$p(\theta|y) \propto p(y|\theta)\, p(\theta)$

we are maximizing log posterior : $\log p(\theta|y) \propto \log p(y|\theta) + \log p(\theta)$

$\Leftrightarrow$ minimizing $-\log p(\theta|y)$

poisson likelihood, given $\lambda_i = e^{x_i^T\theta}$

$$p(y|\theta) = \prod_{i=1}^{n} \frac{(e^{x_i^T\theta})^{y_i}\, e^{-e^{x_i^T\theta}}}{y_i!} \xrightarrow{\text{take log}} \log p(y|\theta) = \sum_{i=1}^{n} y_i x_i^T\theta - e^{x_i^T\theta} - \log(y_i!)$$

normal prior : $\theta \sim N(0, \sigma^2 I)$ $\qquad \Sigma^{-1} = (\sigma^2 I)^{-1} = \frac{1}{\sigma^2} I$

$$p(\theta) = \frac{1}{\sqrt{2\pi^d|\Sigma|}} e^{-\frac{1}{2}((\theta-\mu)^T\Sigma^{-1}(\theta-\mu))} \Rightarrow \frac{1}{\sqrt{2\pi^d|\Sigma|}} \cdot e^{-1/2(\theta^T\cdot\theta)\cdot\frac{1}{\sigma^2}}$$

$$\therefore \log p(\theta) = \log\left(\frac{1}{\sqrt{2\pi^d|\Sigma|}}\right) - \frac{1}{2\sigma^2}\cdot(\theta^T\theta)$$

$$\Rightarrow -\log p(\theta|y) = -\log p(y|\theta) - \log p(\theta)$$

$$= \sum_{i=1}^{n} e^{x_i^T\theta} + \underbrace{\log(y_i!)}_{\approx 0} - y_i x_i^T\theta + \frac{1}{2\sigma^2}\theta^T\theta - \underbrace{\log\left(\frac{1}{\sqrt{2\pi^d|\Sigma|}}\right)}_{\approx 0}$$

$$\therefore \frac{\partial -\log p(\theta|y)}{\partial\theta} = \sum_{i=1}^{n} x_i e^{x_i^T\theta} - y_i x_i + \frac{\theta}{\sigma^2}$$

in vector form

$$\Rightarrow \sum_{i=1}^{n} x_i\lambda_i - y_i x_i + \frac{\theta}{\sigma^2} \Rightarrow \underline{X^T(\lambda - y) + \frac{\theta}{\sigma^2}}$$

$$\Rightarrow \frac{\partial -\log p(\theta|y)}{\partial\theta} = X^T(\lambda - y) + \frac{\theta}{\sigma^2}$$

Hessian: $X^T(e^{X^T\theta} - y) + \frac{\theta}{\sigma^2} = X^T \operatorname{diag}(\lambda)\cdot X + I\cdot\frac{1}{\sigma^2}$

$$\frac{\partial e^{X^T\theta}}{\partial\theta} = X^T\cdot\lambda = \operatorname{diag}(\lambda)\cdot X \quad ; \lambda = e^{X^T\theta}$$