The Large Scale Structure of the Cosmic Microwave Background

Edvard B. Rørnes,[*] Anton A. Brekke,[†] and Isak O. Rukan[‡]
Institute of Physics, University of Oslo,
0371 Oslo, Norway
(Dated: September 3, 2024)

asdfasdf

CONTENTS

1. Introduction 1

2. Theory 1
   2.1. OLS 1
   2.2. Ridge 1
   2.3. LASSO 1
   2.4. Resampling 1
   2.5. Bias-Variance 1

3. Implementation 1

4. Results 1
   4.1. OLS 1
   4.2. Ridge 1
   4.3. LASSO 1

5. Discussion 1

6. Conclusion 1

Part d) 1

Part e) 2

1. INTRODUCTION

intro thingy

2. THEORY

The general structure of all our models is that we have some data set $\{x_i, y_i\}$ where $i \in \{0, 1..., n-1\}$ where x_i are independent variables whilst y_i are dependent variables. The data is assumed to be described by

$$\boldsymbol{y} = f(\boldsymbol{x}) + \boldsymbol{\varepsilon} \tag{1}$$

where f is some continuous function which takes $\boldsymbol{x}$ as input and $\boldsymbol{\varepsilon}$ is a normal distributed error $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma^2)$. The function f will then be approximated with a model $\tilde{\boldsymbol{y}}$ in which we will consider a polynomial expansion with coefficients β_i:

$$\tilde{y}_i = \sum_{j=0}^{p-1} \beta_j x_i^j \tag{2}$$

defining the $n \times p$ design matrix $(\boldsymbol{X})_{ij} = (x_i)^j$ we can rewrite this as

$$\tilde{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\beta} \tag{3}$$

2.1. OLS

2.2. Ridge

2.3. LASSO

2.4. Resampling

2.5. Bias-Variance

3. IMPLEMENTATION

4. RESULTS

4.1. OLS

4.2. Ridge

4.3. LASSO

5. DISCUSSION

6. CONCLUSION

PART D)

Show that the expectation value and variance of $\boldsymbol{y}$ is

$$\mathbb{E}(y_i) = \sum_j x_{ij}\beta_j = \boldsymbol{X}_{i,*}\boldsymbol{\beta}, \quad \mathrm{Var}(y_i) = \sigma^2$$

where $\boldsymbol{y}$ is defined by $\boldsymbol{y} = f(\boldsymbol{x}) + \boldsymbol{\varepsilon}$. Here $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$ is a normal distributed error and $f(\boldsymbol{x})$ is the approximated function given our model $\tilde{\boldsymbol{y}}$ obtained by minimizing $(\boldsymbol{y} -$

[*] e.b.rornes@fys.uio.no
[†] asdf
[‡] asdfafs

$\tilde{\boldsymbol{y}})^2$ with $\tilde{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\beta}$. With the OLS expression for $\hat{\boldsymbol{\beta}}$ also show that

$$\mathbb{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$$

and

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}$$

Solution:
Trivially $\mathbb{E}(\varepsilon_i) = 0$ from its definition. Thus from the definition of $\boldsymbol{y}$ we have that

$$\mathbb{E}(y_i) = \mathbb{E}(f(x_i)) = \boldsymbol{X}_{i,*}\beta$$

Similarly the variance is given by

$$\begin{aligned}\mathrm{Var}(y_i) &= \mathbb{E}\{[y_i - \mathbb{E}(y_i)]^2\} = \mathbb{E}\{(\boldsymbol{X}_{i,*}\beta + \varepsilon_i)^2\} - (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2 \\ &= (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2 + \mathbb{E}(\varepsilon_i^2) + 2\mathbb{E}(\varepsilon_i)\boldsymbol{X}_{i,*}\boldsymbol{\beta} - (\boldsymbol{X}_{i,*}\boldsymbol{\beta})^2 \\ &= \mathrm{Var}(\varepsilon_i^2) = \sigma^2\end{aligned}$$

The optimal parameters β for OLS are given by

$$\hat{\boldsymbol{\beta}}_{\mathrm{OSL}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

which yields the expectation value

$$\mathbb{E}(\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}) = \mathbb{E}[(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}] = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\mathbb{E}[\boldsymbol{y}] = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}.$$

and the variance

$$\begin{aligned}\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}) &= \mathbb{E}\{[\boldsymbol{\beta} - \mathbb{E}(\boldsymbol{\beta})][\boldsymbol{\beta} - \mathbb{E}(\boldsymbol{\beta})]^T\} \\ &= \mathbb{E}\{[(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{\beta}][(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{\beta}]^T\} \\ &= (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\mathbb{E}\{\boldsymbol{y}\boldsymbol{y}^T\}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} - \boldsymbol{\beta}\boldsymbol{\beta}^T \\ &= (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T[\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}^T\boldsymbol{X}^T + \sigma^2]\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} - \boldsymbol{\beta}\boldsymbol{\beta}^T \\ &= \boldsymbol{\beta}\boldsymbol{\beta}^T + \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1} - \boldsymbol{\beta}\boldsymbol{\beta}^T = \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}\end{aligned}$$

PART E)

Show that you can rewrite

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

as

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

where

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

and

$$\mathrm{Var}[\tilde{y}] = \mathbb{E}\left[(\tilde{\boldsymbol{y}} - \mathbb{E}[\tilde{\boldsymbol{y}}])^2\right] = \frac{1}{n}\sum_i(\tilde{y}_i - \mathbb{E}[\tilde{\boldsymbol{y}}])^2$$