Assignment 2.2

The *Mean Squared Error* (MSE) is an often used *Loss Function* in Machine Learning and it is defined as followed:

$$MSE := \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2 \simeq E_\theta[(\theta - \hat{\theta})^2] =: MSE(\hat{\theta})$$

Show that the MSE decomposes into the *Variance* of the estimate and the *Bias* between the estimate and the unknown but 'true' model! What can we conclude from that observation? The following equalities are useful:

- $(a+b)^2 = a^2 + 2ab + b^2$
- $E[X+Y] = E[X] + E[Y]$
- $E[aX+b] = a \cdot E[X] + b$

to prove the connection:

$$MSE(\hat{\theta}) := E_\theta[(\theta - \hat{\theta})^2] = Bias(\theta, \hat{\theta})^2 + Var_\theta[\hat{\theta}]$$

Bias is defined as the difference between the *mean of the estimates* and the *actual value*.

Mathematically, Bias = $E\left[\hat{\theta}\right] - \theta$

We have given, $E\left[\left(\hat{\theta} - \theta\right)^2\right]$

$\Rightarrow E\left[\left(\hat{\theta} - \mu + \mu - \theta\right)^2\right]$ *(Adding and subtracting by μ)*

$\Rightarrow E\left[\left(\hat{\theta} - \mu\right)^2 + (\mu - \theta)^2 - 2\left(\hat{\theta} - \mu\right)(\mu - \theta)\right]$ *(By using the first equality)*

$\Rightarrow E\left[\left(\hat{\theta} - \mu\right)^2\right] + E\left[(\mu - \theta)^2\right] - 2 \cdot E\left[\left(\hat{\theta} - \mu\right)(\mu - \theta)\right]$ *(By using the third equality)*

Let $E\left[\hat{\theta}\right] = \mu$,

So as a resultant the third term in the equation becomes zero (μ being a constant and $E\left[\hat{\theta}\right] - \mu$ = 0)

So we will be left with,

$\Rightarrow E\left[\left(\hat{\theta} - E\left(\hat{\theta}\right)\right)^2\right] + E\left[\left(E\left(\hat{\theta}\right) - \theta\right)^2\right]$

$\Rightarrow Var_\theta\left[\hat{\theta}\right] + Bias\left(\theta, \hat{\theta}\right)^2$