

Linær regresjoni

Eirik vetle Winnness

October 7, 2018

Abstract

Introduction

I dette prosjektet skal jeg se på hvordan de tre forskjellige metodene OLS, ridge, og Lasso best greier å etterligne den ekte funksjonen (Frankfunksjon). Linær regresjon er en metode man kan bruke for å kunne spå hvordan ny data vil oppføre seg, basert på gammel data. Jeg skal studere MSE(), R2, Bias, og. Jeg har skal studere random data ved å generere det i python programmet ankonda og dens pakker. Jeg skal så se på de fire faktorene og studere OLS, ridge og Lasso som om de fungerer annerledes. Om noen av metodene er bedre til å etterligne datan. jeg kommer til å bruke trenings data for å lage en model, så ser jeg om denne modellen tilpasser seg godt opp mot den virkelige funksjonen.

Teori

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \tilde{y}_i)^2$$
$$R2(y, \hat{y}) = 1 - \frac{1}{n} \frac{\sum_{i=0}^{n-1} (y_i - \tilde{y}_i)^2}{\sum_{i=0}^{n-1} (y_i - \bar{y})^2}$$

Gjennomsnittet av \hat{y} er definert som:

$$\bar{y} = \frac{1}{n} \sum_{i=0}^{n-1} (y_i)$$

$$\bar{y} = \frac{1}{n} \sum_{i=0}^{n-1} (y_i)$$

Ordinary Least squar

Ridge Regression

Lasso Regression

