Simulated Tempering

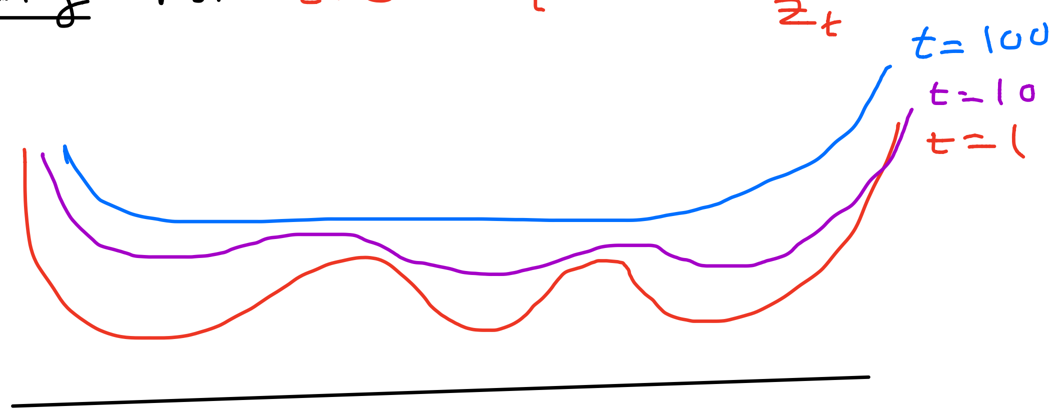
$\mathcal{X} = \mathbb{R}^p \qquad \pi(x) = \frac{1}{Z} e^{-U(x)} \qquad Z = \int e^{-U(x)} dx$

Goal: Sample from π

Issue: U may be multimodal

↳ MC may get stuck causing a slow or inconsistent sample.

Tempering For $t > 0$ $\quad \pi_t(x) \propto \frac{1}{Z_t} e^{-U(x)/t}$



Rmk: For $t \gg 1$ sampling from π_t is very easy.

Simulated Tempering Fix $1 = t_1 < t_2 < \cdots < t_K$

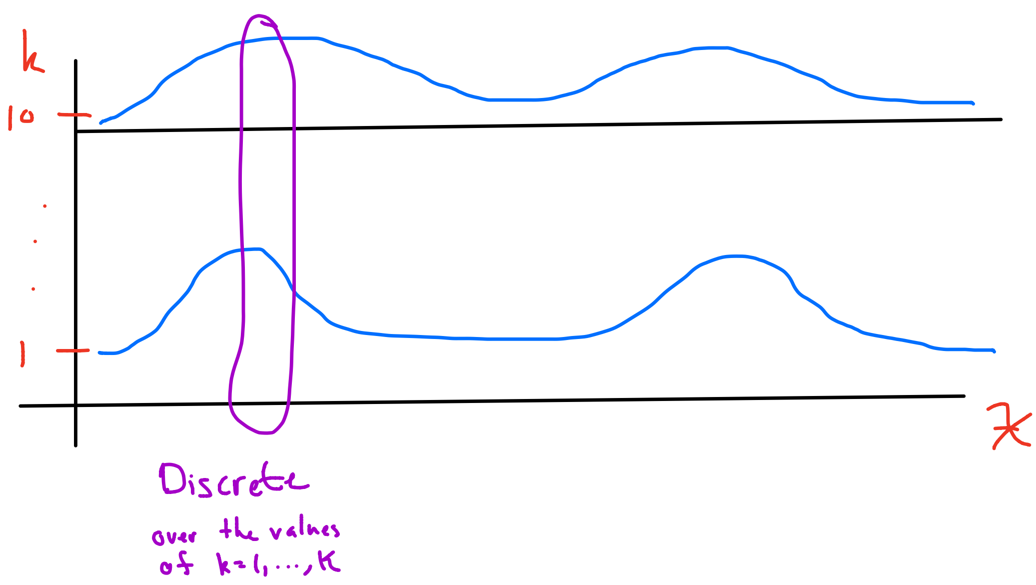
and set $\pi_k \equiv \pi_{t_k}$. Would like to sample

$$\bar{\pi}(k, x) = \frac{\frac{1}{c_k} e^{-U(x)/t_k}}{\sum_{k=1}^{K} \frac{1}{c_k} \int e^{-U(x)/t_k} \, dx} \qquad k = 1, \ldots, K$$

for some positive constants $\{c_k\}_{k=1}^{K}$

Rmk: For some fixed k, marginally

$$\boxed{\bar{\pi}(x \mid k) = \pi_k(x)}$$



$$\pi(k|x) = \frac{\frac{1}{c_k} e^{-U(x)/t_k}}{\sum_{j=1}^{K} \frac{1}{c_j} e^{-U(x)/t_j}}$$

This inspires a Gibbs type sampler.

