Homework Notes

$$Y_i \sim \text{Bern}\left(\frac{e^{f_\theta(x_i)}}{1+e^{f_\theta(x_i)}}\right), \quad i = 1, 2, \dots, n$$

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

$$\pi(\theta | y) \propto p(\theta) \exp(\ell(\theta; y))$$

$$\ell(\theta, y) = \sum_{i=1}^{n} \left(y_i f_\theta(x_i) - \log(1 + \exp(f_\theta(x_i))) \right)$$

Computation of $\ell(\theta, y)$:

$$\theta = [b, a_0, a] \in \mathbb{R}^{K \times (p+2)}$$

$$y \in \mathbb{R}^n, \; X \in \mathbb{R}^{n \times p}$$

$$b^T s(\overline{a_0} + a X^T) \Rightarrow (f_\theta(x_1), \dots, f_\theta(x_n))$$

Metropolis Hastings:

Given θ, $\log \pi$

1. Update θ_i by proposing $\overline{\theta}_i = \theta_i + \sigma_i^{1/2} Z$
2. Set $\overline{\theta} = \theta$, $\overline{\theta}_i = \overline{\theta}_i$
3. Compute $\text{logpi\_prop} = \ell(\overline{\theta}, y)$
4. If $u \leq \min\{1, \exp[\theta_i^2 - \overline{\theta}_i^2 + \text{log\_pi\_prop} - \text{log\_pi}]\}$

a) $\theta = \bar{\theta}$

b) log_p_i = log_p_i_prp

Setting up the Temperatures for P.T.

Choose the temperatures temp_val = $(1, 1.1, 1.2, 1.3, 2, \dots, 5, 10, 15, 20)$

Typically want dense temps near 1.

Compute acceptance for each swap level.

Ex: $(.2, \underbrace{.001}_{\text{probably want to add}}, \overbrace{.73}^{\text{probably want to remove}}, \dots)$ want to choose temps to stabilize the acceptance probability.

Rmk: The model is not identifiable. The rows of θ are completely row invariant. So the alg. will try to explore several local modes all corresponding to the same equivalence class.

Unbiased MCMC

Standard MCMC: $(X_n)_{n=1}^{\infty}$, s.t. $\mathbb{P}(X_n \in A) \longrightarrow \pi(A)$

but for any fixed n, $\mathbb{P}(X_n \in A) \neq \pi(A)$

Cannot average MCMC output across cores because of this bias.

<u>Rmk</u>: Unbiased MCMC will yield more variance.

We will not have $\mathcal{L}X_n = \pi$ but instead

$$\mathbb{E}[h(X_\tau)] = \int h(x)\pi(x)dx, \quad h \in \mathcal{H}, \ \tau \ \text{a.s.}$$

<u>Simple Problem</u> Suppose you want to compute

$$x = \sum_{k \geq 0} x_k, \ x_k \in \mathbb{R}$$

<u>Idea</u>: Truncation $\Rightarrow$ Bias

Unbiased Estimate? Let $q = \{q_k, k \geq 0\}$ a dist.

Draw $K \sim q$. $X = \frac{x_K}{q_K}$ $\mathbb{E}[X] = \sum_{k=0}^{\infty} \frac{x_k}{q_k} q_k = x$

but $\mathbb{E}(X^2) = \sum_{k=0}^{\infty} \frac{x_k^2}{q_k}$ can be infinite if q isn't carefully chosen

<u>Other Estimators</u>: $N \sim q$, $\hat{X} = \sum_{k=0}^{N} \frac{x_k}{\mathbb{P}(N \geq k)}$

$$\mathbb{E}(\hat{X}) = \mathbb{E}\left[\sum_{k \geq 0} \frac{x_k \mathbb{1}(N \geq k)}{\mathbb{P}(N \geq k)}\right]$$

If $\sum |x_k| < +\infty$ $\mathbb{E}(\hat{X}) = \sum_{k \geq 0} x_k$

<u>Application</u>: $\pi(h) = \lim_{n \to \infty} \mathbb{E}[h(X_n)]$

Define $\Delta_0 = h(X_0)$, $\Delta_k = h(X_k) - h(X_{k-1})$, $k \geq 1$

$$\sum_{k=0}^{n} \mathbb{E}(\Delta_k) = \mathbb{E}[h(X_0)] + \sum_{k\geq 1}^{n} \mathbb{E}[h(X_k) - h(X_{k-1})]$$

$$= \mathbb{E}(h(X_n))$$

$$\pi(h) = \lim_{n \to \infty} \sum_{k=0}^{n} \mathbb{E}[\Delta_k] = \sum_{k\geq 0} \mathbb{E}(\Delta_k)$$

$$\stackrel{?}{=} \mathbb{E}\left[\sum_{k\geq 0} h(X_k)\right]$$