4.c) write the loss function $l(\theta)$, the NLL of the dist.

Calculate the Hessian and show that it is always PSD

$$l(\theta) = -\sum_{i=1}^{n} \log p(y^{(i)} | x^{(i)}; \theta)$$

$$= -\sum_{i=1}^{n} \log b(y^{(i)}) + y^{(i)} \cdot \theta^T \cdot x^{(i)} - a(\theta^T \cdot x^{(i)})$$

$$H_{jk} = \frac{\partial l(\theta)}{\partial \theta_j \partial \theta_k}$$

$$= \sum_{i=1}^{n} a''(\theta^T x^{(i)}) x_j^{(i)} x_k^{(i)} \text{ , for each element ij of H}$$

$$z^T \cdot H \cdot z = \sum_{j,k=1}^{d} z_j \cdot \left(\sum_{i=1}^{n} a''(\theta^T \cdot x^{(i)}) \cdot x_j^{(i)} \cdot x_k^{(i)} \right) \cdot z_k$$

$$= \sum_{j,k=1}^{d} \sum_{i=1}^{n} a''(\theta^T \cdot x^{(i)}) \underbrace{x_j^{(i)} \cdot x_k^{(i)} \cdot z_k \cdot z_j}$$

$$= \sum_{j,k=1}^{d} \sum_{i=1}^{n} a''(\theta^T \cdot x^{(i)}) (z^T \cdot x)^2$$

since $a''(\theta^T \cdot x)$ is $\text{Var}(y|x;\theta)$, it is non-negative, so $z^T \cdot H \cdot z \geq 0$ will always be true and $l(\theta)$ is convex