

# Compulsory exercise 1: Group 16

TMA4268 Statistical Learning V2022

Weicheng Hua, Emil Johannesen Haugstvedt, Torbjørn Baadsvik

28 March, 2022

## Contents

<b>Problem 1</b>	<b>1</b>
a) . . . . .	1
b) . . . . .	1
c) . . . . .	1
d) . . . . .	1
<b>Problem 2</b>	<b>1</b>
a) . . . . .	1
b) . . . . .	1
<b>Problem 3</b>	<b>1</b>
a . . . . .	1
b . . . . .	2
a) . . . . .	2
b) . . . . .	2
<b>Problem 4</b>	<b>2</b>
a) . . . . .	2
b) . . . . .	2
c) . . . . .	2
d) . . . . .	2
<b>Problem 5</b>	<b>2</b>
a) . . . . .	2
b) . . . . .	2

**Problem 6** **2**

- a) ..... 2
- b) ..... 2
- c) ..... 2
- d) ..... 2

**Problem 1**

- a)
- b)
- c)
- d)

**Problem 2**

- a)
- b)

**Problem 3**

**a**

- (i) True. The proof for this is pretty long and complicated. You can see it in the link
- (ii) False. By increasing the number of cutpoints each step function will get more and more affected by the points within their range, and thus overfit.
- (iii) False. The penalty term is  $\int g'(t)^2 dt$ .
- (iv) True. With high  $k$  the number more neighbors are needed in order to classify a point, thus the variance will be low and the bias will increase.

**b**

Fit an additive model on 'boston.train' using the function 'gam()' from package 'gam' with the following conditions, and plot the resulting curves.

- response: 'medv'; predictors: 'rm', 'ptratio', 'lstat' (use these three predictors only).
- 'rm' is a linear function
- 'ptratio' is a smoothing spline with 'df=3'.
- 'lstat' is a polynomial of degree 2.

```
# Fit model

model <- gam(medv ~ rm + s(ptratio, k = 3) + poly(lstat, df = 2), data = boston.train)

# Plot model with training data

plot(model, boston.train)
''''
```

a)

b)

#### Problem 4

a)

b)

c)

d)

#### Problem 5

a)

b)

#### Problem 6

a)

b)

c)

d)