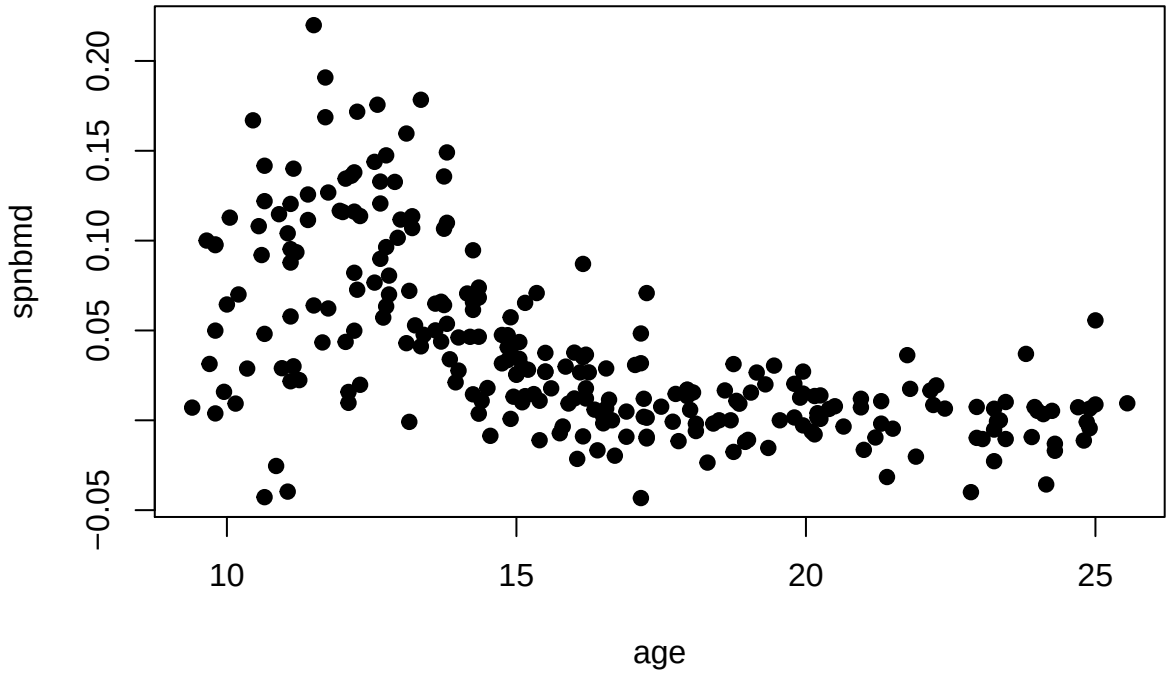
STAT432_HW9

Taiga Hasegawa(taigah2)

2019/4/4

Question1

```
library(ElemStatLearn)
data(bone)
traindata = bone[bone$gender == "female", ]
traindata=traindata[order(traindata$age),]
plot(spnbmd~ age, data = traindata, pch = 19)
```