Ordinary Least squar

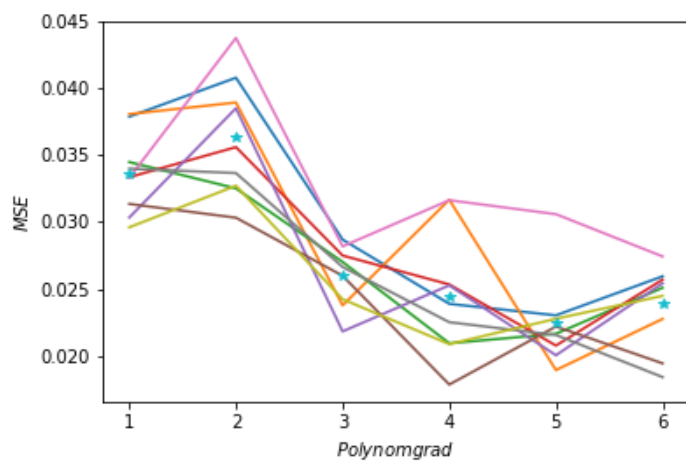


Figure 1: Rød(posetivt lad) blå er negativt lad

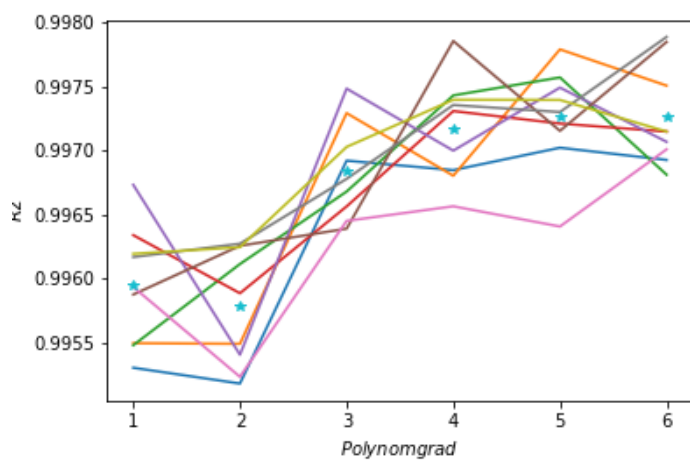


Figure 2: Rød(posetivt lad) blå er negativt lad

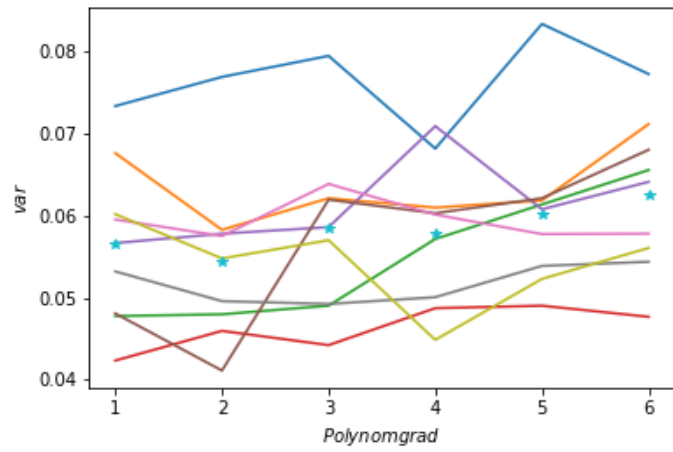


Figure 3: Rød(posetivt lad) blå er negativt lad

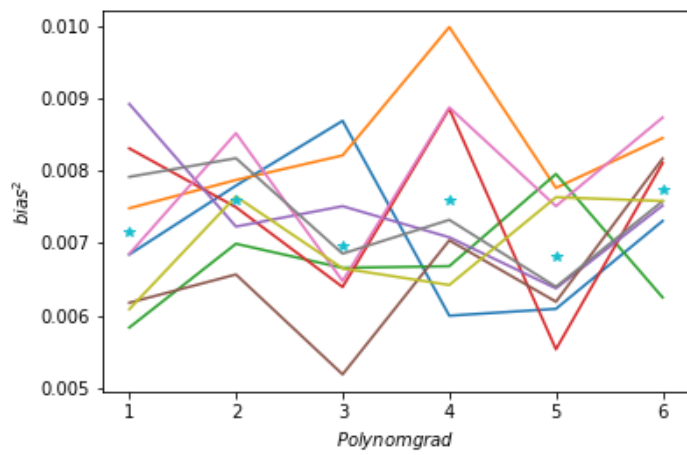


Figure 4: Rød(posetivt lad) blå er negativt lad

Ridge Regression

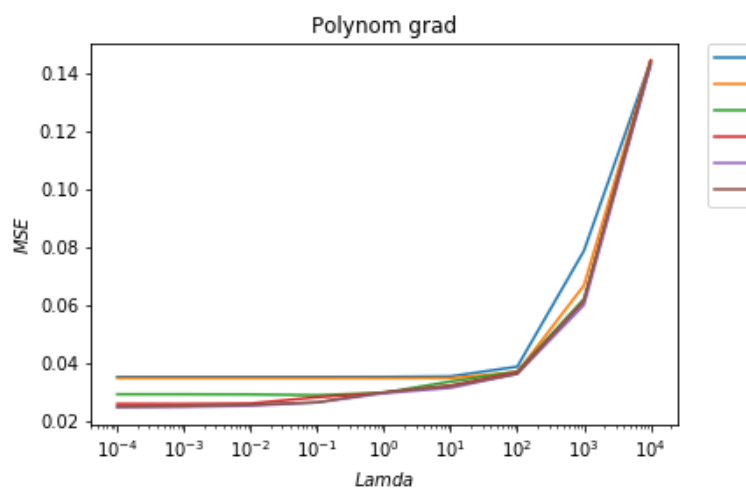


Figure 5: Rød(posetivt lad) blå er negativt lad

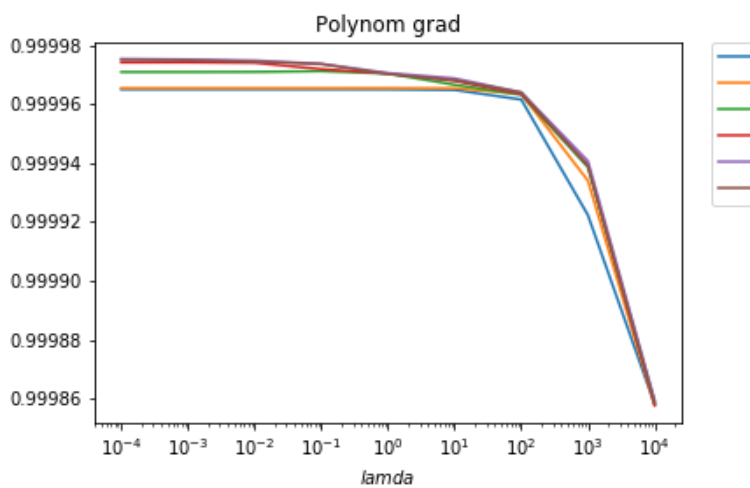


Figure 6: Rød(posetivt lad) blå er negativt lad

