**STA4502: Topics in Stochastic Processes** Winter 2018

# Lecture 4: Random Walks on Groups — March 21, 2018

*Lecturer: Jeffrey Rosenthal* *Scribe: Tiantian Zheng*

## An aside on the choice of norm

In our last lecture we obtained bounds on the convergence rate in terms of eigenvalues. The norm we used for this exercise was the $L^2(\pi)$ norm, defined between vectors $v$ and $w$ as:

**Definition 31.** $\langle v, w\rangle_{L^2(\pi)} := \sum_{x\in\mathcal{X}} v(x)\overline{w(x)}\pi(x)$

Under this norm, $P$ is reversible iff $\langle v, Pw\rangle = \langle Pv, w\rangle$, $\forall v, w$ .

Other norms might also be used, such as:

**Definition 40.** $\langle v, w\rangle^* := \sum_{x\in\mathcal{X}} \frac{v(x)\overline{w(x)}}{\pi(x)}$

This norm can be thought of as acting on densities instead of vectors w.r.t $\pi$ as

$$\langle v, w\rangle^* = \sum_{x\in\mathcal{X}} \frac{v(x)}{\pi(x)}\frac{\overline{w(x)}}{\pi(x)}\pi(x)$$

In this case, $P$ is reversible iff $\langle v, wP\rangle^* = \langle vP, w\rangle^*$, $\forall v, w$.

**Challenge 41.** Check requirements for reversibility under the two norm definitions.

In both cases, $P$ being self adjoint implies that there exists an orthonormal basis $\{v_i\}$ which can be used to improve bounds on the convergence rate. Furthermore, the norm according to Def. 40 has some nice properties when we consider how it bounds the total variation distance:

$$\begin{aligned}
2\|\mu_k - \pi\|_{TV} &= \sum_{x\in\mathcal{X}} |\mu_k(x) - \pi(x)| \\
&= \sum_{x\in\mathcal{X}} |\frac{\mu_k(x)}{\pi(x)} - 1|\pi(x) \\
&= \|\mu_k - \pi\|_{L^1(*)} \\
&\le \|\mu_k - \pi\|_{L^2(*)}
\end{aligned}$$

where the inequality comes from the Cauchy-Schwarz inequality.

This bound does not depend on $\pi$, whereas using the $L^2(\pi)$ norm, the coefficient for the bound depends on $\pi$:

$$2\|\mu_k - \pi\|_{TV} \le \sqrt{\frac{n}{\min\limits_x \pi(x)}}\|\mu_k - \pi\|_{L^2(\pi)}$$

However, as we will see, in the case of random walks on groups, $\pi$ is uniform on $\mathcal{X}$, and both definitions of the norm work, and we will keep using the $L^2(\pi)$ norm.

# 7 Random Walks on Groups

In the last section we saw that we could get various bounds on convergence in terms of eigenvalues and eigenvectors of the transition matrix $P$. However, in general, it is usually hard to find these if $\mathcal{X}$ is large.

In the case of random walks on groups, however, it is always possible to obtain explicit forms for the eigenvalues and eigenvectors.

**Definition 42.** A random walk on a group is a Markov chain on a state space of some general discrete group $\mathcal{X}$. Transition probabilities are written as $P(x, y) = Q(x^{-1}y)$ where $Q(\cdot)$ is some fixed step distribution on $\mathcal{X}$. The increment distributions defined by $Q$ are i.i.d.

**Example 43.** Our previous example of card shuffling is a random walk on the group $S_n$, the symmetric group of permutations.

**Example 44.** Random walk on $\mathbb{Z}$:

- For the random walk with $\frac{1}{2} - \frac{1}{2}$ probabilities of moving by +1 and -1:

$$Q(+1) = Q(-1) = \frac{1}{2}$$

- For the random walk with $\frac{1}{3} - \frac{1}{3} - \frac{1}{3}$ probabilities of moving by +1, 0 and -1:

$$Q(+1) = Q(0) = Q(-1) = \frac{1}{3}$$

It is possible to work with continuous groups, e.g. $O(n)$, the orthogonal group, containing the set of all $n \times n$ orthogonal matrices. For now, we restrict our discussion to finite, abelian groups. As the law of composition is commutative on these groups, we use addition notation, i.e. replace $Q(x^{-1}y)$ with $Q(y - x)$ to represent $P(x, y)$.

