Moden Spectral Graph Theory	**Winter 2022**

Lecture 2: Spectral Profile

Lecturer: Shayan Oveis Gharan — *01/06/22*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

Definition 2.1 (Variance)**.** *Given a graph $G = (V, E)$ with random walk matrix P and stationary distribution π_0, for a function $f : V \to \mathbb{R}$ define*

$$\mathrm{Var}(f) = \langle f - \mathbb{E}f, f - \mathbb{E}f\rangle_{\pi_0} = \langle f, f\rangle_{\pi} - 2\mathbb{E}f\langle f, \mathbf{1}\rangle_{\pi_0} + (\mathbb{E}f)^2 = \langle f, f\rangle_{\pi_0} - (\mathbb{E}f)^2$$

where as usual we used $\|\mathbf{1}\| = 1$ and $\mathbb{E}f = \langle f, \mathbf{1}\rangle$.

In this lecture we see that Variance can be seen as a distance function between the starting distribution of the Markov Chain and the stationary distribution. A Markov Chain is fast mixing if this distance is reduced "significantly" in every step.

The Dirichlet form of the random walk operator P with respect to two functions f, g is defined as follows:

$$\mathcal{E}_P(f, g) = \langle (I - P)g, f\rangle = \mathbb{E}_{u\sim\pi}\mathbb{E}_{\{u,v\}|u} f(u)(g(u) - g(v)).$$

The $I - P$ matrix is the well-known normalized Laplacian matrix. In particular, since all eigenvalues of P are in the range $[-1, +1]$ it follows that the eigenvalues of $I - P$ are in the range $[0, 2]$, i.e., $I - P$ is a PSD matrix. This can be seen immediately by writing the quadratic forms a function f with respect to this operator. In particular, in the special case that $f = g$ the above equation simplifies to

$$\mathcal{E}_P(f, f) = \frac{1}{2}\mathbb{E}_{\{u,v\}\sim\pi_1}(f(u) - f(v))^2. \tag{2.1}$$

Note that $\lambda_1(I - P) = 0$ as $(I - P)\mathbf{1} = 0$. So, by the variational characterization of eigenvalues,

$$\lambda_2(I - P) = \min_{f:\langle f,\mathbf{1}\rangle=0} \frac{\langle (I - P)f, f\rangle}{\langle f, f\rangle} = \min_{f\neq\text{constant}} \frac{\langle (I - P)f, f\rangle}{\mathrm{Var}(f)}$$

Lemma 2.2. *For any function $f : V \to \mathbb{R}$, $\mathrm{Var}(f) - \mathrm{Var}(Pf) = \mathcal{E}_{P^2}(f, f)$. So, in particular, we always have, $\mathrm{Var}(Pf) \leq \mathrm{Var}(f)$.*

Proof. We can write

$$\mathrm{Var}(Pf) = \langle Pf, Pf\rangle - (\mathbb{E}Pf)^2 = \langle P^2 f, f\rangle - \langle Pf, \mathbf{1}^2 = \langle P^2 f, f\rangle - (\mathbb{E}f)^2$$

Therefore,

$$\mathrm{Var}(f) - \mathrm{Var}(Pf) = \langle f, f\rangle - \langle P^2 f, f\rangle = \langle (I - P)f, f\rangle = \mathcal{E}_{P^2}(f, f).$$

□

Lemma 2.3. *Suppose that for any function $f : V \to \mathbb{R}_{\geq 0}$, $\mathrm{Var}(Pf) \leq (1 - \delta)\mathrm{Var}(f)$. Then, for any vertex $u \in V$, the walk started at a vertex u mixes in $\frac{\log \epsilon^{-1}\pi(u)^{-1}}{\delta}$ steps.*

Note that by the previous lemma, the assumption implies that $\lambda_2(I - P^2) \geq \delta$. This inparticular, implies that $\lambda^*(P) = \max\{|\lambda_2(P), |\lambda_n(P)|\} \geq \sqrt{1-\delta}$. So, the above lemma can be proven immediately by the theorem we proved in the last lecture. However, here, we give a directed proof based on changes on the variance.

