DSA 595 Bayesian computations for machine learning
Problem set 6

February 26, 2025

1. Derive the likelihood function for data drawn independently from the logistic regression model:
$$Y_i \mid \beta \sim \text{Bernoulli}\left(\frac{1}{1+e^{-x_i\beta}}\right),$$
for $i \in \{1, \ldots, n\}$, where $\beta, x_i \in \mathbb{R}$.

2. Taking the prior distribution on β to be normal$(0, \tau^2)$, derive the kernel of the posterior distribution of β.

3. Write a Metropolis-Hastings algorithm to draw samples from the posterior $\pi(\beta \mid y_1, \ldots, y_n)$. Generate synthetic data from the logistic regression model, for some fixed choice of β, and some fixed covariate values $x_1, \ldots, x_n$ (you can generate these as well, by then they are regarded as fixed).

4. Write code to compute the Gelman-Rubin statistic, and then use it to assess/diagnose convergence of your MCMC algorithm in the previous problem.