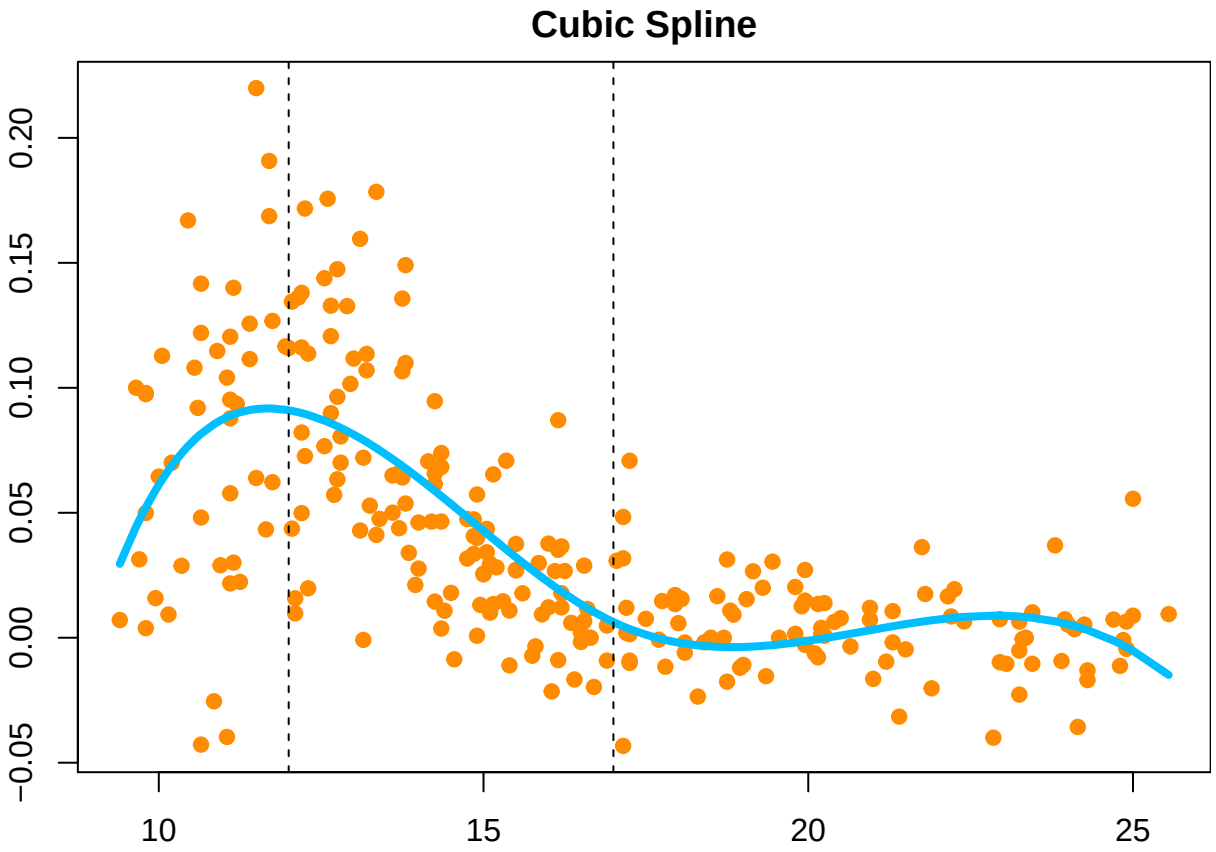



```
myknots = c(12, 17)
```

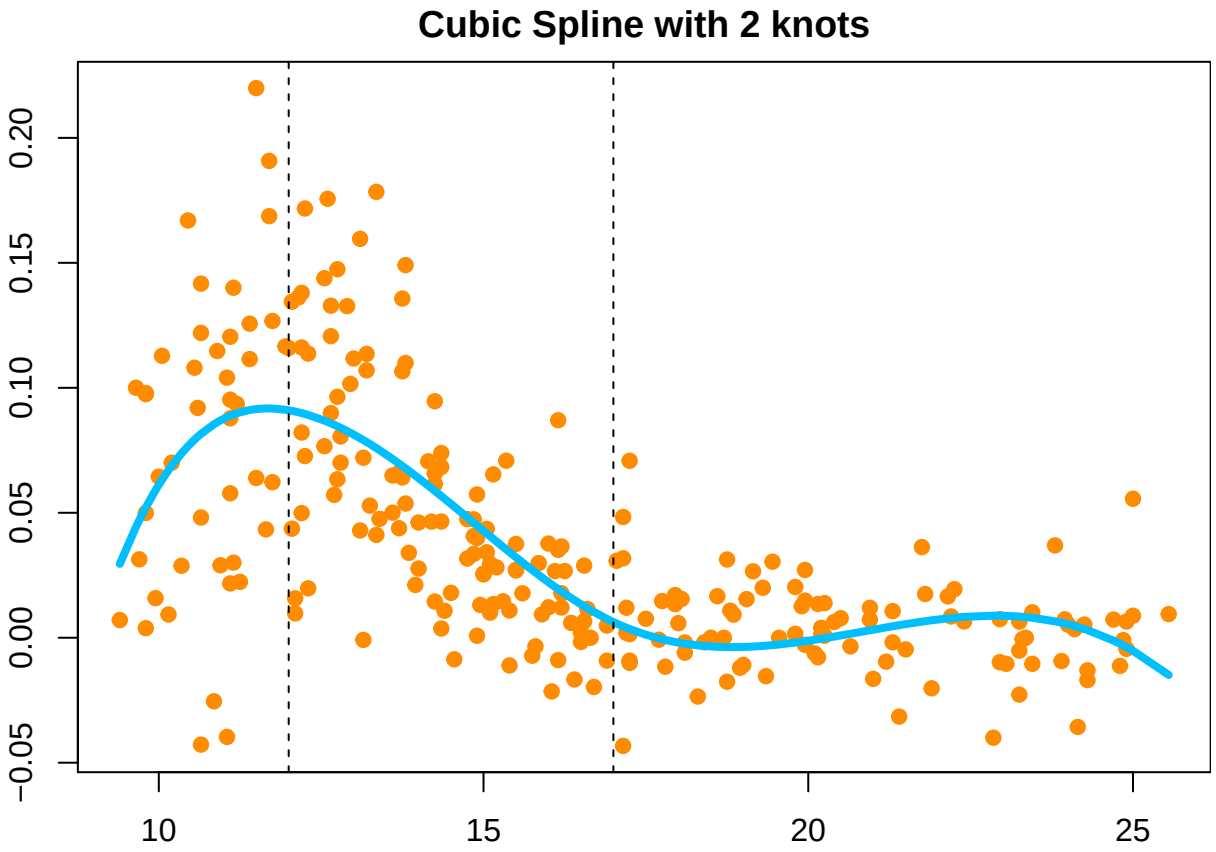
```
pos <- function(x) x^3*(x>0)
mybasis = cbind("int" = 1, "x_1" = traindata$age, "x_2"=(traindata$age)^2,"x_3"=(traindata$age)^3,
                "x_4" = pos(traindata$age - myknots[1]),
                "x_5" = pos(traindata$age - myknots[2]))
df=2+3
```

