Vedlegg B - Ligningene for Bias-variance tradeoff

Her skal vi utlede ligningene for bias-variance tradeoff som danner grunnlaget for tolkningen av bootstrapberegningene som benyttes i prosjektet. Denne utledningen baserer seg på utkesoppgaver gjort i kurset FYS-STK4155 i uke 38.

Vi tar for oss et datasett $\mathcal{L}$ som består av dataene $\mathbf{X}_\mathcal{L} = \{(y_j, \boldsymbol{x}_j), j = 0 \ldots n-1\}$.Og antar at de virkelige dataene kan uttrykkes som en (ukjent) funksjon + et ledd med normalfordelt støy, ϵ, med gjennomsnitt 0 og standerverdi.

$$\boldsymbol{y} = f(\boldsymbol{x}) + \boldsymbol{\epsilon}.$$

Modellen for funksjonen f er uttrykt ved parameterne $\boldsymbol{\beta}$ og designmatrisen $\boldsymbol{X}$

$$\tilde{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\beta}.$$

Og parameterne $\boldsymbol{\beta}$ er funnet (OLS) gjennom å optimalisere funksjonen

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

hvor $\mathbb{E}$ er forventningsverdien for datasettet.

Vi skal finne et uttrykk for $MSE = \mathbb{E}\left[(\boldsymbol{y} - \tilde{\boldsymbol{y}})^2\right] = \mathbb{E}\left[(\boldsymbol{y} - \tilde{\boldsymbol{y}})^2\right] = \mathrm{Bias}[\tilde{y}] + \mathrm{var}[\tilde{y}] + \sigma^2$ og vise at dette kan skrives som en sum av tre ledd hvor det ene representerer biasen til modellen og det andre leddet representerer variansen i modellen mens det siste bestemmes av støyen i dataene.

Begynner med å skrive ut uttrykket som vi skal ta forventnignsverdien av, og tar deretter for oss ett ledd om gangen.

$$\begin{aligned} MSE &= \mathbb{E}\left[(\boldsymbol{y} - \tilde{\boldsymbol{y}})^2\right] \\ &= \mathbb{E}\left[\boldsymbol{y}^2\right] - 2\mathbb{E}\left[\boldsymbol{y}\tilde{\boldsymbol{y}}\right] + \mathbb{E}\left[\tilde{\boldsymbol{y}}^2\right] \end{aligned}$$

Ser først variansen til modellen $\tilde{y}$. Vi bruker at

$$\begin{aligned} \mathrm{var}[\tilde{y}] &= \mathbb{E}\left[(\tilde{\boldsymbol{y}} - \mathbb{E}[\tilde{\boldsymbol{y}}])^2\right] \\ &= \mathbb{E}\left[\tilde{\boldsymbol{y}}^2 - 2\tilde{\boldsymbol{y}}\mathbb{E}[\tilde{\boldsymbol{y}}] + (\mathbb{E}[\tilde{\boldsymbol{y}}])^2\right] \\ &= \mathbb{E}\left[\tilde{\boldsymbol{y}}^2\right] - 2\mathbb{E}[\tilde{\boldsymbol{y}}]\mathbb{E}[\tilde{\boldsymbol{y}}] + (\mathbb{E}[\tilde{\boldsymbol{y}}])^2 \\ &= \mathbb{E}\left[\tilde{\boldsymbol{y}}^2\right] - (\mathbb{E}[\tilde{\boldsymbol{y}}])^2 \end{aligned}$$

Som igjen gir at $\mathbb{E}\left[\tilde{\boldsymbol{y}}^2\right]$ i MSE-en er gitt ved

$$\mathbb{E}\left[\tilde{\boldsymbol{y}}^2\right] = \mathrm{var}[\tilde{y}] + (\mathbb{E}[\tilde{\boldsymbol{y}}])^2$$

Ser deretter på det første leddet i uttrykket for MSE-verdien. Dette kan skrives som