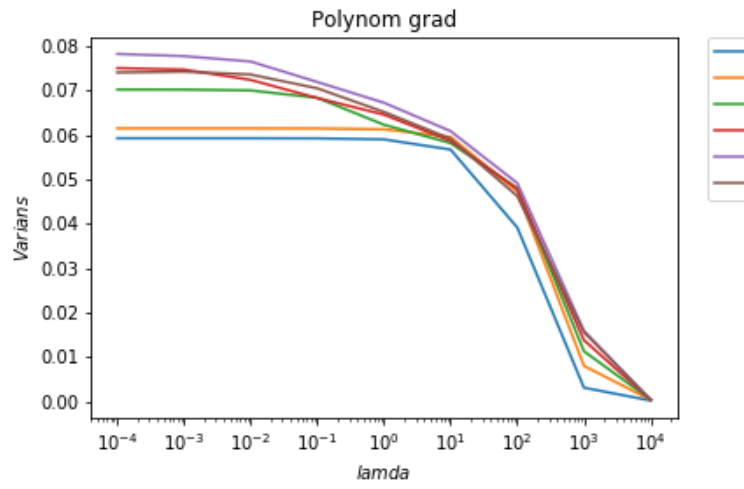


Figure 7: Rød(posetivt lad) blå er negativt lad

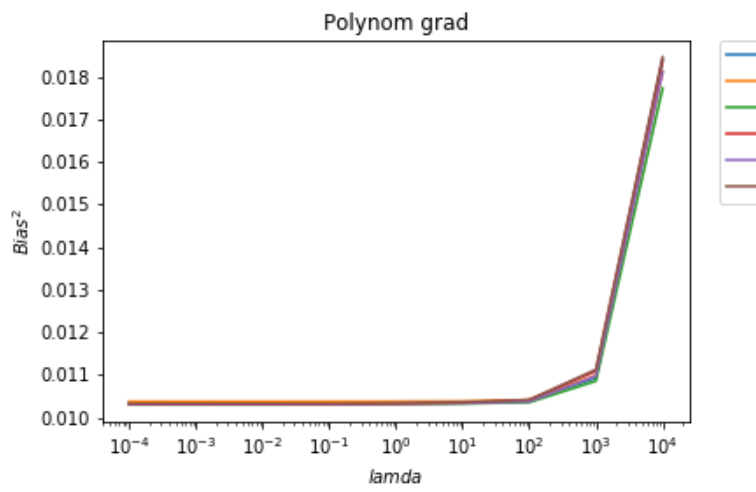


Figure 8: Rød(posetivt lad) blå er negativt lad

Lasso Regression

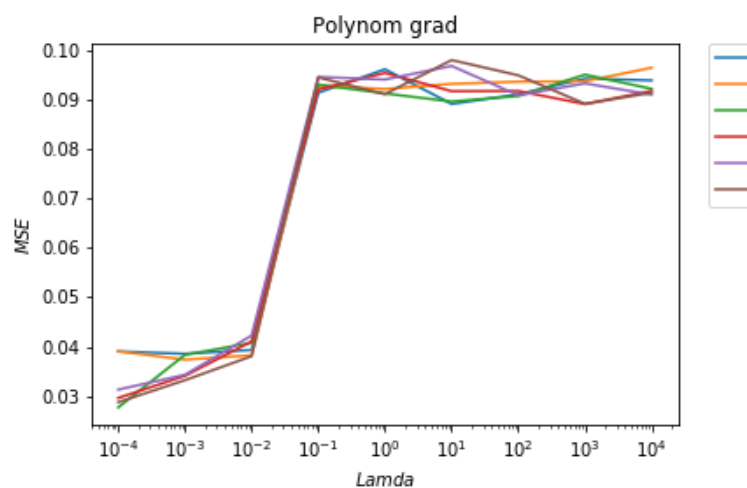


Figure 9: Rød(posetivt lad) blå er negativt lad

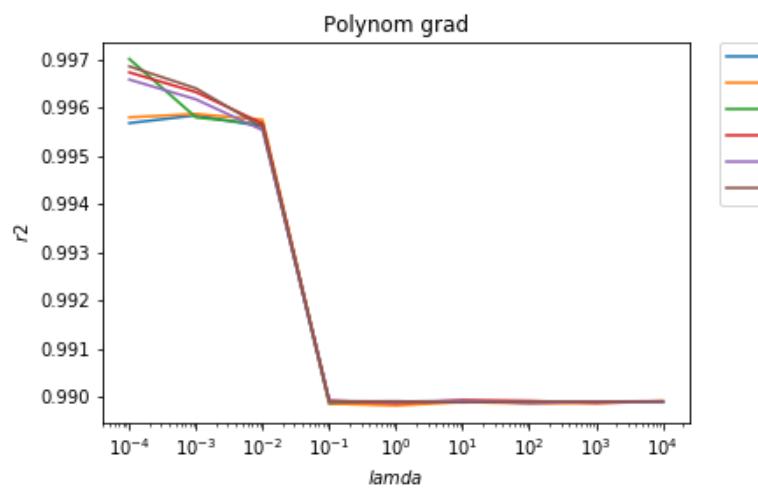


Figure 10: Rød(posetivt lad) blå er negativt lad

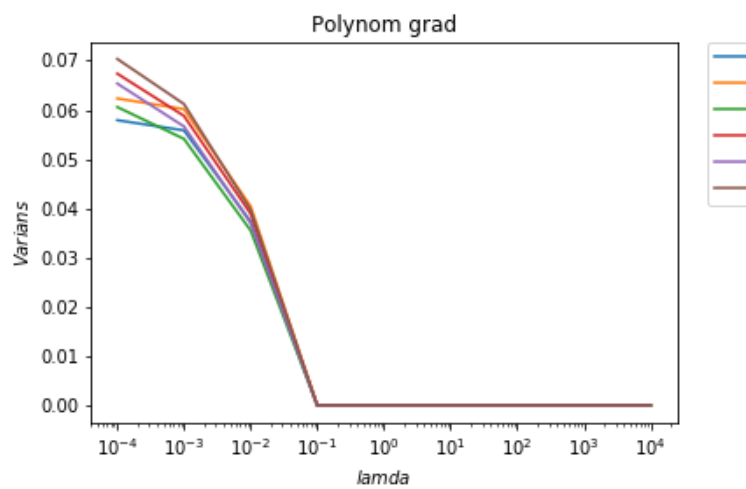


Figure 11: Rød(posetivt lad) blå er negativt lad

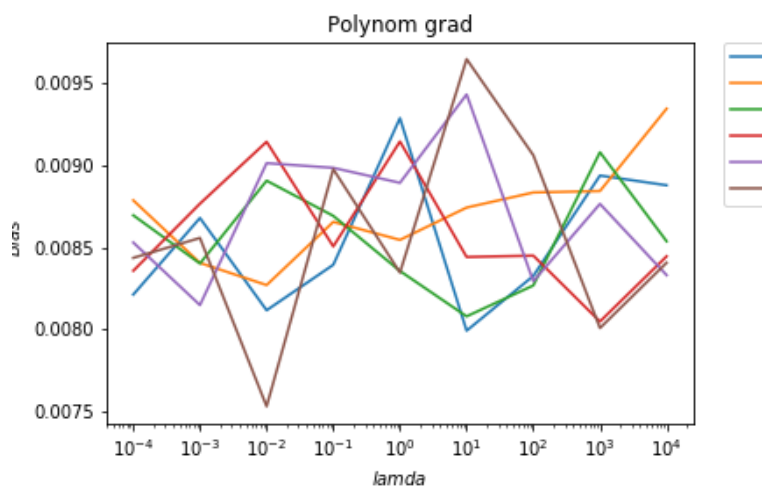


Figure 12: Rød(posetivt lad) blå er negativt lad

Real data

konklusjon