Equivalence of Starting Point Cutoff and the Concentration of Hitting Times on a General State Space: Web Appendix

David Ledvinka[*] Jeffrey S. Rosenthal[†]

A Detailed Proof of Starting Point Cutoff for a Toy Example

In this web appendix we give a detailed proof of the existence of the starting point cutoff phenomenon for a toy Markov chain. This Markov chain was used as an example in [1] for a different purpose. Consider the Markov kernel defined on $\mathbb{R}^d$ by

$$P(x, \cdot) \sim \mathrm{Normal}\left(\frac{x}{2}, \frac{3}{4} I_d\right)$$

where I_d is the d-dimensional identity matrix. One can show that P is a regular Markov kernel with stationary distribution $\pi \sim \mathrm{Normal}(0, I_d)$. In fact one can compute that

$$P^t(x, \cdot) \sim \mathrm{Normal}\left(\frac{x}{2^t}, (I - 4^{-t}) I_d\right)$$

Now let x_n be a sequence of starting points in $\mathbb{R}^d$ such that $\lim_{n\to\infty} \|x_n\| = \infty$. We will show that P has starting point cutoff at time $t_n = \log_2(\|x_n\|)$ starting from x_n. First let $c > 1$. For n sufficiently large such that $\|x_n\| \geq 1$ we have

$$\begin{aligned}
&d_{x_n}(ct_n)\\
&= \left\| P^{\lfloor ct_n \rfloor}(x_n, \cdot) - \pi \right\|_{TV}\\
&\leq \sqrt{\frac{1}{2} D_{KL}(P^{\lfloor ct_n \rfloor}(x_n, \cdot) || \pi)} \quad \text{by Pinsker's inequality}\\
&= \frac{1}{2}\sqrt{-d\log\left(1 - 4^{-\lfloor ct_n \rfloor}\right) - d4^{-\lfloor ct_n \rfloor} + \|x_n\|^2 4^{-\lfloor ct_n \rfloor}} \quad \text{by the formula for KL divergence of Gaussians}\\
&\leq \frac{1}{2}\sqrt{d(4^{-\lfloor ct_n \rfloor} + 16^{-\lfloor ct_n \rfloor}) - d4^{-\lfloor ct_n \rfloor} + \|x_n\|^2 4^{-\lfloor ct_n \rfloor}} \quad \text{by a Taylor series bounds on logarithm}\\
&= \left[\frac{1}{2}\sqrt{d4^{-\lfloor ct_n \rfloor} + \|x_n\|^2}\right] \left(\frac{1}{2}\right)^{\lfloor ct_n \rfloor}\\
&\leq \left(\frac{1}{2}\sqrt{d+1}\right) \|x_n\| \left(\frac{1}{2}\right)^{c\|x_n\| - 1}\\
&= 2\left(\frac{1}{2}\sqrt{d+1}\right) \|x_n\|^{1-c}
\end{aligned}$$

therefore $\lim_{n\to\infty} d_{x_n}(ct_n) = 0$. Now suppose $c < 1$. For $x \in \mathbb{R}^d$ and $r > 0$ define

$$A_x^r = \{y \in \mathbb{R}^d : y \cdot x \leq r\|x\|\}$$

[*]Department of Mathematics, University of Toronto, Canada. E-mail: david.ledvinka@mail.utoronto.ca
[†]Department of Statistical Sciences, University of Toronto, Canada. E-mail: jeff@math.toronto.edu

Then for any $x \in \mathbb{R}^d$ and $r > 0$

$$\begin{aligned}\pi(A_x^r) &= \mathbb{P}\left[W \cdot x \leq r\|x\|\right] \quad \text{where } W \sim \mathrm{Normal}(0, I_d) \\ &= \mathbb{P}\left[Z \leq r\right] \quad \text{where } Z \sim \mathrm{Normal}(0,1) \\ &= \Phi(r)\end{aligned}$$

and for $x \in \mathbb{R}^d$, $r > 0$ and $t \in \mathbb{N}$

$$\begin{aligned}P^t(x, A_x^r) &= \mathbb{P}\left[\left(\frac{x}{2^t} + (\sqrt{1-4^{-t}})W\right) \cdot x \leq r\|x\|\right] \quad \text{where } W \sim \mathrm{Normal}(0, I_d) \\ &= \mathbb{P}\left[Z \leq \frac{(r2^t - \|x\|)}{\sqrt{4^t - 1}}\right] \quad \text{where } Z \sim \mathrm{Normal}(0,1) \\ &= \Phi\left(\frac{r2^t - \|x\|}{\sqrt{4^t-1}}\right)\end{aligned}$$

Now define a sequence $r_n = \|x_n\|^{\frac{1-c}{2}}$. Then for each $n \in \mathbb{N}$,

$$\begin{aligned}d_{x_n}(ct_n) &= CPVersion)(Copy)\left\|P^{\lfloor ct_n \rfloor}(x_n, \cdot) - \pi\right\|_{TV} \\ &\geq |P^{\lfloor ct_n \rfloor}(x_n, A_{x_n}^{r_n}) - \pi(A_{x_n}^{r_n})| \\ &= \left|\Phi\left(\frac{r_n 2^{\lfloor ct_n \rfloor} - \|x_n\|}{\sqrt{4^{\lfloor ct_n \rfloor} - 1}}\right) - \Phi(r_n)\right|\end{aligned}$$

but since

$$\lim_{n\to\infty} \Phi(r_n) = 1$$

and

$$\frac{r_n 2^{\lfloor ct_n \rfloor} - \|x_n\|}{\sqrt{4^{\lfloor ct_n \rfloor} - 1}} \leq \|x_n\|^{\frac{1-c}{2}} - \|x_n\|^{1-c}$$

implies

$$\lim_{n\to\infty} \Phi\left(\frac{r_n 2^{\lfloor ct_n \rfloor} - \|x_n\|}{\sqrt{4^{\lfloor ct_n \rfloor} - 1}}\right) = 0$$

we have that

$$\lim_{n\to\infty} d_{x_n}(ct_n) = 1$$

This completes the proof that P has starting point cutoff at time $t_n = \log_2(\|x_n\|)$ starting from x_n.

References

[1] Qian Qin and James Hobert, *Wasserstein-based methods for convergence complexity analysis of mcmc with application to albert and chib's algorithm*, 10 2018.