Refined mixing and hitting relations

[Basu - Hermon - Peres]

$f, g : S \to \mathbb{R}$ $\qquad \langle f, g \rangle_\pi = \sum_{x \in S} f(x) g(x) \pi(x)$

If P is reversible wrt π (invariant distr.), then there exist $(\lambda_j)_{j=1}^{|S|}$ eigenvalues with corresponding eigenvectors $(f_j)_{j=1}^{|S|}$ and $\lambda_1 = 1$ and $f_1 = (1, \dots, 1)$.

Then $$\frac{P^t(x,y)}{\pi(y)} = 1 + \sum_{j=2}^{|S|} \lambda_j^t f_j(x) f_j(y)$$

Proof Idea $\quad A(x,y) = \frac{\sqrt{\pi(x)}}{\sqrt{\pi(y)}} P(x,y) \rightsquigarrow$ symmetric $\square$

If P is reversible we write the eigenvalues in decreasing order

$$1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_n \geq -1 .$$

Define $\lambda_* = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } P \text{ with } \lambda \neq 1\}$

and $\gamma_* = 1 - \lambda_*$: absolute spectral gap

$\gamma = 1 - \lambda_2$: spectral gap.

For lazy chains $\gamma_* = \gamma$.

Def. $t_{rel} = \frac{1}{\gamma_*}$: relaxation time.

$f : S \to \mathbb{R}$ $\qquad \mathbb{E}_\pi[f] = \sum_x f(x)\pi(x)$ and $\mathrm{Var}_\pi(f) = \mathbb{E}_\pi\left[(f - \mathbb{E}_\pi[f])^2\right]$.

Poincaré $\qquad \mathrm{Var}_\pi(P^t f) \leq e^{-2t/t_{rel}} \mathrm{Var}_\pi(f)$ $\quad$ P reversible, lazy
$\forall f, \ \forall t \geq 0$.

$Pf(x) = \sum_y P(x,y) f(y)$

Define for $\varepsilon, \alpha \in (0,1)$

$$\mathrm{hit}_\alpha(\varepsilon) = \min\{t : \max_{x, A : \pi(A) \geq \alpha} \mathbb{P}_x(\tau_A > t) \leq \varepsilon\}$$

Theorem BHP

Let X be a reversible and lazy Markov chain on S (finite) with P & π.

Then $\forall \varepsilon \in (0, \frac{1}{2})$ $\qquad t_{mix}(2\varepsilon) \leq \mathrm{hit}_{1-\varepsilon}(\varepsilon) + \left\lceil 2 t_{rel} \cdot \log\left(\frac{2}{\varepsilon^3}\right) \right\rceil$ $\qquad (\star)$

and $\quad t_{mix}(1-\varepsilon) \leq hit_{1-\varepsilon}(1-2\varepsilon) + \left\lceil 2\, t_{rel} \cdot \log\left(\frac{8}{\varepsilon^3}\right)\right\rceil. \quad (**)$ Exercise

Remark: Mixing happens in 2 stages: in the first one ($hit_{1-\varepsilon}(\varepsilon)$) we have to wait to escape some small set with high prob and in the second one we wait for t_{rel} steps.

Proof of $(*)$ Set $t = hit_{1-\varepsilon}(\varepsilon)$ and $s = \left\lceil 2\, t_{rel} \cdot \log\left(\frac{2}{\varepsilon^3}\right)\right\rceil$.

Want to show that $\forall\, x, A$

$$|P^{t+s}(x,A) - \pi(A)| \leq 2\varepsilon.$$

Idea: Want to define an intermediate set G s.t. we hit it whp before time t and conditional on hitting it by time t we want to be close to $\pi(A)$ at time $t+s$ by at most ε, i.e

$$|\mathbb{P}_x(X_{t+s} \in A \mid \tau_G \leq t) - \pi(A)| \leq \varepsilon.$$

If $\mathbb{P}_x(\tau_G > t) < \varepsilon$, then we will be done, because

$$|\mathbb{P}_x(X_{t+s} \in A) - \pi(A)| = |\mathbb{P}_x(X_{t+s} \in A \mid \tau_G \leq t)\, \mathbb{P}_x(\tau_G \leq t) +$$

$$+ \mathbb{P}_x(X_{t+s} \in A \mid \tau_G > t)\, \mathbb{P}_x(\tau_G > t) - \pi(A)|$$

$$\leq |\mathbb{P}_x(X_{t+s} \in A \mid \tau_G \leq t) - \pi(A)| + \mathbb{P}_x(\tau_G > t).$$

$$\leq \sup_{\substack{y \in G \\ r \geq s}} |\mathbb{P}_y(X_r \in A) - \pi(A)| + \mathbb{P}_x(\tau_G > t)$$

Define $G = \{y : \sup_{r \geq s} |\mathbb{P}_y(X_r \in A) - \pi(A)| \leq \varepsilon\}$.

Then

$$|\mathbb{P}_x(X_{t+s} \in A) - \pi(A)| \leq \varepsilon + \mathbb{P}_x(\tau_G > t).$$

So suffices to show that $\mathbb{P}_x(\tau_G > t) \leq \varepsilon$.

It suffices to prove that $\pi(G) > 1-\varepsilon$, since $t = hit_{1-\varepsilon}(\varepsilon)$.

