1 Coding (Double-well potential revisited).

Apply parallel tempering to sample from a double-well potential

\[ \pi(x) \propto \exp\left(-x^2 - 4\right). \]

What can you do to optimize? How often does the sampler transition from the right well (centered at \(x = +2\)) into the left well (centered at \(x = -2\))? 

2 Coding (Harlem Shake).

Apply parallel tempering or the Goodman-Weare algorithm to sample from one of the letters in “HARLEM SHAKE”. You need to come up with a formula for the 2-d density and sample from it.

3 Coding (The banana distribution).

The Rosenbrock function is a banana-shaped function that is hard to optimize. It is defined as

\[ H(x, y) = (x - 1)^2 + 100(y - x^2)^2. \]

We can obtain a hard sampling problem by targeting the density \(\pi(x, y) \propto \exp(-H(x, y))\). Can you use HMC to effectively sample from the Rosenbrock density? Can you use parallel tempering?

4 Coding (Goodman-Weare in high dimensions).

Apply the Goodman-Weare algorithm to sample from a Gaussian \(\mathcal{N}(0, I)\) distribution in \(\mathbb{R}^d\).

(a) What is the minimum number of ensemble members to ensure a valid sampler? How does the minimum number of ensemble members depend on \(d\)?

(b) Plot the acceptance rate as a function of \(d\).

5 Math (Optimization of parallel tempering).

Parallel tempering generates a vector of samples \(X_t = (X^{(1)}_t, \ldots, X^{(N)}_t)\), with each \(X^{(i)}_t\) targeting a different density

\[ \pi_i(x) = \pi(x)^{t_i/T}, \]

defined by temperatures \(t_1 > t_2 > \cdots > t_N = T\).

(a) If we propose a swap move \((X^{(i)}_t, X^{(i+1)}_t) \leftrightarrow (X^{(i+1)}_t, X^{(i)}_t)\), what is the proper Metropolis-Hastings acceptance probability so that \(X_t\) is a time-reversible sampler with respect to the product density \(\pi_1(x^{(1)}) \cdots \pi_N(x^{(N)})\)?

(b) What is the average acceptance rate for a swap move? Write down an integral formula.

(c) When \(\pi\) is a multivariate Gaussian \(\mathcal{N}(\mu, \Sigma)\), write down the acceptance probability and acceptance rate explicitly. How can we define temperatures \(t_1 > t_2 > \cdots > t_N = T\) so that each swap move has the same acceptance rate for \(i = 1, 2, \ldots, N - 1\)?
6 Math (Acceptance rate for Goodman-Weare sampler).

The Goodman-Weare sampler is a gradient-free ensemble sampler $X_t = (X^{(1)}_t, \ldots, X^{(N)}_t)$, with each $X^{(i)}_t$ targeting the same density

$$\pi(x), \quad x \in \mathbb{R}^d. \quad (1)$$

The sampler is designed so that $X_t$ that targets the product density $\pi(X^{(1)}_t) \cdots \pi(X^{(N)}_t)$. In the standard formulation, the Goodman-Weare sampler applies a “stretch move" where we move an ensemble member $X^{(i)}_t$ toward or away from another ensemble member $X^{(j)}_t$:

$$X^{(i)}_t \leftarrow X^{(i)}_t + (1 - U^{2/3})(X^{(j)}_t - X^{(i)}_t), \quad U \sim \text{Unif}(1/\sqrt{8}, \sqrt{8}).$$

(a) Write down the transition density $T(x, y) dy$ for the proposed stretch move and argue it is symmetric $T(x, y) = T(y, x)$.

(b) What is the proper Metropolis-Hastings acceptance probability for the stretch move? Write down a justification using spherical coordinates.

(c) What if we replaced $U^{2/3}$ with another random variable $V$ supported on $(1/2, 2)$ with density $f(v)$? What would be the proper Metropolis-Hastings acceptance probability then?