Assignment 3 Question 1 Steven Maharaj 695281

Due: Friday 25 October 2019

There are places in this assignment where R code will be required. Therefore set the random seed so assignment is reproducible.

```
set.seed(965281)  #Please change random seed to your student id number.
```

Question One (17 marks)

To explore some properties of Expectation propagation and Hamiltonian Monte Carlo, consider the dataset `Warpbreaks.csv`, which is on LMS and previously analysed in assignment 2. This dataset contains information of the number of breaks in a consignment of wool. In addition, Wool type (A or B) and tension level (L, M or H) was recorded. As the observed data consists of integer counts, it was assumed that a Poisson distribution should be used to model counts. The probability mass function of a Poisson distribution is

$$Pr(y_i|\lambda_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!}.$$

a) Assuming that the canonical link for observation i can be represented as $\mathbf{X}_i\boldsymbol{\beta}$, determine the following:

- The likelihood, $p(\mathbf{y}|\boldsymbol{\beta})$ and log-likelihood.
- The first derivative of the log-likelihood with respect to $\boldsymbol{\beta}$.

Answer:

Given n observations we have

$$p(\boldsymbol{y}|\boldsymbol{\beta}) = \prod_{i=1}^{n} \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!}$$

thus the log-liklihood

$$\begin{aligned}
\log(p(\boldsymbol{y}|\boldsymbol{\beta})) &= \sum_{i=1}^{n} y_i \log(\lambda_i) - \lambda_i - \log(y_i!) \\
&= \sum_{i=1}^{n} y_i \boldsymbol{X}_i\boldsymbol{\beta} - e^{\boldsymbol{X}_i\boldsymbol{\beta}} - \log(y_i!) && \text{(since } \log(\lambda) = \boldsymbol{X}_i\boldsymbol{\beta}\text{)}
\end{aligned}$$

Taking the derivate

$$\frac{d\log(p(\boldsymbol{y}|\boldsymbol{\beta}))}{d\boldsymbol{\beta}_j} = \sum_{i=1}^{n} \boldsymbol{X}_{ij} y_i - \boldsymbol{X}_{ij} e^{\boldsymbol{X}_i\boldsymbol{\beta}}$$

Thus,

$$\frac{d\log(p(\boldsymbol{y}|\boldsymbol{\beta}))}{d\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{y} - e^{\boldsymbol{X}\boldsymbol{\beta}})$$

b) If you wish to construct a Bayesian analogue to Poisson regression, what prior(s) would you use?

c) Fit a Poisson regression to the warpbreak data, with Wool type and tension treated as factors, using Hamiltonian Monte Carlo. To ensure identifiability, make Wool type A and tension type H the reference category. You are expected to code this in `R`, as opposed to fitting the model using `Stan`. Consider the following values for the number of leapfrog steps $L = 2, 3, 4$. Assume the momentum variable ϕ is drawn from a multivariate normal distribution with zero mean and variance-covariance matrix $5\mathbf{X}'\mathbf{X}$. Run a single chain for each choice of L for 10000 iterations, and remove 30 % of iterations as burn-in. Report the following.

- The posterior mean, standard deviations and 90 % credible intervals for all parameters, combining the results for all chains. Interpret the 90 % credible interval.
- The acceptance rate for each choice of L.

d) Check each chain obtained converged to the same distribution. For each chain and parameter, create acf plots. Based on this, what do you think was the best choice for L?

e) Fit the same model using an expectation propagation algorithm. Report the approximate posterior means, and 90 % credible interval. Comparing the results obtained using expectation propagation to Hamiltonian Monte Carlo, what 'bias' do you think has been caused by using approximate inference.