1 Coding (MCMC for a Gaussian).

Design a Metropolis-Hastings sampler for the $\mathcal{N}(0,1)$ distribution based on local perturbations. Does the sampler give convergent estimates of $\mathbb{E}[Z \mid Z \sim \mathcal{N}(0,1)]$? What can you do to optimize?

2 Coding (MCMC for two Gaussians).

Design a Metropolis-Hastings sampler for a mixture of two Gaussians: $\frac{1}{2}\mathcal{N}(-2,1)+\frac{1}{2}\mathcal{N}(2,1)$. Check that the sampler gives convergent estimates of the mean. What can you do to optimize?

3 Coding (MCMC error bars).

Evaluate the bias and the variance of an MCMC estimate. Which is bigger?

4 Choose-your-own-adventure (MCMC for many Gaussians).

Consider the Gaussian random field $\boldsymbol{Z} = (Z_1, \ldots, Z_N)^T$ with density

$$\pi(\boldsymbol{z}) = \frac{1}{\mathcal{Z}_\beta} \exp\left(-\frac{\beta}{2} \sum\nolimits_{i \sim j} (z_i - z_j)^2\right).$$

The summation runs over neighbors $i \sim j$, e.g., $|i-j| = 1$ for the 1-d chain with an open boundary.

(a) Write down an exact sampler for the 1-d chain that samples Z_1, then Z_2, then Z_3, etc.

(b) Write down a Gibbs sampler that updates Z_i by sampling from $\pi(\cdot | Z_j,\ j \neq i)$.

(c) Write down a Gibbs sampler that updates $Z_i \leftarrow Z_i + \delta$ for all i in a block S of variables. Hint: the density of δ only depends on Z_j for $j \notin \mathsf{S}$ and the acceptance probability is 1.

5 Choose-your-own-adventure (independence sampler).

Consider an "independence" sampler for the density π where we propose a random transition $X \to Y$, with Y randomly drawn from the density $g(y)$ independent of the starting point X.

(a) Write down a general formula for the acceptance probabilities.

(b) Write down a formula for the acceptance probabilities when the target is a Gaussian $Z \sim \mathcal{N}(0,1)$ conditional on $Z > 10$ and we propose $Y \sim \mathcal{N}(10,1)$.

(c) Identify the "small" sets that satisfy the one-step minorization condition. If the whole space $\mathcal{X}$ is a small set, write down a simple geometric bound on $\rho_{s,f} = \mathrm{Corr}[f(X_0), f(X_s) | X_0 \sim \pi]$.

6 Math (Spectral computations).

(a) Assume detailed balance, $\pi(dx)p(x,dy) = \pi(dy)p(y,dx)$ for all $x, y \in \mathcal{X}$, and show that the forward operator $[\mathbb{P}f](x) = \int p(x,dy)f(y)$ is self-adjoint in $L^2(\pi)$.

(b) Bound $\rho_{s,f} = \mathrm{Corr}[f(X_0), f(X_s)|X_0 \sim \pi]$ using a general formula involving $\mathbb{P}$.

(c) Bound $\rho_{s,f}$ for the single-flip update sampler for the Ising model with $\beta = 0$. Hint: every flip is accepted, and the eigenvectors are $f_{\mathsf{S}}(\boldsymbol{\sigma}) = (-1)^{\sum_{i \in \mathsf{S}} \mathbb{1}\{\sigma_i = +1\}}$.

(d) Bound $\rho_{s,f}$ for the autoregressive process $X_t = \sqrt{1-\alpha}X_{t-1} + \sqrt{\alpha}Z_t$ where $Z_t \sim \mathcal{N}(0,1)$.

(e) If kernels p and q satisfy detailed balance with the same stationary distribution and $q(x,dy) \geq p(x,dy)$ for all $x \neq y$, show that q gives smaller correlations $\rho_{f,1}$ for all f (Peskun ordering).

7 Math (uncountable ergodic theorem).

To complete the proof of the uncountable ergodic theorem, fill in the missing steps 2(a)-(b).

1. We assume a one-step minorization condition

$$p(x,dy) \geq \delta\mu(dy), \quad x \in \mathsf{A},$$

involving a "small" set $\mathsf{A} \in \mathcal{X}$, a constant $\delta > 0$, and a probability measure μ. This allows us to construct a split-chain $X'_t = (X_t, R_t)$ with transition probabilities

$$\begin{cases} q((x,r),(dy,0)) = p(x,dy), & x \notin A, \\ q((x,r),(dy,0)) = p(x,dy) - \delta\mu(dy), & x \in A, \\ q((x,r),(dy,1)) = \mu(dy), & x \in A. \end{cases}$$

2. We define renewal times for the split chain as

$$t_1 = \min\{t > 0 : R_t = 1\}, \quad t_i = \min\{t_i > t_{i+1} : R_t = 1\}.$$

We then assume a Foster-Lyapunov drift criterion

$$\int p(x,dy)V(y) \leq (1-\lambda)V(x)\mathbb{1}\{x \in \mathsf{A}\} + b\mathbb{1}\{x \notin \mathsf{A}\}$$

for a function $V : \mathcal{X} \mapsto [1,\infty)$ and constants $\lambda > 0$ and $b > 0$.

(a) Using the drift criterion, prove that each $t_i < \infty$.

(b) Using the drift criterion, prove that $\mu = \mathbb{E}[t_1|X_0 \sim \mu, R_0 = 1] < \infty$. Hint: first show

$$\frac{\delta\lambda}{b} \leq \liminf_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \mathbb{P}\{R_s = 1\}.$$

Then set $N_t = \sum_{s=1}^{t} \mathbb{1}\{R_s = 1\}$ and argue using the strong law of large numbers that $\lim_{t\to\infty} N_t/t = 1/\mu$ and consequently $\lim_{t\to\infty} \mathbb{E}[N_t]/t = 1/\mu$.

3. For any bounded f, we conclude that the segments $\sum_{t=t_i}^{t_{i+1}} f(X_t)$ are iid with finite mean and

$$\frac{1}{T}\sum_{t=0}^{T-1} f(X_t) \overset{T\to\infty}{\to} \frac{1}{\mu}\mathbb{E}\left[\sum_{t=0}^{t_1-1} f(X_t) \,\middle|\, X_0 \sim \mu\right].$$