

Task 0: Calculaton:

The statistical model can be written as:

$$\epsilon_i = y_i - f(x_i; \theta) \sim N(0, \sigma^2)$$

Which is normally distributed. Thus the likelihood for one observation is:

$$f(x_i | \theta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - f(x_i; \theta))^2}{2\sigma^2}}$$

Which gives the following likelihood function for all obs:

$$L = \left[\frac{1}{\sqrt{2\pi\sigma^2}} \right]^N \prod_{i=1}^N e^{-\frac{(y_i - f(x_i; \theta))^2}{2\sigma^2}}$$

Taking natural logs gives the log likelihood function

$$N \cdot \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - f(x_i; \theta))^2$$

Maximizing the log likelihood function wrt. theta gives the maximum likelihood estimator. The first order condition for maximization can be written as

$$\frac{1}{\sigma^2} \sum_{i=1}^N (y_i - f(x_i; \theta)) \cdot \frac{\partial f(x_i; \theta)}{\partial \theta} = 0$$

Multiplying through with sigma squared we get the same first order condition as when minimizing the MSE. Thus:

$$\theta_{mle} = \arg \min_{\theta} MSE(x, y; \theta)$$