**Fact 45.** A random walk $P$ on a finite group always has $\pi = \text{Unif}(\mathcal{X})$, i.e. $\pi(x) = \frac{1}{n}, \forall x \in \mathcal{X}$, since $P$ is doubly stochastic (i.e. $\sum_x P(x, y) = 1, \forall y$).

*Proof.*

$$\sum_x P(x, y) = \sum_x Q(y - x) = \sum_z Q(z) = 1$$

$\square$

**Fact 46.** Finite abelian groups are always of the form $\mathcal{X} = \mathbb{Z}/(n_1) \times \mathbb{Z}/(n_2) \times ... \times \mathbb{Z}/(n_r)$

**Example 47. Frog walk**: A frog jumps on a circular arrangement of 20 lilypads, each time moving clockwise or counterclockwise by one lilypad. The state space is given by $\mathcal{X} = \mathbb{Z}/(20)$.

**Example 48. Bit flipping**: A set of bits of length $d$ can have each bit, or no bit flipped at each timestep with probability $\frac{1}{d+1}$.

- The state space is given by $\mathcal{X} = \big(\mathbb{Z}/(2)\big)^d$.
- The step distribution is given by $Q(0) = Q(\mathbf{e}_1) = ... = Q(\mathbf{e}_d)$, where $\mathbf{e}_i = (0, 0, 0, ..., 0, 1, 0, ...)$, i.e. all entries are 0, except for the $i$th entry, which is replaced by 1.

For the above examples, it is not immediately obvious what the eigenvalues and eigenvectors are. To derive these, we introduce **characters**, which are the beginnings of representation theory of groups.

## 7.1 Characters

**Definition 49.** $\chi_m : \mathcal{X} \to \mathbb{C}$ is a character defined for $m = (m_1, m_2, ..., m_r) \in \mathcal{X}$,

$$\chi_m(x) = e^{2\pi i\left(\frac{m_1 x_1}{n_1} + \frac{m_2 x_2}{n_2} + ... \frac{m_r x_r}{n_r}\right)}$$

**Note 50.** (some identities)

1. $\chi_m(x + y) = \chi_m(x)\chi_m(y)$
2. $\chi_m(0) = 1$
3. $|\chi_m(x)| = 1$
4. $\chi_m(-x) = \overline{\chi_m(x)}$
5. $\sum_{m\in\mathcal{X}} \chi_m(x) = \begin{cases} n, x = 0 \\ 0, x \neq 0 \end{cases} \quad = n\delta_{x0}$, where $n = n_1 n_2 ... n_r = |\mathcal{X}|$
6. $\langle \chi_m, \chi_j \rangle_{L^2(\pi)} = \sum_{x\in\mathcal{X}} \chi_m(x)\overline{\chi_j(x)}\pi(x) = \sum_{x\in\mathcal{X}} \chi_m(x)\chi_j(x)\frac{1}{n} = \begin{cases} 1, m = j \\ 0, m \neq j \end{cases} \quad = \delta_{mj}$

Identity 5 for $x \neq 0$ follows from the fact that $\chi_m(x)$ are equally distributed on the unit circle in the complex plane, or alternatively, noting that this is a product of geometric sums that evaluate to 0

$$\sum_{m\in\mathcal{X}} \chi_m(x) = \sum_{m\in\mathcal{X}} \Big(\prod_{j=1}^{r} e^{2\pi i \frac{m_j x_j}{n_j}}\Big) = \prod_{j=1}^{r} \Big(\sum_{m_j=0}^{n_j-1} e^{\frac{2\pi i m_j x_j}{n_j}}\Big) = \prod_{j=1}^{r} \frac{1 - e^{2\pi i x_j}}{1 - e^{\frac{2\pi i x_j}{n_j}}} = 0$$

In Identity 6, we have made use of the fact that $\pi = \text{Unif}(\mathcal{X})$, and when $m \neq j$, the sum reduces to $\sum_x \chi_{m-j}(x)$, which is zero for the same reason as in Identity 5.

