STAT 31512 Spring 2024
Monday, April 1, 2024
Scribe: Qingqi Zhang
Lecturer: Frederic Koehler

Spectral theory of reversible Markov chains

1 Part 1: Reversibility

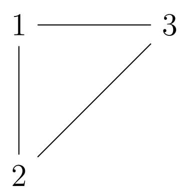
Definition 1. A Markov chain P is reversible with respect to π if and only if ΠP is symmetric, where $\Pi = \mathrm{diag}(\pi_1, \dots, \pi_s)$.

Remark 1. Reversibility is equivalent to $\pi_i P_{ij} = \pi_j P_{ji}$ for all i, j. Additionally, it implies that $\Pi^{1/2} P \Pi^{-1/2}$ is symmetric, indicating that P is diagonalizable.

Fact 1. *If P is symmetric, then π is uniform, specifically $\pi = \frac{1}{s}$, and consequently, ΠP is symmetric.*

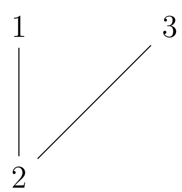
Proof. Since $P^T \mathbb{1} = P \mathbb{1} = \mathbb{1}$, and by the fundamental theorem of algebra, $\mathbb{1}$ is the unique right eigenvector of P corresponding to the eigenvalue 1, it follows that π is proportional to $\mathbb{1}$. □

Example 1. Random walk on a d-regular graph:



$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \Bigg| \quad P = \frac{1}{d} A = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \quad \Bigg| \quad \pi = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right)$$

Example 2. Periodic Graph



$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \Bigg| \quad P = \frac{1}{\mathrm{deg}} A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \quad \Bigg| \quad \pi = \left(\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}\right)$$

where $\mathrm{deg} = \mathrm{diag}(\mathrm{deg}_1, \mathrm{deg}_2, \mathrm{deg}_3) = \mathrm{diag}(1, 2, 1)$.

Fact 2. ***Reversibility:*** *Suppose $X_0 \sim \pi$, then the Markov chain $\cdots \to X_{-2} \to X_{-1} \to X_0 \to X_1 \cdots$ can be reversed, i.e., the law of $\{X_t\}_{t \geq 0}$ is the same as the law of $\{X_{-t}\}_{t \geq 0}$.*

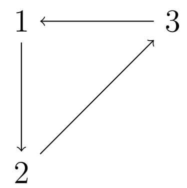
Proof. Consider transitions between states in a reversible Markov chain:

$$X_0 \to X_1 \quad \stackrel{d}{=} \quad X_0 \leftarrow X_1,$$

$$\begin{aligned} \Pr(X_0 = j | X_1 = i) &= \frac{\Pr(X_0 = j) \Pr(X_1 = i | X_0 = j)}{\Pr(X_1 = i)} \\ &= \frac{\pi_j P_{ji}}{\pi_i} \\ &= P_{ji}, \end{aligned}$$

which confirms the reversibility as the forward and reverse transitions obey the same probability law. □

Example 3. Non-reversible Case:



In the forward direction, the walk proceeds as $1 \to 2 \to 3 \to 1 \to 2 \cdots$, while in the reverse direction, it goes $1 \to 3 \to 2 \to 1 \to 3 \cdots$

2 Part 2: Spectral Theorem

Theorem 1 (Concrete Version)**.** *For any symmetric matrix $M \in \mathbb{R}^{n\times n}$, it can be represented as $M = U\Lambda U^T = \sum_i \lambda_i q_i q_i^T$, where $q_i^T q_j = \delta_{ij}$, and each λ_i is a real number.*

Theorem 2 (Abstract Version)**.** *Let $\langle \cdot, \cdot \rangle_\pi$ be an inner product on $\mathbb{R}^n$. If P is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\pi$, meaning $\langle \alpha, P\beta \rangle_\pi = \langle P\alpha, \beta \rangle_\pi$ for all α, β, then P can be diagonalized as $Pf = \sum_i \lambda_i \langle f_i, f \rangle_\pi f_i$, where $\langle f_i, f_j \rangle_\pi = \delta_{ij}$, the f_i are real orthonormal eigenvectors, and λ_i are the corresponding real eigenvalues. In matrix form, $P = F\Lambda F^{-1}$, where F is the matrix with columns $f_1, \ldots, f_n$.*

Definition 2. The inner product $\langle f, g \rangle_\pi$ for vectors f and g with respect to the stationary distribution π is defined as $\langle f, g \rangle_\pi = f^T \Pi g$, where Π is a diagonal matrix with entries of π. This is equal to $\sum_i f(i)g(i)\pi_i = \mathbb{E}_{X\sim\pi}[f(X)g(X)]$.

