CS6190: Probabilistic Modeling Homework 3 MCMC

Gopal Menon

01 April, 2018

1.

2. Say you are given data (X, Y), with $X \in \mathbb{R}^d$ and $Y \in \{0, 1\}$. The goal is to train a classifier that will predict an unknown class label $\tilde{y}$ from a new data point $\tilde{x}$. Consider the following model:

$$\begin{aligned} Y &\sim Ber\left(\frac{1}{1 + e^{-X^T\beta}}\right), \\ \beta &\sim N(0, \sigma^2 I). \end{aligned}$$

This is a **Bayesian logistic regression** model. Your goal is to derive and implement a Hamiltonian Monte Carlo sampler for doing Bayesian inference on β.

(a) Write down the formula for the unormalized posterior of $\beta|Y$, i.e.,

$$\begin{aligned} p(\beta|y;x,\sigma) &\propto \prod_{i=1}^{n} p(y_i|\beta;x_i)p(\beta;\sigma) \\ &= \prod_{i=1}^{n} \left(\left[\frac{1}{1+e^{-x_i^T\beta}}\right]^{y_i} \left[1 - \frac{1}{1+e^{-x_i^T\beta}}\right]^{1-y_i} \right) \frac{1}{\sqrt{(2\pi)^k|\sigma^2 I|}} e^{\left(-\frac{1}{2}\beta^T\left(\sigma^2 I\right)^{-1}\beta\right)} \end{aligned}$$

(b) Show that this posterior is proportional to $exp(-U(\beta))$, where

$$U(\beta) = \sum_{i=1}^{n} (1 - y_i)x_i^T\beta + \log(1 + e^{-x_i^T\beta}) + \frac{1}{2\sigma^2}\|\beta\|^2$$

$$\begin{aligned} p(\beta|y;x,\sigma) &\propto \prod_{i=1}^{n} \left(\left[\frac{1}{1+e^{-x_i^T\beta}}\right]^{y_i} \left[1 - \frac{1}{1+e^{-x_i^T\beta}}\right]^{1-y_i} \right) \frac{1}{\sqrt{(2\pi)^k|\sigma^2 I|}} e^{\left(-\frac{1}{2}\beta^T\left(\sigma^2 I\right)^{-1}\beta\right)} \\ &= \prod_{i=1}^{n} \left(\left[\frac{1}{1+e^{-x_i^T\beta}}\right]^{y_i} \left[\frac{e^{-x_i^T\beta}}{1+e^{-x_i^T\beta}}\right]^{1-y_i} \right) \frac{1}{\sqrt{(2\pi)^k|\sigma^2 I|}} e^{\left(-\frac{\|\beta\|^2}{2\sigma^2}\right)} \\ \log\left(p(\beta|y;x,\sigma)\right) &\propto -\sum_{i=1}^{n} y_i \log\left(1 + e^{-x_i^T\beta}\right) + \sum_{i=1}^{n}(1-y_i). - x_i^T\beta - \sum_{i=1}^{n}(1-y_i)\log\left(1+e^{-x_i^T\beta}\right) - \frac{1}{2\sigma^2}\|\beta\|^2 \\ &= -\sum_{i=1}^{n} y_i \log\left(1 + e^{-x_i^T\beta}\right) - \sum_{i=1}^{n}(1-y_i)x_i^T\beta - \sum_{i=1}^{n}\log\left(1+e^{-x_i^T\beta}\right) + \sum_{i=1}^{n} y_i \log\left(1 + e^{-x_i^T\beta}\right) - \frac{\|\beta\|^2}{2\sigma^2} \\ &= -\sum_{i=1}^{n}(1-y_i)x_i^T\beta - \sum_{i=1}^{n}\log\left(1+e^{-x_i^T\beta}\right) - \frac{1}{2\sigma^2}\|\beta\|^2 \end{aligned}$$

$$\begin{aligned} &\Rightarrow p(\beta|y;x,\sigma) \propto exp\left(-U(\beta)\right), \text{where} \\ &U(\beta) = \sum_{i=1}^{n}(1-y_i)x_i^T\beta + \sum_{i=1}^{n}\log\left(1+e^{-x_i^T\beta}\right) + \frac{1}{2\sigma^2}\|\beta\|^2 \end{aligned}$$