$$h_r(s) = \sum_{s' \in C_r} P(s, s') + \sum_{u \in D} P(s, u) h_r(u). \tag{16}$$

Hence, h_r is a non-negative solution to this equation. It can be shown (by induction on time n) that h_r is the *smallest* non-negative solution of this equation.

Let $s \in D$ and $s' \in C_r$ for some r. Then,

- If C_r is transient or null recurrent, $\lim_{n\to\infty} P^n(s, s') = 0$.
- If C_r is positive recurrent and aperiodic, $\lim_{n\to\infty} P^n(s, s') = \pi_r(s') h_r(s)$, where π_r is the stationary distribution for the chain restricted to C_r.
- If C_r is positive recurrent and periodic with period d_r

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n} P^n(s, s') = \pi_r(s') h_r(s) \tag{17}$$

as above, but $P^n(s, s')$ itself does not converge.

Reference 11. Another Drift Criterion for Transience

Let Z be an irreducible, countable-state Markov chain on $\mathcal{S}$ with drift operator Δ. Then, Z is transient if and only if there exists an $s_0 \in \mathcal{S}$ and a bounded, nonconstant function V such that

$$\Delta V(s) = 0 \qquad \text{for } s \neq s_0. \tag{18}$$

Reference 12. Random Walks

We've already seen that the one-dimensional random walk is recurrent. To see that it's null recurrent, note that the transition probabilities are doubly stochastic. (Why?) So, $\rho = \rho P$ has a solution $\rho(i) \equiv 1$. This is not normalizable, so the chain is null recurrent.

Let's consider the case of a symmetric, two dimensional random walk which moves up, down, left, or right with probability 1/4 each. The chain is irreducible because we can find a non-zero probability path from any one state to another. It is also has period 2 for the same reason as in the one-dimensional case.

Thus, to assess the recurrence of the chain, we can look at only one point, let's say 0. We know, then, that $P^{2n+1}(0,0) = 0$ for all $n \geq 0$ and, by considering all paths two and from zero, we get that

$$P^{2n}(0,0) = \sum_{\substack{i,j \\ i+j=n}} \frac{(2n)!}{i!j!i!j!} \left(\frac{1}{4}\right)^{2n}. \tag{19}$$

Why?

Our goal is to compute $\sum_m P^m(0,0)$ to test recurrence for the chain. By multiplying $P^{2n}(0,0)$ by $(n!)^2/(n!)^2$, we get

$$\sum_m P^m(0,0) = \sum_n P^{2n}(0,0) \tag{20}$$

$$= \sum_n \sum_{\substack{i,j \\ i+j=n}} \frac{(2n)!}{i!j!i!j!} \left(\frac{1}{4}\right)^{2n} \tag{21}$$

$$= \sum_n \sum_{i=0}^{n} \frac{(2n)!}{i!(n-i)!i!(n-i)!} \left(\frac{1}{4}\right)^{2n} \tag{22}$$

$$= \sum_n 4^{-2n} \binom{2n}{n} \sum_{i=0}^{n} \binom{n}{i}\binom{n}{i} \tag{23}$$

$$= \sum_n 4^{-2n} \binom{2n}{n} \sum_{i=0}^{n} \binom{n}{i}\binom{n}{n-i} \tag{24}$$

$$= \sum_n 4^{-2n} \binom{2n}{n}\binom{2n}{n} \qquad \text{(why?)} \tag{25}$$

$$\sim \sum_n \frac{1}{\pi n} \tag{26}$$

$$= \infty. \tag{27}$$

The penultimate equation follows by Stirling's approximation $n! \sim \sqrt{n} n^n e^{-n} \sqrt{2\pi}$, from which $\binom{2n}{n} \sim 2^{2n}/\sqrt{\pi n}$. Hence, the chain is recurrent, and because again the matrix is doubly stochastic, the invariant measure has infinite mass. The chain is thus positive recurrent.

Next, consider three dimensional case. Using the same logic, we have that

$$\sum_m P^m(0,0) = \sum_n P^{2n}(0,0) \tag{28}$$

$$= \sum_n \sum_{\substack{i,j \\ 0 \le i+j \le n}} \frac{(2n)!}{i!j!(n-i-j)!i!j!(n-i-j)!} \left(\frac{1}{6}\right)^{2n} \tag{29}$$

$$= \sum_n 2^{-2n} \binom{2n}{n} \sum_{\substack{i,j \\ 0 \le i+j \le n}} \binom{n}{i\ j}\binom{n}{i\ j} 3^{-2n} \tag{30}$$

$$\le \sum_n 2^{-2n} \binom{2n}{n} c_n 3^{-n} \sum_{\substack{i,j \\ 0 \le i+j \le n}} \binom{n}{i\ j} 3^{-n} \tag{31}$$

$$= \sum_n 2^{-2n} \binom{2n}{n} c_n 3^{-n}, \tag{32}$$

where $c_n = \max\limits_{\substack{i,j \\ 0 \le i+j \le n}} \binom{n}{i\ j}$. We can show that $c_n \sim \binom{n}{n/3\ n/3}$. We thus get that

$$P^{2n}(0,0) \le 2^{-2n} \binom{2n}{n} \frac{n!}{(n/3)!(n/3)!(n/3)!} 3^{-n} \sim C n^{-3/2}, \tag{33}$$

by Stirling's approximation. Thus, the chain is transient.

Reference 13. Let Δ be the drift operator $\Delta V = PV - V$ for the symmetric random walk in d dimensions. When $d = 1$,

$$\Delta V(i) = \frac{1}{2}\left(V(i+1) + V(i-1)\right) - V(i) \tag{34}$$

$$= \frac{V(i+1) - 2V(i) + V(i-1)}{2} \tag{35}$$

$$= \frac{(V(i+1) - V(i)) - (V(i) - V(i-1))}{2}. \tag{36}$$

In general, we have that

$$\Delta V(i_1, \ldots, i_d) = \tag{37}$$

$$\frac{1}{d}\sum_{k=1}^{d} \frac{V(i_1, \ldots, i_{k-1}, i_k + 1, i_{k+1}, \ldots, i_d) - 2V(i_1, \ldots, i_d) + V(i_1, \ldots, i_{k-1}, i_k - 1, i_{k+1}, \ldots, i_d)}{2}, \tag{38}$$

which looks much worse than it is. Notice that we have second-order divided differences in each variable, reminiscent of second derivatives.

Reference 14. Let X be a symmetric random walk on $\mathcal{S} = \mathbb{Z}^d$. Let $S_0 \subset \mathcal{S}$ have boundary ∂S_0. Then, Δ restricted to S_0 corresponds to a symmetric matrix on S_0 and thus has eigenvalues (λ_k) and eigenfunctions (ϕ_k) and the solution to the above problem takes the form

$$f(s) = \frac{1}{d}\sum_k \left(\sum_{\substack{u \in S_0, v \in \partial S_0 \\ u, v \text{adjacent}}} \phi_k(u) g(v) \right) \frac{\phi_k(s)}{\lambda_k}. \tag{39}$$

Again, this looks bad, but don't worry about the details. The key point is that given the computable properties of the operator Δ and given boundary conditions, we can find what we need quite simply.