Task 1

Show that minimizing the negative log likelihood for $\sigma(x_n) = \text{const}$ is equivalent to minimizing the MSE

Likelihood ~~is given~~ ~~for~~ is given by:

$$\arg\min_\theta -\prod_{i=1}^{m} p_{model}(x_i | \theta) \quad \text{for dataset } X=\{x_1, \dots, x_m\}$$

• Now we use the log ....

$$\Rightarrow \arg\min_\theta -\sum_{i=1}^{m} \log(p_{model}((x_i, \theta)))$$

• Use $p(y_n | x_n) = \mathcal{N}(\mu(x_n), \sigma(x_n))[y_n]$ as p_{model}

$$\Rightarrow \arg\min_\theta -\sum_{i=1}^{m} \underbrace{\log(p(y_i | x_i))}_{\text{inspect this closer}}$$

$$\Rightarrow \log(\mathcal{N}(\mu(x_i), \sigma(x_i))) = \log\left(\underbrace{\frac{1}{\sigma(x_i)\cdot\sqrt{2\pi}}}_{=a} \cdot \exp\left(-\frac{1}{2}\cdot\left(\frac{x_i-\mu(x_i)}{\sigma(x_i)}\right)^2\right)\right)$$

$$= \log(a) - \cancel{\log}\,\frac{1}{2}\cdot\left(\frac{x_i-\mu(x_i)}{\sigma(x_i)}\right)^2$$

ignore

$$\Rightarrow \arg\min_\theta \sum_{i=1}^{m} \log(a) + \cancel{\frac{1}{2}\sum_{i=1}^{m}} \frac{1}{2\sigma(x_i)} \sum_{i=1}^{m} (x_i - \mu(x_i))^2$$

ignore since constant!

= MSE