Proof. Let $f : V \to \mathbb{R}_{\geq 0}$ be a non-negative function. Since P is a positive operator, $P^t f$ is non-negative for all $t \geq 0$. BY the assumption of the lemma for $t \geq \frac{\log \epsilon^{-1}}{\delta}$ we have

$$\operatorname{Var}(P^t f) \leq (1-\delta) \operatorname{Var}(P^{t-1} f) \leq \cdots \leq (1-\delta)^t.$$

Note that $\operatorname{Var}(P^t f) = \|P^t f - \mathbb{E} f\|^2$, so following the same proof as the Corollary 1.6 we get that the chain started at a vertex u mixes in $\frac{\log(\epsilon^{-1}\pi(u)^{-1})}{\delta}$ steps. □

Corollary 2.4. *If $\frac{\mathcal{E}_{P^2}(f,f)}{\operatorname{Var}(f)} \geq \delta$ for all non-constant functions $f : V \to \mathbb{R}_{|geq0}$, then the walk mixes in $O(\frac{1}{\delta} \log \epsilon^{-1}\pi(u)^{-1})$ steps.*

2.1 Spectral Profile

The main topic of this lecture is spectral profile. We see that we can significantly improve the mixing time of random walk if we have better bounds for variance reduction for functions with bounded support. The material of this lecture is based on a paper of Goel, Montenegro and Tetali [GMT06].

Definition 2.5 (Spectral Profile)**.** *For a non-empty subset $S \subseteq V$, define*

$$\lambda(S) = \inf_{f:S\to\mathbb{R}_{\geq 0}} \frac{\mathcal{E}(f,f)}{\operatorname{Var}(f)}.$$

where the infimum is over non-constant functions (note if $S \neq V$, then any such f is not constant. Now, for $r \geq 0$, define the spectral profile,

$$\Lambda(r) := \inf_{S:\pi(S)\leq r} \lambda(S).$$

Lemma 2.6. *For any non-constant function $f : V \to \mathbb{R}_{\geq 0}$,*

$$\frac{\mathcal{E}_P(f,f)}{\operatorname{Var}(f)} \geq \frac{1}{2}\Lambda\left(\frac{4(\mathbb{E}f)^2}{\operatorname{Var} f}\right)$$

In particular, there is a function $g : V \to \mathbb{R}_{\geq 0}$ that is a level set *of f such that $\pi_0(\operatorname{supp}(g)) \leq \frac{4(\mathbb{E}f)^2}{\operatorname{Var}(f)}$ and*

$$2\frac{\mathcal{E}_P(f,f)}{\operatorname{Var}(f)} \geq \frac{\mathcal{E}_P(g,g)}{\operatorname{Var}(g)}.$$

Proof. Let $t \geq 0$ be a threshold that we choose later. For a function g, write g_+ to denote the point-wise maximum of g and 0. We have

$$\begin{aligned}
\mathcal{E}_P(f,f) &\underset{\text{Dirichlet for shift inv}}{=} \mathcal{E}_P(f-t, f-t) \\
&\underset{\forall a,b:(a-b)^2\geq(a_+-b_+)^2}{\geq} \mathcal{E}_P((f-t)_+, (f-t)_+) \\
&= \operatorname{Var}((f-t)_+) \inf_{g:S_{f>t}\to\mathbb{R}_{\geq 0}} \frac{\mathcal{E}_P(g,g)}{\operatorname{Var}(g)} \\
&= \operatorname{Var}((f-t)_+)\Lambda(\mathbb{P}[f \geq t])
\end{aligned}$$

Where we wrote $S_{f>t}$ to denote the set of vertices u with $f(u) > t$. Further,

$$\begin{aligned}\mathrm{Var}((f-t)_+) &= \mathbb{E}(f-t)_+^2 - (\mathbb{E}(f-t)_+)^2 \\ &\underset{\forall a,b (a-b)_+^2 \geq a^2 - 2ab}{\geq} \mathbb{E}f^2 - 2t\mathbb{E}f - (\mathbb{E}(f-t)_+)^2 \\ &\underset{f\geq 0}{\geq} \mathbb{E}f^2 2t\mathbb{E}f - (\mathbb{E}f)^2 = \mathrm{Var}(f) - 2t\mathbb{E}f = \frac{\mathrm{Var}(f)}{2}\end{aligned}$$

where to get the last identity you need to let $t = \frac{\mathrm{Var}(f)}{4\mathbb{E}f}$. Furthermore, by Markov's inequality, $\mathbb{P}[f \geq t] \leq \frac{\mathbb{E}f}{t} = \frac{4(\mathbb{E}f)^2}{\mathrm{Var}(f)}$. Putting these together both statement follows. □