It follows therefore from Identity 6 that $\{\chi_m\}$ are orthonormal. It remains to be shown that they are eigenvectors.

$$\begin{aligned}
\big(\overline{\chi_m}P\big)(y) &= \sum_{x\in\mathcal{X}} \overline{\chi_m(x)}P(x,y) \\
&= \sum_{x\in\mathcal{X}} \chi_m(-x)P(x,y) \\
&= \sum_{x\in\mathcal{X}} \chi_m(-x)Q(y-x)
\end{aligned}$$

Making the change of variable $z = y - x, -x = z - y$:

$$\begin{aligned}
\big(\overline{\chi_m}P\big)(y) &= \sum_{z\in\mathcal{X}} \chi_m(z-y)Q(z) \\
&= \sum_{z\in\mathcal{X}} \chi_m(z)\chi_m(-y)Q(z) \\
&= \overline{\chi_m(y)} \sum_{z\in\mathcal{X}} \chi_m(z)Q(z) \\
&= E_Q(\chi_m)\overline{\chi_m(y)}
\end{aligned}$$

Therefore, $\{\chi_m\}$ is the set of eigenvectors, with corresponding eigenvectors, $\{\lambda_m\}$ being the expectation of the characters under $Q$. As usual, we set $\lambda_0 = 1$ and define $\lambda_* = \max_{m\neq 0}|\lambda_m|$.

Finally we want to be convinced that in the case when the random walker starts in a designated position, i.e. $\mu_0 = \delta_0(\cdot)$ is a point mass, we can still write $\mu_0$ as a linear combination of the eigenvectors of $P$. I.e. we want to show that $\mu_0 = \sum_m a_m v_m$ for some set of complex coefficients $\{a_m\}$.

This can be done by simply observing that $a_m = \frac{1}{n}$ since $\sum_m \overline{\chi_m(x)} = n\delta_{x0} = \sum_m v_m$. This leads us to conclude that $\mu_0 - \pi = \frac{1}{n}(\sum_m v_m - \mathbf{1}) = \frac{1}{n}(\sum_{m\neq 0} v_m)$.

Therefore $\mu_k = \frac{1}{n}\sum_m (\lambda_m)^k v_m$, and $\mu_k - \pi$. From this we obtain

$$\sum_{x\in\mathcal{X}} |\mu_k(x) - \pi(x)|^2\pi(x) = \sum_{m\neq 0} |a_m|^2|\lambda_m|^{2k}$$

as $\{v_m\}$ are orthonormal.

And as $\pi(x) = \frac{1}{n} = a_m$

$$\sum_{x\in\mathcal{X}} |\mu_k(x) - \pi(x)|^2 = \frac{1}{n}\sum_{m\neq 0} |\lambda_m|^{2k}$$

We therefore obtain a bound on the total variation distance

$$
\begin{aligned}
\Big(2\|\mu_k-\pi\|_{TV}\Big)^2 &= \Big(\sum_{x\in\mathcal{X}}|\mu_k(x)-\pi(x)|\Big)^2\\
&= \Big(n\sum_{x\in\mathcal{X}}|\mu_k(x)-\pi(x)|\pi(x)\Big)^2\\
&= n^2\Big(\langle\mu_k-\pi,\mathbf{1}\rangle\Big)^2\\
&\le n^2\|\mu_k-\pi\|^2_{L^2(\pi)}\|\mathbf{1}\|^2_{L^2(\pi)}\\
&= n^2\|\mu_k-\pi\|_{L^2(\pi)}\\
&= \sum_{m\neq 0}|\lambda_m|^{2k}
\end{aligned}
$$

where again the inequality comes from the Cauchy-Schwarz inequality.

