Adaptive MCMC

General: π target dist.

$\{P_\theta, \theta \in \Theta\}$ a collection of Markov kernels

Such that $\pi P_\theta = \pi \ \forall \theta \in \Theta$

Possibility: choose any $\theta_0 \in \Theta$

and run

Alg: Given X_n draw $X_{n+1} \sim P_{\theta_0}(X_n, \cdot)$

Adaptive MCMC: Given $x_n, \theta_n = \theta$

draw $X_n \sim P_\theta(x, \cdot)$

update $\theta_{n+1} = \theta_n + \gamma H(\theta_n, X_{n+1})$

if we choose

$$\boxed{\theta = \underset{\tilde{\theta}}{\operatorname{argmax}} \int H(\tilde{\theta}, x)\, \pi(dx)}$$

gives the "best" choice and

$\gamma_n \to 0$ s.t. $\sum_{n=1} \gamma_n = +\infty$

<u>Ex:</u> (R.W.M.)

$X_n = x$

$Y \sim N(x, \Sigma)$

$$X_{n+1} = \begin{cases} Y & \min\left(\frac{\pi(Y)}{\pi(x)}, 1\right) \\ x & \text{o.w.} \end{cases}$$

Given X_n, Σ_n, μ_n

$X_n \sim P_{\Sigma_n}(x_n, \cdot)$

$$\mu_{n+1} = \mu_n + \frac{1}{n+1}(X_n - \mu_n)$$

$$\Sigma_{n+1} = \Sigma_n + \frac{1}{n+1}\left[(X_{n+1} - \mu_{n+1})(X_{n+1} - \mu_{n+1})^T - \Sigma_n\right]$$

<u>Another class of</u>

<u>Adaptive MCMC</u> $*$

(N.N.) $(y_i, x_i)_1^n$, $y_i \in \mathbb{R}$ $x_i \in \mathbb{R}^{p-1}$

$$Y_i = f_\theta(x_i) + \varepsilon_i \sim N(0, e\ )$$

$$f_\theta(x) = \sum_{k=1}^{K} b_k\, s\left(a_{0k} + \langle a, x \rangle\right)$$

$$\theta = \left((b_0, a_{01}, a_1), \ldots, (b_k, a_{0k}, a_k) \right) \in \mathbb{R}^{K(p+1)}$$

<u>Prior</u>: $$p(\theta | s_1) \equiv \prod_{k=1}^{K} p(b_k | s_1)\, p(a_{0k} | v_2)\, p(a | v_2)$$

<u>Post</u>: $$\pi(\theta | y_{1:n}, v_1, v_2) =$$

$$p(\theta | v_2) \prod_{i=1}^{n} e^{-v_1/2}\, e^{-v_2/2\, (y_i - f_\theta(x_i))^2}$$

<u>Q</u>: Choosing (v_1, v_2)

<u>A</u>: 1. Priors. 2. Empirical Bayes*

*Estimate $\hat{v}_1$ $\hat{v}_2$ from data.

<u>How</u>? Find $f(y_{1:n} | v_1, v_2)$ and do MLE

$$(\hat{v}_1, \hat{v}_2) = \underset{(v_1, v_2)}{\arg\max}\ f(y_{1:n} | v_1, v_2)$$

$$f(y_{1:n} | v_1, v_2) = \int f(y_{1:n} | \theta, v_1, v_2)\, p(\theta | v_1, v_2)\, d\theta$$

intractable

But we really care about

$$\frac{\partial}{\partial v_1} f(y_{1:n} | v_1, v_2) \quad \text{or} \quad \frac{\partial}{\partial v_1} \log f(y_{1:n} | v_1, v_2)$$

$$\frac{\partial}{\partial v_1} \ell(v_1, v_2)$$

$$= \int \left[\frac{\partial}{\partial v_1} \log f(y_{1:n} | \theta, v_1) \right] \boxed{\frac{f(y_{1:n} | \theta, v_1)\, p(\theta | v_2)}{\int f(y | u, v_1)\, p(u | v_2)\, du}}\, d\theta$$

density $f = \pi(\theta \mid y_{1:n}, \nu_1, \nu_2)$

$$\frac{\partial}{\partial \nu_1} \ell(\nu_1, \nu_2) = \int \left[\frac{d}{d\nu_1} \log f(y_{1:n} \mid \theta, \nu_1) \right] \pi(\theta \mid y_{1:n}, \nu_1, \nu_2) \, d\theta$$

$$= \mathbb{E}_{\pi_{\theta \mid y_{1:n}, \nu_1 \nu_2}} \left[\partial / \partial \nu_1 \log f \right]$$

$$\frac{\partial}{\partial \nu_2} \ell(\nu_1, \nu_2) = \int \left[\frac{d}{d\nu_2} \log f(y_{1:n} \mid \theta, \nu_1) \right] \pi(\theta \mid y_{1:n}, \nu_1, \nu_2) \, d\theta$$

$$= \mathbb{E}_{\pi_{\theta \mid y_{1:n}, \nu_1 \nu_2}} \left[\partial / \partial \nu_2 \log f \right]$$

<u>Idea:</u> Plug-in estimator for π.

<u>Alg:</u> · Given (ν_1, ν_2) let $P_{(\nu_1, \nu_2)}$ be MK.

with inv. dist. $\pi(\theta \mid y_{1:n}, \nu_1, \nu_2)$

· Given $(\theta^{(k)}, \nu_1^{(k)}, \nu_2^{(k)})$ then

$\theta^{(k+1)} \sim P \quad (\theta^{(k)}, \cdot)$

$\theta \sim \pi_{v_1^{(k)} v_2^{(k)}}$

Partial derivatives

$$S_1 = -\frac{n}{2} + \frac{e^{-v_1^{(k)}}}{2} \sum_{i=1}^{n} \left(y_i - f_{\theta^{(k+1)}}(x_i)\right)^2$$

$$S_2 = \frac{-K(p+1)}{2} + \frac{e^{-v_2^{(k)}}}{2} \sum_{i=1}^{K} \left(b_k^2 + a_{0k}^2 + \|a_k\|_2^2\right)$$

$$v_1^{(k+1)} = v_1^{(k)} + \gamma_{k+1} S_1$$

$$v_2^{(k+1)} = v_2^{(k)} + \gamma_{k+1} S_2$$

NR updates.

Connection with EM

$$Q(v_1, v_2 \mid \bar{v}_1, \bar{v}_2) = \int \log\left[f(y_{1:n} \mid \theta, v_1, v_2)\, p(\theta \mid v_1, v_2)\right] \pi(\theta \mid y_{1:n}, \bar{v}_1, \bar{v}_2)\, d\theta$$

EM: Given $(v_1^{(k)}, v_2^{(k)})$

$$(v_1^{(k+1)}, v_2^{(k+1)}) = \underset{v_1, v_2}{\arg\max}\ Q(v_1, v_2 \mid v_1^{(k)}, v_2^{(k)})$$

Problem: Both steps intractable.

↳ approximation leads to our solution.