<u>Rmk</u>: Use extra dimension to jump far in the $*$ space.

<u>Issues</u>

1. Choice of $\{c_k\}_{k=1}^{K}$

2. Sampling from $\pi(x|k)$ still difficult.

<u>Note</u>: $\pi(k) = \int \pi(k,x)\, dx$

$$= \frac{\frac{1}{c_k} e^{-U(x)/t_k}}{\sum_{j=1}^{K} \frac{1}{c_j} \int e^{-U(x)/t_j} \, dx}$$

Rmk: $c_k \equiv 1 \quad \forall k$

$$\Longrightarrow \quad \pi(k) \propto Z_k \sim \exp(t_k)$$

Choosing c_k

Adaptive approach, based on quality of chain exploration at level k.

One idea:

Let P_k be a Markov Kernel w. initial distribution π_k

Then adapted simulated tempering (AST) can be written as.

Algo: Given $(k^{(j)}, X^{(j)})$ and $C_{1:K}^{(j)}$

- $X^{(j+1)} \sim P_{k^{(j)}}(X^{(j)}, \cdot)$
- $k^{(j+1)} \sim \left\{ \frac{e^{-u(X^{(j+1)})/t_k}}{C_k^{(j)}},\ k = 1, \dots, K \right\}$
- $C_k^{(j+1)} = \begin{cases} C_k^{(j)}+1 & k^{(j+1)} = k \\ C_k^{(j)} & \text{o.w.} \end{cases}$

Last step downweights current marginal.

Encourages more exploration.

Rmk: Only accept samples from π_1. π_k just helps us jump around the target quickly.

Rmk: Closely related to Wang-Landau.

New space: $\bigcup_{k=1}^{K} \{k\} \cup \mathcal{X}$ and reweight each piece.

Parallel Tempering (Replica Exchange)

$$\overline{\pi}(x_1, \dots, x_K) \propto \prod_{k=1}^{K} \pi_k(x_k) \quad \text{where}$$

$$\pi_k(x) \propto e^{-U(x)/t_k}$$

Algo: Given $(x_1^{(j)}, \dots, x_K^{(j)})$

- For $k = 1, \dots, K$

  $\overline{X}_k \sim P_k(X_k^{(j)}, \cdot)$

- Randomly draw $I \sim U_{1:(K-1)}$
- With prob α,

  $X_I^{(j+1)} = \overline{X}_{k+1} \quad X_{I+1}^{(j+1)} = \overline{X}_I, \; X_\ell^{(j+1)} = \overline{X}_\ell$

- With prob $1 - \alpha$

  $X_\ell^{(j+1)} = \overline{X}_\ell \quad \forall \ell.$

$$\alpha = \min\left[1, \frac{\pi_{I+1}(x_I)\, \pi_I(x_{I+1})}{\pi_I(x_I)\, \pi_{I+1}(x_{I+1})}\right]$$

$$= \min \left\{ 1, \ \frac{e^{u(x_I)\left(\frac{1}{t_I} - \frac{1}{t_{I+1}}\right)}}{e^{u(x_{I+1})\left(\frac{1}{t_I} - \frac{1}{t_{I+1}}\right)}} \right]$$

Justification

(Deterministic MH)

Fix $\mathcal{X} \equiv \mathbb{R}^p$ $\quad T: \mathbb{R}^p \longmapsto \mathbb{R}^p$ that is invertible

with $T^{-1} = T$

Fix π on $\mathbb{R}^p$.

Alg: Given $X^{(k)} = x$ propose $Y = T(X^{(k)})$

$$X^{(k+1)} = \begin{cases} Y & \alpha \\ X^{(k)} & 1-\alpha \end{cases}$$

$$\alpha = \min \left\{ 1, \ \frac{\pi(y)}{\pi(x)} \quad \frac{1}{|J_T(x)|} \right\}$$

Prop: If $X^{(0)} \sim \pi$ $X^{(1)} \sim \pi$.

Rmk: $J_T(x) = \det(\nabla T(x))$