Fact 3. *If P is reversible w.r.t π, then it is also self-adjoint with respect to $\langle \cdot, \cdot \rangle_\pi$.*

Proof. The inner product of f and Pg under π is given by:

$$\begin{aligned}
\langle f, Pg \rangle_\pi &= \langle f, \Pi P g \rangle \\
&= \langle f, (\Pi P)^T g \rangle \\
&= \langle f, P^T \Pi g \rangle \quad (\text{since } P^T \text{ is the adjoint of } P \text{ with respect to } \langle \cdot, \cdot \rangle) \\
&= \langle Pf, \Pi g \rangle \\
&= \langle Pf, g \rangle_\pi.
\end{aligned}$$

This equality demonstrates that P maintains self-adjointness under the given inner product. □

Remark 2. The Spectral Theorem ensures that a matrix P can be expressed as $P = \sum_i \lambda_i f_i f_i^T \Pi = F\Lambda F^T \Pi = F\Lambda F^{-1}$, since $F^T \Pi F = I_s$, where I_s is the identity matrix of size s, so $F^{-1} = F^T \Pi$.

3 Part 3: Poincaré Inequality

Variational Characterization of Eigenvectors

For a matrix P:

$$\lambda_2(P) = \sup \left\{ \frac{\langle f, Pf \rangle_\pi}{\langle f, f \rangle_\pi} : \langle f, 1 \rangle_\pi = 0 \right\}.$$

This implies:

$$\lambda_2(P) \geq \frac{\langle f, Pf \rangle_\pi}{\langle f, f \rangle_\pi}, \quad \text{for all } f \text{ such that } \langle f, 1 \rangle_\pi = 0.$$

It follows that:

$$\lambda_2(P) \langle f, f \rangle_\pi \geq \langle f, Pf \rangle_\pi,$$

$$-\lambda_2(P)\langle f, f\rangle_\pi \le -\langle f, Pf\rangle_\pi,$$

$$(I-\lambda_2(P))\langle f, f\rangle_\pi \le \langle f, (I-P)f\rangle_\pi, \quad \text{for all } f \text{ such that } \langle f, 1\rangle_\pi = 0.$$

Define the centered function:

$$\hat{f} = f - \mathbb{E}_\pi[f]\cdot \mathbb{1}, \quad \text{since } \langle \hat{f}, 1\rangle_\pi = 0.$$

This leads to the inequality:

$$(1-\lambda_2(P))\langle \hat{f}, \hat{f}\rangle_\pi \le \langle \hat{f}, (I-P)\hat{f}\rangle_\pi,$$

$$(1-\lambda_2(P))\langle \hat{f}, \hat{f}\rangle_\pi \le \langle f, (I-P)f\rangle_\pi,$$

justified by $(I-P)\mathbb{1}_s = 0$ and the fact that $I-P$ is self-adjoint with respect to $\langle\cdot,\cdot\rangle_\pi$.

Moreover, we find that:

$$\langle \hat{f}, \hat{f}\rangle_\pi = \mathbb{E}_\pi[(f-\mathbb{E}_\pi[f])^2] = \mathrm{Var}_\pi(f-\mathbb{E}_\pi[f]).$$

Definition 3. The **Laplacian** of P is defined as $L = I - P$.

Remark 3.

- L is self-adjoint with respect to $\langle\cdot,\cdot\rangle_\pi$.
- L is positive semi-definite, and $L\mathbb{1} = 0$.
- For a d-regular graph, $L = \frac{1}{d}(dI - A) = \frac{1}{d}\hat{L}$, where $\hat{L}$ is the standard Laplacian of the graph.

Theorem 3 (Poincaré Inequality)**.** *From the calculations above, we establish the Poincaré Inequality:*

$$\mathrm{Var}_\pi[f-\mathbb{E}_\pi[f]] \le \frac{1}{1-\lambda_2(P)}\langle f, Lf\rangle_\pi, \quad \forall f.$$

Here, $E_P(f,f)$ denotes the Dirichlet form.

Remark 4.

- If the magnitude of ∇f is small, then the variance $\mathrm{Var}_\pi(f)$ is also small.
- The second eigenvalue λ_2 can be characterized as:

$$1-\lambda_2 = \inf\left\{\frac{\langle f, Lf\rangle_\pi}{\mathrm{Var}_\pi(f-\mathbb{E}_\pi[f])} : f \ne 0\right\},$$

$$\lambda_2 = 1 - \inf\left\{\frac{\langle f, Lf\rangle_\pi}{\mathrm{Var}_\pi(f-\mathbb{E}_\pi[f])} : f \ne 0\right\}$$

$$= \sup\left\{1 - \frac{\langle f, Lf\rangle_\pi}{\mathrm{Var}_\pi(f)} : f \ne 0\right\}.$$

  For any function f, this provides a lower bound on λ_2.

- Moreover, for any value λ such that the inequality

$$\lambda \ge 1 - \frac{\langle f, Lf\rangle_\pi}{\mathrm{Var}_\pi(f)}$$

  holds for all non-zero f, we have an upper bound on λ_2, specifically $\lambda_2 \le \lambda$.