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{y}^2\right] &= \mathbb{E}\left[(f + \varepsilon)^2\right] \\ &= \mathbb{E}\left[(f^2 + 2f\varepsilon + \varepsilon^2)\right] \\ &= \mathbb{E}\left[f^2\right] + 2\mathbb{E}\left[f\varepsilon\right] + \mathbb{E}\left[\varepsilon^2\right] \\ &= f^2 + 2f\mathbb{E}\left[\varepsilon\right] + \mathbb{E}\left[\varepsilon^2\right] \end{aligned}$$

siden f antas å være en ikke-stokastisk funksjon. Deretter bruker vi at at $\mathbb{E}\left[\varepsilon\right] = 0$ og at $\mathbb{E}\left[\varepsilon^2\right] = \sigma^2$, og finner at

$$\mathbb{E}\left[\boldsymbol{y}^2\right] = f^2 + 0 + \sigma^2$$

Ser til slutt på det midterste leddet i uttrykket for MSE-verdien:

$$\begin{aligned} 2\mathbb{E}\left[\boldsymbol{y}\tilde{\boldsymbol{y}}\right] &= 2\mathbb{E}\left[(f + \varepsilon)\tilde{\boldsymbol{y}}\right] \\ &= 2\mathbb{E}\left[f\tilde{\boldsymbol{y}}\right] + 2\mathbb{E}\left[\varepsilon\tilde{\boldsymbol{y}}\right] \\ &= 2\mathbb{E}\left[f\tilde{\boldsymbol{y}}\right] + 2\mathbb{E}\left[\varepsilon\right]\mathbb{E}\left[\tilde{\boldsymbol{y}}\right] \\ &= 2\mathbb{E}\left[f\tilde{\boldsymbol{y}}\right] + 0 \\ &= 2f\mathbb{E}\left[\tilde{\boldsymbol{y}}\right] \end{aligned}$$

Den totale funksjonen blir da

$$\begin{aligned} MSE &= \mathbb{E}\left[(\boldsymbol{y} - \tilde{\boldsymbol{y}})^2\right] \\ &= \mathbb{E}\left[\boldsymbol{y}^2\right] - 2\mathbb{E}\left[\boldsymbol{y}\tilde{\boldsymbol{y}}\right] + \mathbb{E}\left[\tilde{\boldsymbol{y}}^2\right] \\ &= f^2 + \sigma^2 - 2f\mathbb{E}\left[\tilde{\boldsymbol{y}}\right] + \mathrm{var}[\tilde{y}] + (\mathbb{E}[\tilde{\boldsymbol{y}}])^2 \\ &= f^2 - 2f\mathbb{E}\left[\tilde{\boldsymbol{y}}\right] + (\mathbb{E}[\tilde{\boldsymbol{y}}])^2 + \mathrm{var}[\tilde{y}] + \sigma^2 \\ &= \mathbb{E}[(f - \mathbb{E}[\tilde{\boldsymbol{y}}])^2] + \mathrm{var}[\tilde{y}] + \sigma^2 \end{aligned}$$

Siden vi gjerne ikke kjenner f, så approksimerer vi f med y og får

$$\begin{aligned} MSE &= \mathbb{E}\left[(\boldsymbol{y} - \tilde{\boldsymbol{y}})^2\right] \\ &= \mathbb{E}[(\boldsymbol{y} - \mathbb{E}[\tilde{\boldsymbol{y}}])^2] + \mathrm{var}[\tilde{y}] + \sigma^2 \end{aligned}$$

Og setter inn for $\mathrm{var}[\tilde{y}]$ og får

$$\begin{aligned} MSE &= \mathbb{E}\left[(\boldsymbol{y} - \tilde{\boldsymbol{y}})^2\right] \\ &= \mathbb{E}[(\boldsymbol{y} - \mathbb{E}[\tilde{\boldsymbol{y}}])^2] + \mathbb{E}\left[(\tilde{\boldsymbol{y}} - \mathbb{E}[\tilde{\boldsymbol{y}}])^2\right] + \sigma^2 \end{aligned}$$

Her er det første leddet biasen, altså hvor mye modellen i snitt avviker fra det vi forsøker å modellere, mens det andre leddet gir variansen i prediksjonene som modellern gjør i seg selv. Det siste leddet er konstant,og ukjent, så det ser vi bort ifra når vi skal optimere.