1. 马氏链采样

给定 $f: R^{N_\theta} \to R$ 满足 $\mathrm{Var}_\rho[f] = 1$. 我们有

$$
\begin{aligned}
\mathbb{E}_{\theta_0\sim\rho}\left[\rho^N_{\mathrm{MCMC}}(f) - \mathbb{E}_\rho(f)\right] &= 0,\\
\mathbb{E}_{\theta_0\sim\rho}\left[\left(\rho^N_{\mathrm{MCMC}}(f) - \mathbb{E}_\rho(f)\right)^2\right] &= \frac{\tau_N^2(f)}{N},
\end{aligned}
$$

其中

$$
\tau_N^2(f) = 1 + 2\sum_{n=1}^{N-1}\frac{N-n}{N}\mathrm{Cov}\big(f(\theta_0), f(\theta_n)\big).
$$

证明

由于 $\theta_0 \sim \rho$，是马氏链的不变分布，我们有 $\theta_n \sim \rho$ 对于 $n \geq 1$。因此

$$
\mathbb{E}_{\theta_0\sim\rho}\left[\rho^N_{\mathrm{MCMC}}(f) - \mathbb{E}_\rho(f)\right] = 0
$$

对于方差，我们有

$$
\begin{aligned}
\mathbb{E}_{\theta_0\sim\rho}\left[\left(\rho^N_{\mathrm{MCMC}}(f) - \mathbb{E}_\rho(f)\right)^2\right] &= \frac{1}{N^2}\mathbb{E}_{\theta_0\sim\rho}\left[\Big(\sum_{n=1}^{N}\big(f(\theta_n) - \mathbb{E}_\rho(f)\big)\Big)^2\right]\\
&= \frac{1}{N^2}\left[\sum_{n=1}^{N}\mathrm{Var}\big[f(\theta_n)\big] + 2\sum_{n=1}^{N-1}\sum_{m>n}\mathrm{Cov}\big(f(\theta_n), f(\theta_m)\big)\right]\\
&= \frac{1}{N^2}\left[N + 2\sum_{n=1}^{N-1}\sum_{m>n}\mathrm{Cov}\big(f(\theta_0), f(\theta_{m-n})\big)\right]\\
&= \frac{1}{N^2}\left[N + 2\sum_{n=1}^{N-1}\sum_{k=1}^{N-n}\mathrm{Cov}\big(f(\theta_0), f(\theta_k)\big)\right]\\
&= \frac{1}{N^2}\left[N + 2(N-k)\sum_{k=1}^{N-1}\mathrm{Cov}\big(f(\theta_0), f(\theta_k)\big)\right]\\
&= \frac{1}{N}\left[1 + 2\frac{N-k}{N}\sum_{k=1}^{N-1}\mathrm{Cov}\big(f(\theta_0), f(\theta_k)\big)\right]\\
&= \frac{\tau_N^2(f)}{N}.
\end{aligned}
$$

θ_0 和 θ_n 越独立越好。

2. 马氏链收敛性

对于有限状态空间$S = \{1, 2, \cdots d\}$ 上的马氏链，假设马氏核 p 对应的转移概率矩阵$P \in R^{d\times d}$ 满足

$$
\min_{(i,j)\in S\times S} P_{i,j} \geq \frac{\epsilon}{d}
$$

那么存在唯一不变分布$\rho \in R^d$，满足

$$
\rho^T P = \rho^T
$$

对于马氏链$\{\theta_n\}, \theta_n \sim \rho_n$， 我们有

$$
\|\rho_n - \rho\|_{L_1} \leq 2(1-\epsilon)^n
$$

证明

由于 $P = \epsilon R + (P - \epsilon R)$， 其中 $R_{i,j} = \frac{1}{d}$。 这意味着在每一步中，总会存在一定的残余概率ϵR，使得马氏链可以跳转到任何新状态，与其当前位置无关。随着 n 的增加，每一步的残余概率将减小马氏链对θ_0 的依赖性。

基于这个观测，我们定义独立同分布的Bernoulli的随机变量 $\{b_n\}$，满足 $P(b_n = 1) = \epsilon$， $P(b_n = 0) = 1 - \epsilon$。以及转移概率矩阵

$$
W_{i,j} = \frac{P_{i,j} - \epsilon R_{i,j}}{1-\epsilon}.
$$

我们可以定义对应于转移概率矩阵W 和 R 的 马氏核 w 和 r 以及新的马氏链

$$
\theta'_{n+1} \sim \begin{cases} w(\theta'_n, \cdot), \text{ 当 } b_n = 0\\ r(\theta'_n, \cdot), \text{ 当 } b_n = 1\end{cases}
$$

容易验证 $P(\theta'_{n+1} = j | \theta'_n = i) = p(i,j)$。

对于任意测试函数 $f: S \to R$， 其中 $|f|_\infty \leq 1$。 我们定义

$$
\tau = \min\{n \in N : b_n = 1\}
$$

那么

$$
\begin{aligned}
\mathbb{E}_{\theta'_0\sim\rho_0}[f(\theta'_n)] &= \mathbb{E}_{\theta'_0\sim\rho_0}\left[f\left(\theta'_n\right)|\tau \geq n\right] P\left(\tau \geq n\right) + \sum_{l=0}^{n-1}\mathbb{E}_{\theta'_0\sim\rho_0}\left[f\left(\theta'_n\right)|\tau = l\right] P\left(\tau = l\right)\\
&= \underbrace{\mathbb{E}_{\theta'_0\sim\rho_0}\left[f\left(\theta'_n\right)|\tau \geq n\right] P\left(\tau \geq n\right)}_{|\cdot| \leq (1-\epsilon)^n} + \underbrace{\sum_{l=0}^{n-1}\mathbb{E}_{\theta'_0\sim\mathrm{uniform}(S)}\left[f\left(\theta'_{n-l-1}\right)\right] P\left(\tau = l\right)}_{\text{independent of original initial distribution}}
\end{aligned}
$$

我们有

$$
\begin{aligned}
\|\rho_n - \rho\|_{L_1} &= \sup_{|f|_\infty \leq 1}\left|\mathbb{E}_{\theta'_0\sim\rho_0}[f(\theta'_n)] - \mathbb{E}_{\theta'_0\sim\rho}[f(\theta'_n)]\right|\\
&= \sup_{|f|_\infty \leq 1}\left|\mathbb{E}_{\theta'_0\sim\rho_0}\left[f\left(\theta'_n\right)|\tau \geq n\right] P\left(\tau \geq n\right) - \mathbb{E}_{\theta'_0\sim\rho}\left[f\left(\theta'_n\right)|\tau \geq n\right] P\left(\tau \geq n\right)\right|\\
&= \sup_{|f|_\infty \leq 1}\left|\left(\mathbb{E}_{\theta'_0\sim\rho_0}\left[f\left(\theta'_n\right)|\tau \geq n\right] - \mathbb{E}_{\theta'_0\sim\rho}\left[f\left(\theta'_n\right)|\tau \geq n\right]\right) P\left(\tau \geq n\right)\right|\\
&\leq 2(1-\epsilon)^n
\end{aligned}
$$

In [ ]: