STAT 545, Spring 2026
Homework 5

Due: 2/19/26, 11:59pm, through Brightspace
Total points: 50

Problem 1

[10 + 10 + 15 + 15 points]

A. Consider sampling from a $\mathrm{N}(\mu, \Sigma)$ distribution where

$$\Sigma = (X^T X + D^{-1})^{-1}, \quad \mu = \Sigma X^T y, \tag{1}$$

where $D \in \mathbb{R}^{p\times p}$ is a positive definite diagonal matrix; $X \in \mathbb{R}^{n\times p}$; and $y \in \mathbb{R}^{n\times 1}$. Distributions of this form arise, for example, as the posterior distribution for β under the model $y = X\beta + \epsilon$ where the prior is $\beta \sim \mathrm{N}(0, D)$ and $\epsilon \sim \mathrm{N}(0, I_n)$. Assume $p > n$. What is the computational complexity for directly computing Σ as in Equation (1) as a function of n and p?

B. You are given the Woodbury matrix identity:

$$(A + UBV)^{-1} = A^{-1} - A^{-1}U(B^{-1} + VA^{-1}U)^{-1}VA^{-1},$$

for compatible matrices A, B, V, U, assuming all the required inverses exist. Show how this result can be used to compute Σ so that the computational cost is lower than in Part A.

C. Next, consider the following algorithm, consisting of steps (i) – (iv):

(i) Sample $u \sim \mathrm{N}(0, D)$ and $\delta \sim \mathrm{N}(0, I_n)$ independesntly.
(ii) Set $v = Xu + \delta$.
(iii) Solve $(XDX^T + I_n)w = (y - v)$ to obtain w.
(iv) Set $\theta = u + DX^T w$.

Suppose θ is obtained according to this algorithm. Show that $\theta \sim \mathrm{N}(\mu, \Sigma)$. Also, calculate the computational complexity of the algorithm. Assume that the sampling cost of a univariate normal is $O(1)$.

D. Take $n = 50, p = 100$. Simulate X and ϵ from i.i.d. standard normals. Take β to be a vector of all ones and set $y = X\beta + \epsilon$. Take $D = I$ and implement the sampling schemes of parts (a) , (b) and (c) to generate 1,000 samples each. Report the computational times. You may use the Cholesky decomposition function in R.