

Pseudo-marginal MCMC

The Challenge: Intractable Marginals

The Challenge:

- ▶ Target: $\pi(x) = \int f(x, u) du$
- ▶ $f(x, u)$ is known, e.g. model with latent variables
- ▶ Integral is **intractable**
- ▶ Standard MCMC requires exact $\pi(x)$

The Pseudo-marginal Solution

Key Prerequisites

For pseudo-marginal MCMC to be applicable, we need:

1. Ability to **evaluate** $f(x, u)$ pointwise for any (x, u)
2. Ability to **sample** from an importance distribution $q_x(\cdot)$ over the u -space
3. The importance distribution must have appropriate support: $q_x(u) > 0$ whenever $f(x, u) > 0$

Importance Sampling Estimator:

$$\hat{\pi}(x) = \frac{1}{N} \sum_{i=1}^N \frac{f(x, U_i)}{q_x(U_i)}, \quad U_i \sim q_x(\cdot)$$

Key Property: $\mathbb{E}[\hat{\pi}(x)] = \pi(x)$ (unbiased estimator for $\hat{\pi}(x)$)

The Magic: Replace π with $\hat{\pi}$ in MH acceptance ratio!

$$\alpha = \min \left\{ 1, \frac{\hat{\pi}(y)q(y, x)}{\hat{\pi}(x)q(x, y)} \right\}$$

Result: Still targets correct $\pi(x)$!

Why It Works: Extended Target

One can think of estimator (the “pseudo-marginal”) as the product of the true target and a random variable:

$$\hat{\pi}(x) = \pi(x)Z_x$$

where Z_x satisfies:

1. is non-negative: $Z_x \geq 0$,
2. has density $g_x(\cdot)$: $\int_0^\infty g_x(z)dz = 1$
3. has expectation 1: $\mathbb{E}[Z_x] = 1$.

Why It Works: Extended Target

Extended Target Construction:

$$\bar{\pi}(x, z) = \pi(x) \cdot z \cdot g_x(z)$$

where $g_x(z)$ is the density of Z_x

Key Property: extended target marginalizes to original target:

$$\int \bar{\pi}(x, z) dz = \pi(x)$$

Now apply Metropolis–Hastings with proposal

$$\bar{q}((x, z), (y, w)) := q(x, y) \cdot g_y(w).$$

Intuition:

- ▶ Run exact MCMC on (x, z) space
- ▶ Marginal in x gives correct target
- ▶ z represents the "noise" in estimates

Equivalence to MH on Extended Space

Theorem (Equivalence)

Metropolis-Hastings on the extended target $\bar{\pi}$ with proposal \bar{q} is equivalent to the pseudo-marginal algorithm using estimates $\hat{\pi}$.

Proof Sketch: The MH acceptance ratio on the extended space is:

$$\begin{aligned}\alpha_{\text{ext}} &= \min \left\{ 1, \frac{\bar{\pi}(y, w) \bar{q}((y, w), (x, z))}{\bar{\pi}(x, z) \bar{q}((x, z), (y, w))} \right\} \\ &= \min \left\{ 1, \frac{\pi(y) \cdot w \cdot g_y(w) \cdot q(y, x) \cdot g_x(z)}{\pi(x) \cdot z \cdot g_x(z) \cdot q(x, y) \cdot g_y(w)} \right\} \\ &= \min \left\{ 1, \frac{\pi(y) \cdot w \cdot q(y, x)}{\pi(x) \cdot z \cdot q(x, y)} \right\} = \min \left\{ 1, \frac{\hat{\pi}(y) q(y, x)}{\hat{\pi}(x) q(x, y)} \right\} = \alpha_{pm}\end{aligned}$$

In the last step, we used $\hat{\pi}(x) = \pi(x)z$ and $\hat{\pi}(y) = \pi(y)w$, which is exactly the pseudo-marginal acceptance probability.

Pseudo-marginal MCMC Algorithm

Given $(X^{(t-1)}, \hat{\pi}^{(t-1)})$:

1. **Propose:** $Y \sim q(X^{(t-1)}, \cdot)$

2. **Estimate:**

- ▶ Sample $U_i \sim q_Y(\cdot)$
- ▶ $\hat{\pi}(Y) = \frac{1}{N} \sum_i \frac{f(Y, U_i)}{q_Y(U_i)}$

3. **Accept with probability:**

$$\alpha = \min \left\{ 1, \frac{\hat{\pi}(Y) q(Y, X^{(t-1)})}{\hat{\pi}^{(t-1)} q(X^{(t-1)}, Y)} \right\}$$

4. **Update:**

- ▶ If accept: $(X^{(t)}, \hat{\pi}^{(t)}) = (Y, \hat{\pi}(Y))$
- ▶ Else: $(X^{(t)}, \hat{\pi}^{(t)}) = (X^{(t-1)}, \hat{\pi}^{(t-1)})$

Critical Points:

- ▶ Store estimates with states! In the next iteration, use the stored $\hat{\pi}(X^{(t-1)})$.
- ▶ Works with *any* MH proposal q
- ▶ Small N is cheap to compute but gives large variance and poor mixing;
- ▶ Large N is expensive but improves mixing;

$$\hat{\omega}(x) = \frac{1}{N} \sum_{i=1}^N \frac{f(x, U_i)}{q_x(U_i)}, \quad U_i \stackrel{\text{i.i.d.}}{\sim} q_x(\cdot)$$

$$\mathbb{E}[\hat{\omega}(x)] = \mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N \frac{f(x, U_i)}{q_x(U_i)} \right]$$

$$= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\frac{f(x, U_i)}{q_x(U_i)} \right]$$

$$\mathbb{E} \left[\frac{f(x, U_i)}{q_x(U_i)} \right] = \int \frac{f(x, u)}{q_x(u)} \cdot q_x(u), du = \int f(x, u), du$$