**Conclusion 51.** $\|\mu_k-\pi\|_{TV}\le\frac{1}{2}\sqrt{\sum_{m\neq 0}|\lambda_m|^{2k}}\le\frac{\sqrt{n-1}}{2}(\lambda_*)^k$

## 7.2 Application to examples

### 7.2.1 Frog walk

$\mathcal{X}=\mathbb{Z}/(n)$, $\qquad Q(0)=Q(1)=Q(-1)=\frac{1}{3}$

$$
\begin{aligned}
\chi_m(x) &= e^{2\pi i(\frac{mx}{n})}\\
\lambda_m &= E_Q\big(\overline{\chi_m}\big)\\
&= \frac{1}{3}\overline{\chi_m}(0)+\frac{1}{3}\overline{\chi_m}(1)+\frac{1}{3}\overline{\chi_m}(-1)\\
&= \frac{1}{3}(1)+\frac{1}{3}e^{\frac{-2\pi im}{n}}+\frac{1}{3}e^{\frac{2\pi im}{n}}\\
&= \frac{1}{3}+\frac{2}{3}\cos(\frac{2\pi m}{n})
\end{aligned}
$$

$m=0$ corresponds to $\lambda_m=1$. It can be seen that as $m$ increases, $\cos(\frac{2\pi m}{n})$ decreases, then increases back towards 1, but cannot exceed $\cos(\frac{2\pi}{n})$. The value for $\lambda_*$ is therefore $\frac{1}{3}+\frac{2}{3}\cos\big(\frac{2\pi}{n}\big)$.

$$
\begin{aligned}
\|\mu_k-\pi\| &\le \frac{\sqrt{n}}{2}\big(\frac{1}{3}+\frac{2}{3}\cos(\frac{2\pi}{n})\big)^k\\
&= \frac{\sqrt{n}}{2}\big(1-\frac{2}{3}(1-\cos(\frac{2\pi}{n}))\big)^k
\end{aligned}
$$

Assuming $n\ge 3$, we have that for $0\le x\le\sqrt{6}$, $\cos(x)\le 1-\frac{x^2}{4}$. Further, $1-x\le e^{-x}$, therefore,

$$\|\mu_k - \pi\| \le \frac{\sqrt{n}}{2} e^{-\frac{2\pi^2}{3n^2}k}$$

For $n = 1000$, this gives $k_* = 1120000$. This bound requires $k$ to be on the order of $n^2\log(n)$. Since we know all the eigenvalues, we can obtain a tighter bound

$$\begin{aligned}
\|\mu_k - \pi\|^2 &\le \frac{1}{4} \sum_{m=1}^{n-1} |\lambda_m|^{2k} \\
&\le \sum_{m=1}^{\lceil \frac{n-1}{4} \rceil} e^{-\frac{4\pi^2 m^2}{3n^2}k} \\
&\le \sum_{m=1}^{\infty} e^{-\frac{4\pi^2 m}{3n^2}k} \\
&= \frac{e^{-\frac{4\pi^2}{3n^2}k}}{1 - e^{-\frac{4\pi^2}{3n^2}k}}
\end{aligned}$$

Which for $n = 1000$, gives $k_* = 351000$. This bound now scales with $n^2$. How much tighter still can we make this bound? To answer this question, we look at the lower bound for convergence. First note that as $E_{\mu_k}(\chi_m)$ is the eigenvalue of $P^k$ corresponding to the eigenvector $\overline{\chi_m}$, it is equal to the $k$th power of the corresponding eigenvalue of $P$:

$$E_{\mu_k}(\chi_m) = (E_Q(\chi_m))^k$$

We therefore have

$$\begin{aligned}
\|\mu_k - \pi\| &= \frac{1}{2} \sup_{|f|\le 1} |E_{\mu_k}(f) - E_\pi(f)| \\
&\ge \frac{1}{2} |E_{\mu_k}(\chi_1) - 0| \\
&= \frac{1}{2} |E_Q(\chi_1)|^k \\
&= \frac{1}{2} \left( \frac{1}{3} + \frac{2}{3} \cos(\frac{2\pi}{n}) \right)^k
\end{aligned}$$

where the inequality comes from the fact that the supremum of a set must be greater than or equal to any member of that set.

so for $n = 1000$, $k_* \ge 290000$, therefore our previous bound cannot be improved by much more.