Building on Lemma 2.3 we get the following improved bound on mixing time.

Theorem 2.7. *Given a graph $G = (V, E)$ with a random walk matrix P; let $\Lambda_{P^2}(.)$ be the spectral profile of P^2. For any vertex $u \in V$, the walk started at u mixes in time,*

$$T := 2 \sum_{t=1}^{\log(\pi(u)^{-1})} \frac{1}{\Lambda_{P^2}(2^{-t+2})}.$$

As a motivating application of this lemma we can apply it to bound the mixing time of walks on a hypercube with n vertices. It can be seen that $\lambda_2(P) = 1 - \frac{1}{\log(n)}$ so $\lambda_2(I - P) = \frac{1}{\log n}$. So, from the mixing theorem in lecture 1 we get a mixing-time bound of $O(\log^2 n)$.

But for any $r \geq 0$, $\Lambda(r) \leq \frac{\log r^{-1}}{\log n}$. So, from the above formula the walk mixes in time $O(\log n \log\log n)$.

Proof. As before define $f(u) = 1/\pi(u)$ and zero everywhere else and we have $\mathbb{E}f = 1$ and $\mathrm{Var}(f) = \pi(u)^{-1}$. As showed in Lemma 2.2, $\mathrm{Var}(f) \geq \mathrm{Var}(Pf) \geq \dots$. Furthermore, for any $g : V \to \mathbb{R}_{\geq 0}$, with $\mathbb{E}g = 1$,

$$\frac{\mathrm{Var}(g) - \mathrm{Var}(Pg)}{\mathrm{Var}(g)} = \frac{\mathcal{E}_{P^2}(g,g)}{\mathrm{Var}(g)} \underset{\text{Lemma 2.6}}{\geq} \frac{1}{2}\Lambda_{P^2}\left(\frac{4}{\mathrm{Var}(g)}\right)$$

Therefore, if for $g = P^t f$, $\mathrm{Var}(g) = 2^t$, in $\frac{2}{\Lambda_{P^2}(2^{-t+2})}$) steps, $\mathrm{Var}(g)$ halves. So, after T as defined in the lemma's statement steps we have $\mathrm{Var}(P^T f) \leq \frac{1}{4}$. The rest of the proof is the same as Lemma 2.3 □

Corollary 2.8. *Given a graph $G = (V, E)$, let $\tilde{P}$ be the transition probability matrix of the 1/2-lazy on G. Then the walk mixes in*

$$2 \sum_{t=1}^{\log \pi(u)^{-1}} \frac{1}{\Lambda_{\tilde{P}}(2^{-t+2})}$$

steps.

Proof. Notice $\tilde{P} = (I + P)/2$. Observe that $0 \preceq \tilde{P} \preceq I$ so $\tilde{P}^2 \preceq_\pi \tilde{P}$. This, in particular implies that $I - \tilde{P}^2 \succeq_\pi I - \tilde{P}$ and we can write

$$\mathcal{E}_{P^2}(f,f) = \langle (I - P^2)f, f\rangle \geq \langle (I - \tilde{P}f, f\rangle = \mathcal{E}_{\tilde{P}}(f,f).$$

It follows that for any $0 < r < 1$, $\Lambda_{\tilde{P}^2}(r) \geq \Lambda_{\tilde{P}}(r)$. The statement follows from the theorem. □

References

[GMT06] S. Goel, R. Montenegro, and P. Tetali. "Mixing Time Bounds via the Spectral Profile". In: *Electronic Journal of Probability* 11.none (2006), pp. 1 –26 (cit. on p. 2-2).