NTS $\pi(G) > 1-\varepsilon$

$f : S \to \mathbb{R}$, define $\quad f^*(x) = \sup_{k \geq 0} |P^{2k} f(x)|$

$f_t(x) = P^t(x, A) - \pi(A) = P^t(\mathbb{1}(A) - \pi(A))\ (x)$

$f_t^*(x) = \sup_{k \geq 0} |P^{2k} f_t(x)| = \sup_{k \geq 0} |P^{2k+t}(x, A) - \pi(A)|$

$(Pf_t)^*(x) = \sup_{k \geq 0} |P^{2k} P f_t(x)| = \sup_{k \geq 0} |P^{2k+t+1}(x, A) - \pi(A)|$

$G = \{ y : \ f_s^*(y), (Pf_s)^*(y) \leq \varepsilon \}$

$\pi(G^c) \leq \pi(\{y : f_s^*(y) > \varepsilon\}) + \pi(\{y : (Pf_s)^*(y) > \varepsilon\})$

$\leq \dfrac{\mathbb{E}_\pi[(f_s^*)^2]}{\varepsilon^2} + \dfrac{\mathbb{E}_\pi[((Pf_s)^*)^2]}{\varepsilon^2}$ Markov's ineq. (***)

We write for $p \in (1, \infty)$ $\|f\|_p^p = \mathbb{E}_\pi[|f|^p]$

$\|Pf_s\|_2^2 \underset{\substack{\uparrow \\ P \text{ is a contraction}}}{\leq} \|f_s\|_2^2 = \mathrm{Var}_\pi(P^s(\mathbb{1}(A))) \underset{\substack{\uparrow \\ \text{Poincaré}}}{\leq} e^{-\frac{2s}{t_{rel}}} \mathrm{Var}_\pi(\mathbb{1}(A)) =$

substituting value of s

$= \dfrac{\varepsilon^3}{2} \cdot \pi(A) \cdot (1 - \pi(A)) \leq \dfrac{\varepsilon^3}{8}$

Theorem (Starr's maximal ineq.) P revers. wrt π, $p \in (1, \infty)$

$\forall f : S \to \mathbb{R}$

$$\|f^*\|_p \leq \frac{p}{p-1} \|f\|_p$$

$\|(Pf_s)^*\|_2^2 \leq 4 \cdot \|Pf_s\|_2^2 \leq \dfrac{\varepsilon^3}{2}$ and $\|f_s^*\|_2^2 \leq 4 \cdot \|f_s\|_2^2 \leq \dfrac{\varepsilon^3}{2}$.

Plug into (***) $\rightsquigarrow$ $\pi(G^c) < \dfrac{\varepsilon}{2} + \dfrac{\varepsilon}{2} = \varepsilon$. $\square$

Proof of Starr $f^*(x) = \sup_{k \geq 0} |P^{2k} f(x)|$

Let X be a MC with $X_0 \sim \pi$.

$P^{2n} f(X_0) = \mathbb{E}[f(X_{2n}) | X_0] \underset{\text{tower property}}{=} \mathbb{E}[\mathbb{E}[f(X_{2n}) | X_n, X_0] | X_0] =$

$= \mathbb{E}[\mathbb{E}[f(X_{2n}) | X_n] | X_0]$ (Markov property)

Set $R_n = \mathbb{E}[f(X_{2n}) | X_n]$

Goal: Show that R is a backwards martingale in other words we will show that if N is fixed, then

$(R_{N-n})_{0\le n\le N}$ is a martingale.

Since $X_0 \sim \pi$ and X is reversible, it follows that

$$(X_n, X_{n+1}, \dots, X_{2n}) \sim (X_n, X_{n-1}, \dots, X_0)$$

So $R_n = \mathbb{E}[f(X_{2n}) \mid X_n] = \mathbb{E}[f(X_0) \mid X_n] \underset{\text{Markov property}}{=} \mathbb{E}[f(X_0) \mid X_n, X_{n+1}, \dots]$

Set $\mathcal{F}_n = \sigma(X_n, X_{n+1}, \dots)$

and fix $N \ge 0$. Then $(R_{N-n})_{0\le n\le N}$ is a martingale wrt $(\mathcal{F}_n)$

$$\Big\| \max_{0\le n\le N} R_n \Big\|_p = \Big\| \max_{0\le n\le N} R_{N-n} \Big\|_p \underset{\text{Doob's } L^p\text{- ineq.}}{\le} \frac{p}{p-1} \cdot \|R_0\|_p = \frac{p}{p-1} \|f(X_0)\|_p = \frac{p}{p-1} \|f\|_p.$$

$$\Big| \max_{0\le n\le N} P^{2n} f(X_0) \Big| = \Big| \max_{0\le n\le N} \mathbb{E}[R_n \mid X_0] \Big| \le \mathbb{E}\Big[\Big| \max_{0\le n\le N} R_n \Big| \;\Big|\; X_0 \Big]$$

Conditional Jensen implies that

$$\Big\| \max_{0\le n\le N} P^{2n} f(X_0) \Big\|_p \le \Big\| \max_{0\le n\le N} R_n \Big\|_p \le \frac{p}{p-1} \cdot \|f\|_p$$

Letting $N \to \infty$ and using MONOTONE cvg completes the proof. $\square$