

## LECTURE NOTES

ELI

**Recall:** We're interested in probability measures  $\pi$  which are solutions to  $\pi G = 0$ .

Assume we have a countable (or finite) state space  $X$ .  $g_{i,j}$  represents the transition rate from  $i$  to  $j$ , and  $g_i = \sum_{j \neq i} g_{i,j}$  is the rate of leaving state  $i$ . Suppose  $\max_i g_i = g < \infty$  (uniformly bounded rates).

Let  $P_{i,j} = \frac{g_{i,j}}{g}$  for  $j \neq i$ . Then  $\sum_{j \neq i} P_{i,j} = \frac{g_i}{g} \leq 1$ . Let  $P_{i,i} = 1 - \frac{g_i}{g}$  so that  $\sum_j P_{i,j} = 1$ .  $P$  is a transition matrix with  $G = g(P - I)$ .

Note,  $\pi G = 0$  iff  $\pi P = \pi$ . In continuous time,  $P_t = e^{tG}$  is the operator corresponding to moving forward  $t$  units of time.  $\pi P_t = \pi \sum_{k=0}^{\infty} \frac{t^k G^k}{k!} = \pi$  since nonzero powers of  $G$  vanish under  $\pi$ .  $\pi P_t$  is the distribution at time  $t$ , starting at  $\pi$ , so this says we will remain in  $\pi$  for all times  $t$ .

**Existence:** When is there a solution to  $\pi P = \pi$ ?

Assume  $P$  is irreducible so that for all  $i, j$ , there exists  $t$  such that  $P^t(i, j) > 0$  (meaning every state is reachable by any other state eventually.) Then there exists a solution to  $\pi P = \pi$  but  $\pi$  need not be a finite measure.

**Example:** Consider the simple random walk on the integers with  $P_{i,j} = \frac{1}{2}$  if  $j = i \pm 1$  and  $P_{i,j} = 0$  otherwise. The solution is  $\pi = 1$ .

**Definition:** A chain with transition matrix  $P$  is called positive recurrent if  $\pi P = \pi$  has a finite measure solution  $\pi$ . A state  $i$  is called recurrent if  $\mathbb{P}_i(\text{Return to } i) = 1$ . Otherwise, the state is called transient.

If  $P$  is irreducible, then either all states are recurrent or all states are transient.

**Theorem:** A state  $i$  is recurrent iff  $\mathbb{E}_i[\tau_i^+] < \infty$  where  $\tau_i^+ = \min\{t \geq 1 : X_t = i\}$ . Equivalently,  $\mathbb{P}_i(\tau_i^+ < \infty) = 1$ . (Obvious when the state space  $X$  is finite and irreducible.)

**Example:** A random walk on the non-negative integers where the state increases by one with probability  $p$  and decreases by one with probability  $q = 1 - p$ . (The transition from 0 to 0 takes probability  $q$ .)

$p > q$  implies the chain is not recurrent. If  $q > p$ , then the chain might be recurrent. For  $k \neq 0$  we have  $\pi(k) = \pi(k-1)p + \pi(k+1)q$  since there are two different ways to arrive at state  $k$ . As the special case, we have  $\pi(0) = \pi(0)q + \pi(1)q$ .

We want to solve  $\pi P = \pi$ . We get  $\pi(1) = \frac{p}{q}\pi(0)$  and  $\pi(k) = \left(\frac{p}{q}\right)^k \cdot \pi(0)$  as a solution. Thus,  $\sum_k \pi(k) = \pi(0) \left(1 - \frac{p}{q}\right)$  and  $\pi(0) = 1 - \frac{p}{q}$ . We have a probability distribution when  $\pi(k) = \left(1 - \frac{p}{q}\right) \left(\frac{p}{q}\right)^k$ . Thus, positive recurrence.

**Definition:** A chain is aperiodic if  $\gcd\{t : P^t(x, x) > 0\} = 1$ . In the case of a bipartite graph, we can only return at even times, and  $P^t(x, x) = 0$  if  $t$  is odd. (Not needed in continuous time.)

If the chain is positive recurrent and irreducible (with aperiodicity in the discrete time case), then there exists a unique solution  $\pi$  which solves  $\pi P = \pi$  (equivalently  $\pi G = 0$ ).

**(Convergence) Theorem:** The distribution starting at  $x$  after  $t$  steps converges to  $\pi$  as  $t \rightarrow \infty$  in the sense that  $d(P^t(x, \cdot), \pi) \rightarrow 0$  for some metric  $d$ .

**Example:**  $d_1(P^t(x, \cdot), \pi) = \sum_y |P^t(x, y) - \pi(y)| \rightarrow 0$  and  $t \rightarrow \infty$ . Thus,  $P^t(x, y) \rightarrow \pi(y)$  for all  $x$ .

