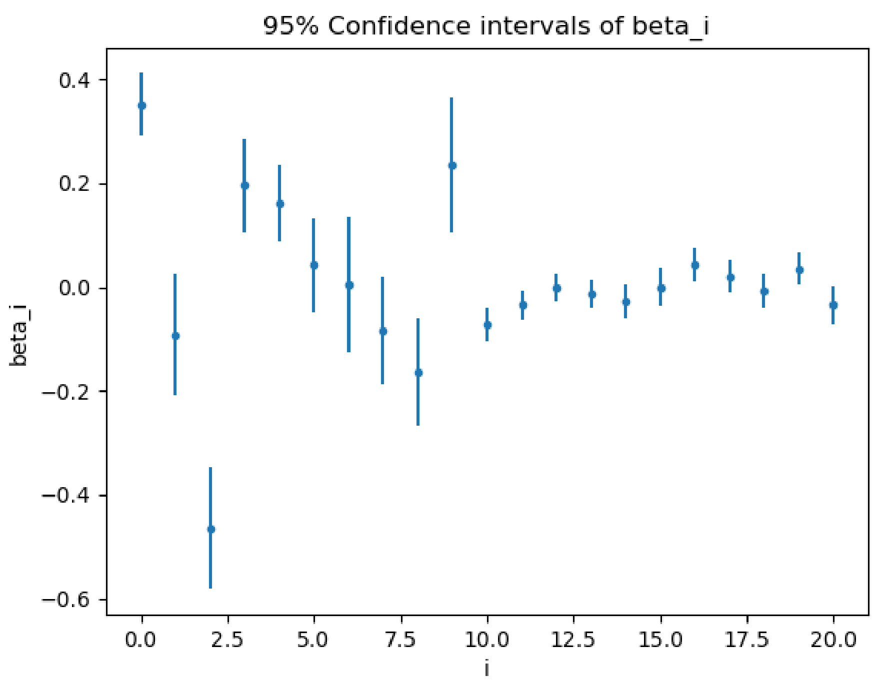
FYS-STK4155 Project 1

Espen Lønes

October 14, 2021

Exercise 1:



We see that the betas corresponding to lower degree parameters are larger in absolute value. Telling us these parameters are more important according to the model. But these also have the largest uncertainty.

Exercise 2:

We have:

$$C(\mathbf{X}, \beta) = \frac{1}{n}\sum_{i=0}^{n-1}(y_i - \tilde{y}_i)^2 = \mathbb{E}[(\mathbf{y} - \tilde{\mathbf{y}})^2]$$

To Derive the wanted equation we use the fact that the variance of $\mathbf{y}$ and ϵ are both σ^2. The mean of ϵ is zero and f is not stochastic for $\tilde{\mathbf{y}}$. And using the more compact notation of expected value, we get.

$$\mathbb{E}[(\mathbf{y} - \tilde{\mathbf{y}})^2] = \mathbb{E}[(\mathbf{f} + \epsilon - \tilde{\mathbf{y}})^2]$$

Then add and subtract $\mathbb{E}[\tilde{\mathbf{y}}]$

$$\mathbb{E}[(\mathbf{y} - \tilde{\mathbf{y}})^2] = \mathbb{E}[(\mathbf{f} + \epsilon - \tilde{\mathbf{y}} + \mathbb{E}[\tilde{\mathbf{y}}] - \mathbb{E}[\tilde{\mathbf{y}}])^2]$$

Then we use the expectation values mentioned above and get.

$$\mathbb{E}[(\mathbf{y} - \tilde{\mathbf{y}})^2] = \mathbb{E}[(\mathbf{y} - \mathbb{E}[\tilde{\mathbf{y}}])^2] + Var[\tilde{\mathbf{y}}] + \sigma^2$$

The tree terms in this equation represent, in order. The square of the bias of the model, the variance of the model and lastly the variance of the error.