Compulsory exercise 1: Group 39

TMA4268 Statistical Learning V2021

Alexander J Ohrt, Jim Totland

08 februar, 2021

Problem 1

a)

We assume that $\mathbf{Y}$ is a multivariate normal, which gives the distribution $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$.

$$
\begin{aligned}
\mathsf{E}(\widetilde{\boldsymbol{\beta}}) &= \mathsf{E}((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{Y}) = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathsf{E}(\mathbf{Y}) \\
&= (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathsf{E}(\mathbf{X}\boldsymbol{\beta} + \varepsilon) = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathsf{Cov}(\widetilde{\boldsymbol{\beta}}) &= \mathsf{Cov}((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{Y}) = ((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T)\mathsf{Cov}(\mathbf{Y})((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T)^T \\
&= ((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T)\sigma^2 I(\mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-T}) = \sigma^2((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-T}) \\
&= \sigma^2((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}),
\end{aligned}
$$

where we have used that $(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I})^{-1} = (\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I})^{-T}$ in the last equality (USIKKER PÅ DENNE SISTE! DET BLIR I HVERT FALL HELT RIKTIG Å BEHOLDE -T). In both these equations it is apparent that the moments are equal to those of the OLS estimator when $\lambda = 0$.

b)

The requested moments of $\widetilde{f}(\mathbf{x}_0)$ are

$$
\mathsf{E}(\widetilde{f}(\mathbf{x}_0)) = \mathsf{E}(\mathbf{x}_0^T\widetilde{\boldsymbol{\beta}}) = \mathbf{x}_0^T\mathsf{E}(\widetilde{\boldsymbol{\beta}}) = \mathbf{x}_0^T(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}
$$

and

$$
\begin{aligned}
\mathsf{Cov}(\widetilde{f}(\mathbf{x}_0)) &= \mathsf{Cov}(\mathbf{x}_0^T\widetilde{\boldsymbol{\beta}}) = \mathbf{x}_0\mathsf{Cov}(\widetilde{\boldsymbol{\beta}})\mathbf{x}_0^T \\
&= \sigma^2\mathbf{x}_0((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1})\mathbf{x}_0^T.
\end{aligned}
$$

c)

The expected MSE at $\mathbf{x}_0$ is

$$
\begin{aligned}
\mathsf{E}[(y_0 - \widetilde{f}(\mathbf{x}_0))^2] &= [\mathsf{E}(\widetilde{f}(\mathbf{x}_0) - f(\mathbf{x}_0)]^2 + \mathsf{Var}(\widetilde{f}(\mathbf{x}_0)) + \mathsf{Var}(\varepsilon) \\
&= [\mathsf{E}(\widetilde{f}(\mathbf{x}_0)) - \mathsf{E}(f(\mathbf{x}_0))]^2 + \mathsf{Cov}(\widetilde{f}(\mathbf{x}_0)) + \sigma^2 I \\
&= [\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} + \mathbf{x}_0^T\boldsymbol{\beta}]^2 + \sigma^2\mathbf{x}_0((\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1})\mathbf{x}_0^T + \sigma^2 I
\end{aligned}
$$

FOR NOE GRISERI

```r
id <- "1X_8OKcoYbng1XvYFDirxjEWr7LtpNr1m"  # google file ID
values <- dget(sprintf("https://docs.google.com/uc?id=%s&export=download", id))
X = values$X
dim(X)
```

```
## [1] 100  81
```

```r
x0 = values$x0
dim(x0)
```

```
## [1] 81  1
```

```r
beta = values$beta
dim(beta)
```

```
## [1] 81  1
```

```r
sigma = values$sigma
sigma
```

```
## [1] 0.5
```

d)

```r
library(ggplot2)
bias = function(lambda, X, x0, beta) {
    p = ncol(X)
    value = ...
    return(value)
}
lambdas = seq(0, 2, length.out = 500)
BIAS = rep(NA, length(lambdas))
for (i in 1:length(lambdas)) BIAS[i] = bias(lambdas[i], X, x0, beta)
dfBias = data.frame(lambdas = lambdas, bias = BIAS)
ggplot(dfBias, aes(x = lambdas, y = bias)) + geom_line(color = "red") + xlab(expression(lambda)) +
    ylab(expression(bias^2))
```

Comments:

Problem 2

a)