

Prosjekt 1

order = m = 5

$[x, y, x^2, y^2, xy, \dots]$ up to 5th order

5th order?

1	x	x ²	x ³	x ⁴	x ⁵
y	yx	yx ²	yx ³	yx ⁴	yx ⁵
y ²	y ² x	y ² x ²	y ² x ³	y ² x ⁴	y ² x ⁵
y ³	y ³ x	y ³ x ²	y ³ x ³	y ³ x ⁴	y ³ x ⁵
y ⁴	y ⁴ x	y ⁴ x ²	y ⁴ x ³	y ⁴ x ⁴	y ⁴ x ⁵
y ⁵	y ⁵ x	y ⁵ x ²	y ⁵ x ³	y ⁵ x ⁴	y ⁵ x ⁵

$$p = (m+1)^2$$
$$p = \frac{(m+1)(m+1)}{2}$$

p = 0

for i in range(m+1)

for j in range(m+1-i):

X[:, p] = (X[:, i] + (y[:, j]))

p += 1

5th order

↳ summen av eksponentene er fem

$$\begin{aligned} E[(\hat{y} - \tilde{y})^2] &= E[(\hat{f} + \hat{\varepsilon} - \tilde{y})^2] = \\ E[(\hat{f} + \hat{\varepsilon} - \tilde{y} + E[\hat{y}] - E[\hat{y}])^2] &= \end{aligned}$$

Brücker \hat{f}, \hat{y} ikke stokastiske

$$E[(\hat{y} - E[\hat{y}])^2] = E[(\hat{\varepsilon} - 0)^2] = \sigma^2$$

$$\text{Var}[\hat{y}] = E[(\hat{y} - E[\hat{y}])^2]$$

$$E[(\hat{f} - E[\hat{y}]) - (\hat{y} - E[\hat{y}] - \hat{\varepsilon})^2] =$$

$$E[(\hat{f} - E[\hat{y}])^2] - 2E[(\hat{f} - E[\hat{y}])(\hat{y} - E[\hat{y}] - \hat{\varepsilon})]$$

$$+ E[(\hat{y} - E[\hat{y}] - \hat{\varepsilon})^2]$$

$$E[(\hat{y} - E[\hat{y}])^2] - 2E[(\hat{y} - E[\hat{y}])(\hat{\varepsilon})] + E[\hat{\varepsilon}^2]$$

Struktur

Regression analysis and resampling methods

Abstract

Introduction

Status of problem and the major objectives

- regression og det store billedet
- presenter innholdet i resten av rapporten

Formalism / methods

- Om OLS, Ridges og Lasso
↳ Algoritme

- Resampling → Algoritme

- Bias-variance, hva leddene står for / bias²
↳
$$E[(\hat{y} - \tilde{y})^2] = \frac{1}{n} \sum_i (f_i - E[\hat{y}])^2 + \frac{1}{n} \sum_i (\tilde{y}_i - E[\hat{y}])^2 + \sigma^2$$

- Hva slags data vi bruker
↳ Franke function og digital termisk data

- Implementasjon

↳ python, scikit-learn

Resultater

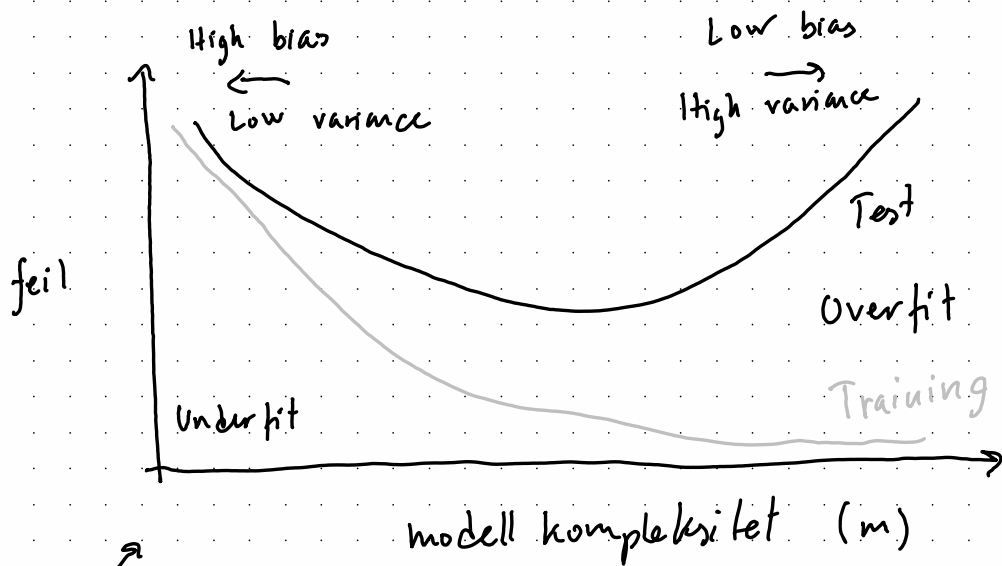
Sammenligne $MSE(\hat{y}, \hat{\hat{y}})$ $R^2(\hat{y}, \hat{\hat{y}})$
for de forskjellige metodene

v-validation

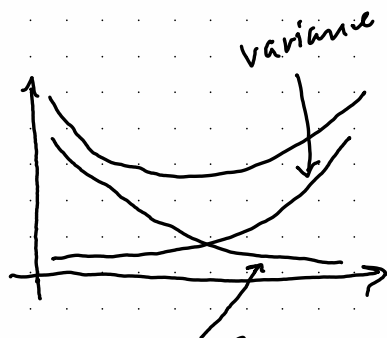
Tabell 2

	MSE	R^2
OLS		
OLS_v		
Ridge λ_1		
Ridge λ_2		
Ridge λ_3		
Lasso		

Se på Bias-variance trade off



Reprodusere denne med OLS
Ridge for forskjellige λ
Lasso



Plotte bias og varians? bias^2

Evaluer OLS, Ridge og Lasso
for våre ekte data

↳ velge ut et område

Diskusjon

Variance - bias tradeoff

Hvor kompleks bør modellen være?

Hva bør λ være?

Hvilken modell passer best til våre
digital terrain data.

Konklusjon

Husk dette!

Siteringer