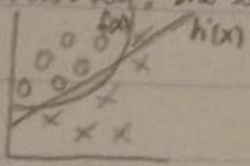
FDA SS2022

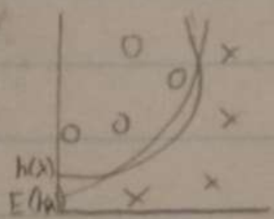
Pen & Paper

a) The bias of a hypothesis class is the difference between the chosen hypothesis function and the function $f(x)$, according to which the dataset S is distributed. The bias is bigger the more one tries to simplify a complex reality.

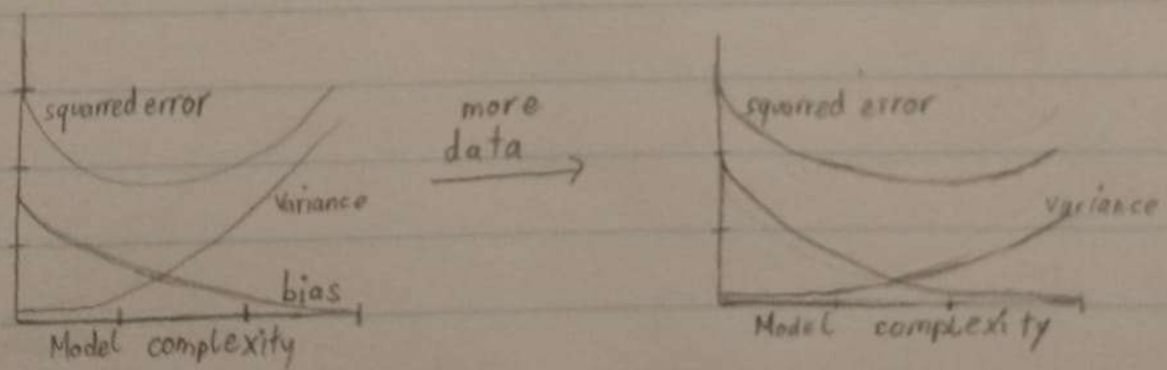


→ big bias

The variance of a hypothesis class is the expected squared error between the expectation of the chosen hypothesis and the chosen hypothesis. The variance increases with the complexity of $h^*(x)$.



c)



The variance sinks with more data. The bias increases slightly.

d) By using Regularisation and penalizing high variance, complexity can be reduced. You can also remove dimensions directly to reduce complexity.

1b)

$$E[(h^*(x)-y)^2] = E[(h^*(x)-f(x)+\varepsilon)^2] = E[(h^*(x)-f(x))^2] = ESE$$

$$E[\varepsilon]=0\ ,\ E[f(x)]=f(x)\ ,\ \sigma^2=0$$

$$Var[h^*(x)] = E[(E[h^*(x)]-h^*(x))^2] = E[(h^*(x)-f(x))^2]-(E[h^*(x)]-f(x))^2$$

$$Bias[h^*(x)] = E[h^*(x)]-f(x)$$

$$g = E[(h^*(x)-f(x))^2]$$

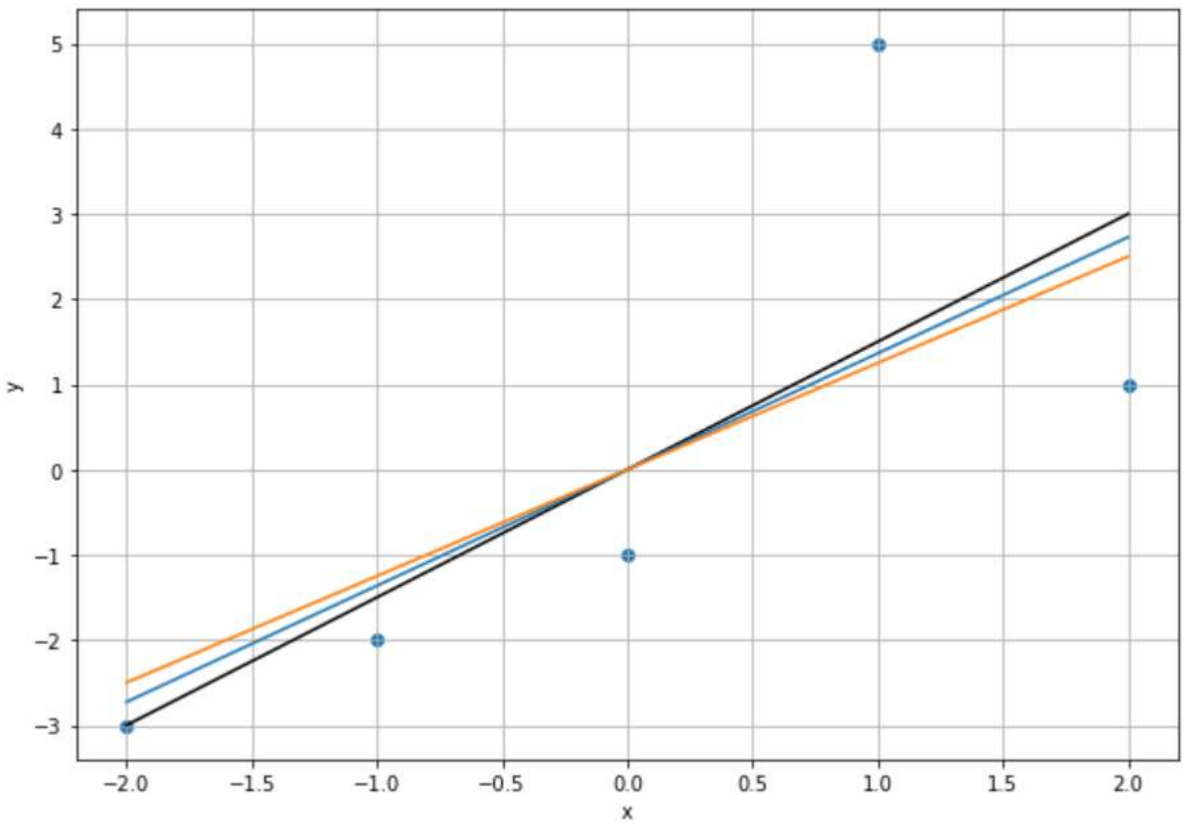
$$k = E[h^*(x)]-f(x)$$

$$ESE = (Var) + Bias^2$$

$$g = (g - k^2) + k^2$$

$$g = g$$

2b)



RidgeRegression(lambda =0) => weight = 1.5

RidgeRegression(lambda =1) => weight = 1.36

RidgeRegression(lambda =2) => weight = 1.25

```
linReg.coef_
```

```
array([1.5])
```

```
ridgeReg1.coef_
```

```
array([1.36363636])
```

```
ridgeReg2.coef_
```

```
array([1.25])
```

```
Observation: The higher the lambda, the more flat the line becomes. If lambda is 0, the result is equal to the
result of linear regression.
```

```
2d)
Through regularization the bias in the trainingsample is increased, so that in the longterm with more testsamples
the overall variance is decreased. It is especially beneficial, if there are outliers, that do not benefit the
learning process much, but scew the regression-line.
```