Now take  $\mu$  to be any distribution. We have  $\mu P^t \rightarrow \pi$ .

**Application:** Tilings of a square by dominoes. We want to select a tiling uniformly at random. The naive way to do this is list (enumerate) them all and then select one using a random number generator. In practice, this is very hard to do because there are so many possibilities. Another way might be to choose positions and orientations of dominoes in sequence until the whole space is filled, but not all tilings will occur with equal probability.

Alternatively, come up with a way to transition between random tilings such as rotating a square of two dominoes or modifying other such sub-tilings. We can construct a Markov chain with the uniform distribution as the stationary distribution  $\pi$ . We start by choosing the starting position according to any distribution  $\mu$  and then proceed by running the chain for sufficiently large  $t$ . The closer  $\mu$  is to be uniform, the faster the convergence, but  $\mu$  could theoretically be very far from uniform.

**Coupling:** Let  $X_t$  be a chain started from  $x$ , and let  $\tilde{Y}_t$  be a chain started from  $\pi$ . Let  $\tau = \inf(t > 0 : X_t = \tilde{Y}_t)$ . When the state space  $X$  is finite, then  $\tau < \infty$ . Define  $Y_t = \tilde{Y}_t$  for  $t \leq \tau$  and  $Y_t = X_t$  for  $t \geq \tau$ .

Check:  $Y_t$  is a Markov chain with transition matrix  $P$  still started with  $\pi$  (comes from  $\tilde{Y}_t$ ), and  $Y_t$  always has distribution  $\pi$ .

$$\begin{aligned}
|P^t(x, z) - \pi(z)| &= |\mathbb{P}(X_t = z) - \mathbb{P}(Y_t = z)| \\
&= |\mathbb{P}(X_t = z, \tau \leq t) + \mathbb{P}(X_t = z, \tau > t) - \mathbb{P}(Y_t = z, \tau \leq t) - \mathbb{P}(Y_t = z, \tau > t)| \\
&= |\mathbb{P}(X_t = z, \tau > t) - \mathbb{P}(Y_t = z, \tau > t)| \\
&\leq \mathbb{P}(X_t = z, \tau > t) + \mathbb{P}(Y_t = z, \tau > t) \\
&\leq 2\mathbb{P}(\tau > t)
\end{aligned}$$

Using  $\mathbb{P}(\tau < \infty) = 1$ , we have  $2\mathbb{P}(\tau > t) \rightarrow 0$  and thus  $d_1(P^t(x, \cdot), \pi) = \sum_z |P^t(x, z) - \pi(z)| \rightarrow 0$  as  $t \rightarrow \infty$ .

In general, solving for the stationary distribution takes work.

**Definition:**  $P$  (or  $G$ ) is reversible with respect to  $\pi$  if  $\pi(x)P(x, y) = \pi(y)P(y, x)$  (or  $\pi(x)G(x, y) = \pi(y)G(y, x)$ ). These are called the detailed balance equations which say the "flow" from  $x$  to  $y$  is the same as from  $y$  to  $x$ .

**Example:** Weighted random walk with  $P(x, y) = \frac{w_{x,y}}{w_x}$  where  $w_x = \sum_z w_{x,z}$ .  $P$  is reversible with  $\pi(x) = w_x / (\sum_z w_z)$ .

If  $P$  is reversible, then  $P$  need not be symmetric but can be diagonalized.  $A(x, y) = \sqrt{\pi(x)/\pi(y)}P(x, y)$  is symmetric. By the Spectral Theorem,  $A$  has real eigenvalues  $\lambda_1 > \dots > \lambda_n$  and the eigenfunctions  $\varphi_1, \dots, \varphi_n$  form an orthonormal basis for  $\mathbb{R}^n$ .

$A = D_\pi^{1/2} P D_\pi^{-1/2}$  where  $D_\pi$  is diagonal with  $D_\pi(i, i) = \pi(i)$ .  $f_j = D_\pi^{1/2} \varphi_j$  are eigenfunctions for  $P$ .

Define  $\langle f, g \rangle_\pi = \sum_x f(x)g(x)\pi(x)$ , then  $\{f_j\}$  are orthonormal with respect to this inner product.  $P^t(x, y) = \sum_j f_j(y)\lambda_j^t f_j(x)\pi(y)$ .

$$\left| \frac{P^t(x, y)}{\pi(y)} - 1 \right| \leq \sum_{j=2}^n |f_j(x)f_j(y)|\lambda_j^t$$

where  $\lambda_j^t$ 's tell us how fast the convergence is.