Modern Computational Statistics, Problem Set 2

October 25, 2019 Due 11/06/2019

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.
(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?
(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?

Problem 2. Consider a univariate normal model with conditional conjugate priors

$$x_i \sim \mathcal{N}(\mu, \sigma^2),\ i = 1, \ldots, n. \qquad \mu \sim \mathcal{N}(\mu_0, \tau_0^2),\ \ \sigma^2 \sim \text{Inv-}\chi^2(\nu_0, \sigma_0^2)$$

where $\mu_0 = 0, \tau_0 = 1, \nu_0 = 1, \sigma_0 = 1$. Download the data from the course website.

(1) Derive the conditional distributions for the Gibbs sampler.
(2) Implement a Gibbs sampler to simulate 1000 samples from the posterior.
(3) Implement a random walk Metropolis algorithm to simulate 1000 samples from the posterior (be careful to tune your step sizes to achieve good performance). Hint: you can use the following symmetrical proposal for σ^2

$$(\sigma^2)' = \exp(\text{Uniform}(-\delta, \delta)) \cdot \sigma^2$$

(4) Compare the samples given by the above samplers. Show the trace plots of μ and σ^2. Choose appropriate number of samples as burn-in based on the trace plot, and show the autocorrelation plots of μ and σ^2. Which sampler is better in this case? Explain your findings.

Problem 3. Consider a logistic regression model with normal priors

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

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.
(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/default/` to compute the KL divergence between two samples.