Derivation of posterior, its gradient and Hessian

Likelihood:

$$y \sim \text{Poisson}(\lambda_i) \xrightarrow{\text{Link}} \log(\lambda_i) = x_i^T\beta \Rightarrow \lambda_i = e^{x_i^T\beta}$$

$$p(y_i \mid x_i, \beta) = \prod_{i=1}^{n} \frac{1}{y_i!} \cdot e^{x_i^T\beta y_i} \cdot e^{-e^{x_i^T\beta}}$$

$$\log p(y_i \mid x_i, \beta) = \sum_{i=1}^{n} \left(\log \frac{1}{y_i!} + x_i^T\beta \cdot y_i - e^{x_i^T\beta} \right)$$

Prior: $p(\beta) = N(0, \alpha I) = \frac{1}{(2\pi)^{k/2} \cdot \det(\alpha I)^{1/2}} e^{-\frac{1}{2}\beta^T \cdot (\alpha I)^{-1} \cdot \beta} =$

$$= \text{const} \cdot e^{-\frac{1}{2}\beta^T \frac{1}{\alpha} \cdot I \beta} = \text{const} \cdot e^{-\frac{1}{2\alpha} \cdot \beta^T\beta}$$

$$\log p(\beta) = -\frac{1}{2\alpha} \cdot \beta^T\beta + \log(\text{const})$$

Posterior:

$$l = \log(p(\beta \mid X, y)) = \sum_{i=1}^{n} \left(x_i^T\beta y_i - e^{x_i^T\beta} \right) - \frac{1}{2\alpha} \beta^T\beta + \text{const} =$$

$$= y^T \cdot X \cdot \beta - \sum_{i=1}^{n} e^{x_i^T\beta} - \frac{1}{2\alpha}\beta^T\beta + \text{const}$$

$$\frac{dl}{d\beta} = X^T y - \sum_{i=1}^{n} x_i^T \cdot e^{x_i^T\beta} - \frac{1}{\cancel{2}\alpha} \cancel{2}\beta = X^T y - X^T e^{X\beta} - \frac{1}{\alpha}\beta$$

$$\frac{d^2 l}{d\beta^2} = - \sum_{i=1}^{n} x_i^T \cdot e^{x_i^T\beta} \cdot x_i - \frac{1}{\alpha} I = -X^T \cdot \text{diag}(e^{x^T \cdot \beta}) \cdot X - \frac{1}{\alpha} I$$