

## CS4070: EXERCISE 2: POISSON REGRESSION

*Hand in before the deadline. Include code as an appendix.*

Consider the Poisson regression model which is a basic model for count-data. So we assume data  $\{(x_i, y_i)\}_{i=1}^n$ , where  $y_i \in \{0, 1, 2, \dots\}$  and  $x_i \in \mathbb{R}^p$ . The model is given by

$$y_i \sim \text{Pois}(\mu_i), \quad \text{where} \quad \log \mu_i = x_i^T \theta$$

for an unknown parameter vector  $\theta \in \mathbb{R}^p$ . Hence, for  $k \in \{0, 1, 2, \dots\}$  we have

$$\mathbb{P}(Y_i = k) = e^{-\mu_i} \mu_i^k / (k!).$$

- (1) Give the loglikelihood, assuming all  $y_i$  are independent.
- (2) Derive an expression for the gradient and Hessian of the loglikelihood.
- (3) In the following we consider the dataset `dataexercise2.csv`. We take a Bayesian point of view, where we assume

$$y_i | \theta \stackrel{\text{ind}}{\sim} \text{Pois}(\mu_i), \quad \text{where} \quad \log \mu_i = x_i^T \theta$$
$$\theta \sim N(0, \tilde{\sigma}^2 I_p).$$

Assume the prior standard deviation is given by  $\tilde{\sigma} = 4$ . Implement a Newton algorithm for computing the Laplace approximation to the posterior distribution. Report mean and covariance matrix of the approximation.

- (4) Implement a random-walk Metropolis Hastings algorithm to sample from the posterior. Take proposals of the form  $\theta^\circ := \theta + \sigma_{\text{proposal}} N(0, I_p)$ . Tune  $\sigma_{\text{proposal}}$  to achieve an acceptance rate of about 25% – 50%. Make a plot of the iterates where you plot  $\theta_2$  versus  $\theta_1$ , with colour indicating the iteration number. Report the Monte-Carlo estimate of the posterior mean (where you “throw away” burnin samples, i.e. initial samples where the chain has not reached its stationary region).
- (5) The results may be sensitive to the choice of  $\tilde{\sigma}$ . For that reason we add an extra layer to the hierarchical model in the following way:

$$y_i | \theta \stackrel{\text{ind}}{\sim} \text{Pois}(\mu_i), \quad \text{where} \quad \log \mu_i = x_i^T \theta$$
$$\theta | \tilde{\sigma} \sim N(0, \tilde{\sigma}^2 I_p)$$
$$\tilde{\sigma}^2 \sim IG(\alpha, \beta).$$

Here  $IG(\alpha, \beta)$  denotes the inverse Gamma distribution with parameters  $\alpha$  and  $\beta$  (its density function is given in exercise 3.12 in RG). Take  $\alpha = \beta = 0.2$ . Implement a Gibbs sampler that iteratively samples from the full conditionals of  $\theta$  and  $\tilde{\sigma}$ .

Include a derivation for the update-step for  $\tilde{\sigma}^2$  in your report. Also include a traceplot of the posterior samples of  $\tilde{\sigma}^2$  (a traceplot is a plot of iterate value versus iterate number).