Stat 238, Spring 2025
Homework 5

Alexander Strang
Due: by **11:00 am** Tuesday, April 29th, 2025

Submission Instructions

Homework assignments may have a written portion and a code portion. Please follow the directions here Homework Guidelines Ed post when submitting, and check the link for homework policies.

This homework focuses on the algorithms introduced in the first half of the computation unit.

Problems eligible for spot grading are marked with an asterisk. These are problems 12.2, 12.3, 6.4 and 6.7. Please mark at least two problems for spot grading. Note, these will be graded for effort, not accuracy, so you should select the problems that show your best problem-solving effort. The goal is to target our attention to give you meaningful feedback on your work. All problems will be checked for completion and basic accuracy.

Problem 1: MCMC Sampling in Mixture Models

Sometimes it helps to expand the space of unknowns we aim to sample to include additional variables (e.g. the temperature in simulated tempering, or the momentum in Hamiltonian Monte Carlo). We've also seen that it can be useful to expand models hierarchically, by expressing a desired distribution as a mixture of simpler distributions.

Let X denote the unknowns of interest, and Z the additional unknowns introduced in the model or sampler. Let $Y = y$ denote the observed data.

In both settings, we usually want to sample from a target distribution on X. In a Bayesian setting, this target is usually either the marginal posterior for X marginalizing out Z (as in HMC, or when Z is a latent parameter in a mixture model), or the conditional posterior for X conditioned on a particular value of Z (as in simulated tempering).

As a specific example, consider the mixture model setting where we first sample $X \sim p_X$, then $Z|X \sim p_{Z|X}$, and finally, $Y|Z \sim p_{Y|Z}$. Suppose, in addition, that we can derive, and sample from, the exact conditional $Z|X = x, Y = y$.

Then, consider the following joint MCMC procedure on X, Z. Given X_t, Z_t:

1. Sample a candidate X' from a symmetric proposal distribution given the current iterate X_t. Let $p_r(x \to x')$ denote the probability that we propose $X' = x'$ given $X_t = x$.

2. Sample $Z'|X', Y = y$ using its exact conditional, $p_{Z|X,Y}$.

3. Accept or reject the joint candidate X', Z' as a single Metropolis-Hasting step on the joint posterior.

(a) Derive the acceptance probability needed so that the MH procedure described above samples from the joint posterior $X, Z|Y = y$.

(b) Describe how you could use this joint sampling procedure to sample from the marginal posterior $X|Y = y$. Is your procedure the same as Metropolis sampling on the target $X|Y = y$ for some choice of proposal?

Book Problems

Please complete the following problems from *Bayesian Data Analysis* (Third Edition).

- Chapter 12: problems 2*, and 3*. You may use Python instead of R for 12.3.
- Chapter 6: problems 1 (a), 4* (you may use the example from Lab 6 or try to replicate the SAT Coaching example), and 7*.

For this Homework it may help to review the “SAT Coaching” example presented in Chapters 5.4, 5.5, and 6.5, and the “Bioassay” example presented in Chapter 3.7 and Lab 6.

Problem 0: Spot Grading

Select two problems from the problems marked with an asterisk for spot grading.