Homework 4

Sohaib Syed

2023-04-02

#Load in Data

```
dataset<-read.table("11502661.txt",
    sep=",",head=T,row.names=1)
x<-dataset$typea #choose single predictor
y<-dataset$chd
```

```
library(splines)
library(scales)
library(ggplot2)

#scale values to be between 0 and 1
x <- rescale(x)


# assume W, weight matrix, is a NxN matrix of 1's
# global linear
H_linear <- bs(x, degree = 1, df=2, intercept = TRUE, Boundary.knots = c(0,1)) # used 2 df because that
sigma_linear <- solve(t(H_linear)%*%H_linear) #Sigma hat
var_linear <- diag(H_linear%*%sigma_linear%*%t(H_linear))#get diagonals

# global cubic
H_cubic <- bs(x, degree = 3, df=4, intercept = TRUE, Boundary.knots = c(0,1)) # used 4 df because that
sigma_cubic <- solve(t(H_cubic)%*%H_cubic)
var_cubic <- diag(H_cubic%*%sigma_cubic%*%t(H_cubic))

# cubic spline  2 knots
H_2s <- bs(x, degree = 3, df=6, intercept = TRUE, knots = c(0.33, 0.66)) # used 6 df because that is wh
sigma_2s <- solve(t(H_2s)%*%H_2s)
var_2s <- diag(H_2s%*%sigma_2s%*%t(H_2s))

# natural cubic spline  6 knots
H_6s <- ns(x, intercept = TRUE, Boundary.knots = c(0.1,0.9), knots = c(.26,.42,.58,.74)) #number of kno
sigma_6s <- solve(t(H_6s)%*%H_6s)
var_6s <- diag(H_6s%*%sigma_6s%*%t(H_6s))


ggplot()+
  geom_point(aes(x,var_linear))+geom_line(aes(x, var_linear,color="Global Linear"))+
  geom_point(aes(x,var_cubic))+geom_line(aes(x, var_cubic,color="Global Cubic"))+
  geom_point(aes(x,var_2s))+geom_line(aes(x, var_2s,color="Cubic Spline-2 Knots"))+
```

```
geom_point(aes(x,var_6s))+geom_line(aes(x, var_6s,color="Natural Spline-6 Knots"))+
ylab("pointwise var")+scale_color_manual(name='Models',
                     breaks=c("Global Linear", "Global Cubic", "Cubic Spline-2 Knots", "Natural Spline-
                     values=c("Global Linear"="orange", "Global Cubic"="red", "Cubic Spline-2 Knots"="g
theme(legend.position = c(0.5, 0.8))
```