Statistical Models and Computing Methods, Problem Set 2

October 26, 2025 Due 11/10/2025

Problem 1.
The standard Laplace distribution has density

$$f(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < \infty$$

(1) Describe how to generate a standard Laplace random variable by inverting the CDF. *(5 points)*
(2) Describe and implement a rejection sampling algorithm to simulate random draws from the standard normal distribution using (a multiple of) the Laplace density as the envelop function. Hint: how do you choose the constant multiple to make sure that this is a valid envelop? *(10 points)*
(3) Can one simulate Laplace random variables using rejection sampling with a multiple of the standard normal density as the envelop? Why or why not? *(5 points)*

Problem 2.
Consider a target distribution of the following form

$$\pi(x) = \int \pi(x, z)\, dz$$

(1) Let K be a positive integer. Show that the following estimate of $\pi(x)$ is unbiased *(5 points)*

$$\hat{\pi}(x) = \frac{1}{K}\sum_{i=1}^{K} \frac{\pi(x, z_i)}{q(z_i)}, \quad z_i \sim q(z).$$

(2) Construct a Markov chain as follows: at the current state x, propose a new state $x' \sim Q(x'|x)$ and accept it as the next state with the following probability

$$\alpha = \min\left(1, \frac{\hat{\pi}(x')Q(x|x')}{\hat{\pi}(x)Q(x'|x)}\right)$$

where $\hat{\pi}(x)$ and $\hat{\pi}(x')$ are both estimated using (1); otherwise x is the next state. Show that π is a stationary distribution of the Markov chain. Hint: you may want to consider the corresponding Markov chain in an augmented space. *(10 points)*

Problem 3.

Consider the probit regression model, with $x_i \in \mathbb{R}$ and $Y_i \in \{0, 1\}$:

$$\begin{aligned} p(Y_i = 1|Z_i) &= p(Z_i \geq 0) \\ Z_i &\sim \mathcal{N}(x_i\beta, \sigma^2) \\ \beta &\sim \mathcal{N}(0, 10^2) \\ \sigma^2 &\sim \text{Inv-}\chi^2(3, 1), \end{aligned}$$

where (x_i, y_i), $i = 1, \ldots, n$ are the observe data. Download the data from the course website.

(1) Implement a Gibbs sampler for approximating the posterior of (β, σ^2). Initialize the sampler at $(25, 25)$ and run it for $20,000$ iterations (with no burn-in). Plot the samples on top of the contours of the true posterior. *(15 points)*

(2) Now consider an alternative sampler that inserts a Metropolis-Hasting step after each Gibbs cycle which scales the current state of the Markov chain, $(\beta^{(t)}, (\sigma^2)^{(t)})$, by a factor s which is drawn from an exponential distribution, Exp(1). The rescaled state is accepted or rejected according to the usual Metropolis-Hastings procedure. Implement this sampler and conduct a similar simulation as in part (1), again plotting the samples on top of the contours of the true posterior. Which sampler mixes more quickly? *(15 points)*

Problem 4.

Consider a logistic regression model with normal priors

$$y_i \sim \text{Bernoulli}(p_i),\ p_i = \frac{1}{1 + \exp(-x_i^T\beta)},\quad i = 1, \ldots, n. \quad \beta \sim \mathcal{N}(0, \sigma_\beta^2)$$

where $\sigma_\beta = 1$. Download the data from the course website.

(1) Implement a Hamiltonian Monte Carlo sampler to collect 500 samples (with 500 discarded as burn-in), show the scatter plot. Test the following two strategies for the number of leapfrog steps L: (1) use a fixed L; (2) use a random one, say Uniform$(1, L_{\max})$. Do you find any difference? Explain it. *(15 points)*

(2) Run HMC for 100000 iterations and discard the first 50000 samples as burn-in to form the ground truth. Implement stochastic gradient MCMC algorithms including SGLD, SGHMC and SGNHT. Show the convergence rate of different SGMCMC algorithms in terms of KL divergence to the ground truth as a function of iterations. You may want to use the ITE package `https://bitbucket.org/szzoli/ite-in-python/src/` to compute the KL divergence between two samples. *(20 points)*