Degree of freedom is 5.

```
lmfit <- lm(traindata$spnbmd ~ ., data = data.frame(mybasis))
par(mar = c(2,3,2,0))
plot(traindata$spnbmd~traindata$age, pch = 19, col = "darkorange")
lines(traindata$age, lmfit$fitted.values, lty = 1, col = "deepskyblue", lwd = 4)
abline(v = myknots, lty = 2)
title("Cubic Spline")
```




```
lmfit=lm(spnbmd~splines::bs(age,degree=3,knots = myknots),data=traindata)
par(mar = c(2,3,2,0))
plot(spnbmd~age, pch = 19, col = "darkorange",data = traindata)
lines(traindata$age, lmfit$fitted.values, lty = 1, col = "deepskyblue", lwd = 4)
abline(v = myknots, lty = 2)
title("Cubic Spline with 2 knots")
```

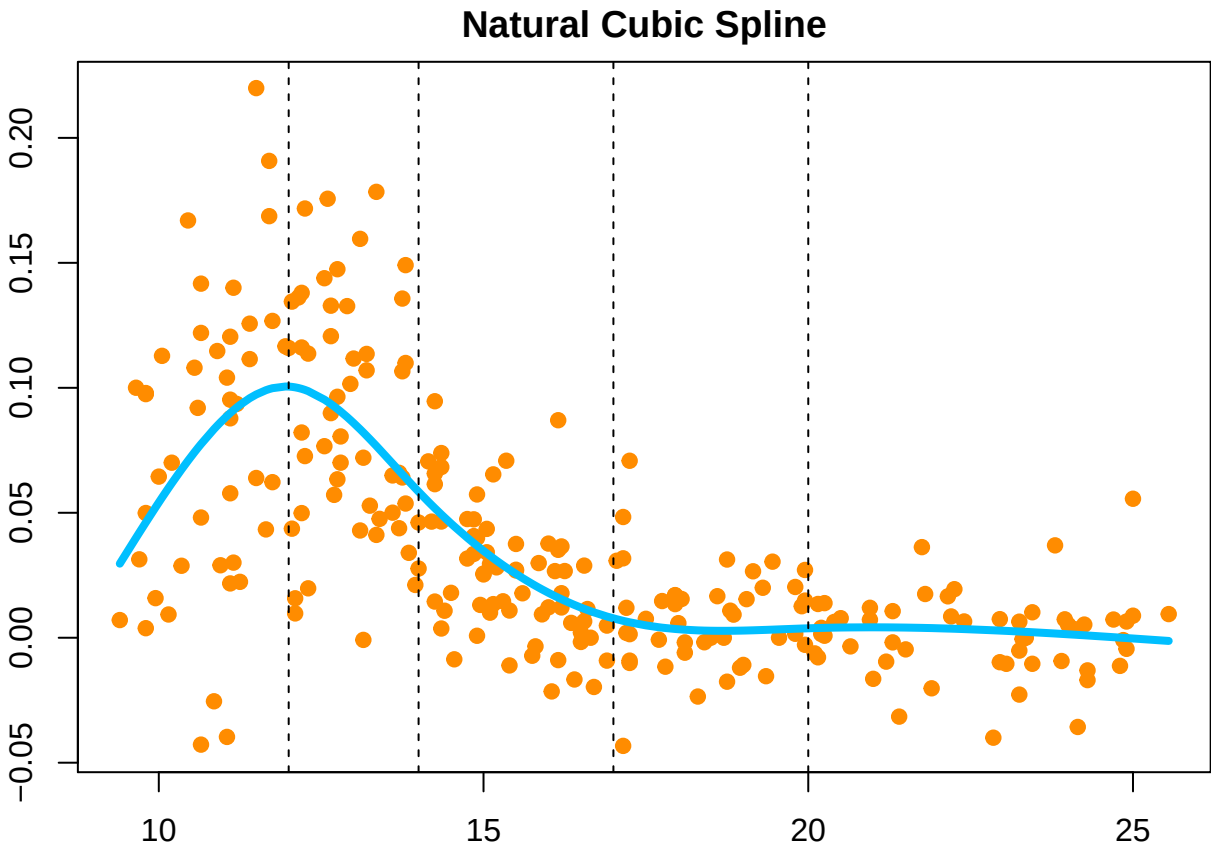



```
myknots=c(12,14,17,20)
```

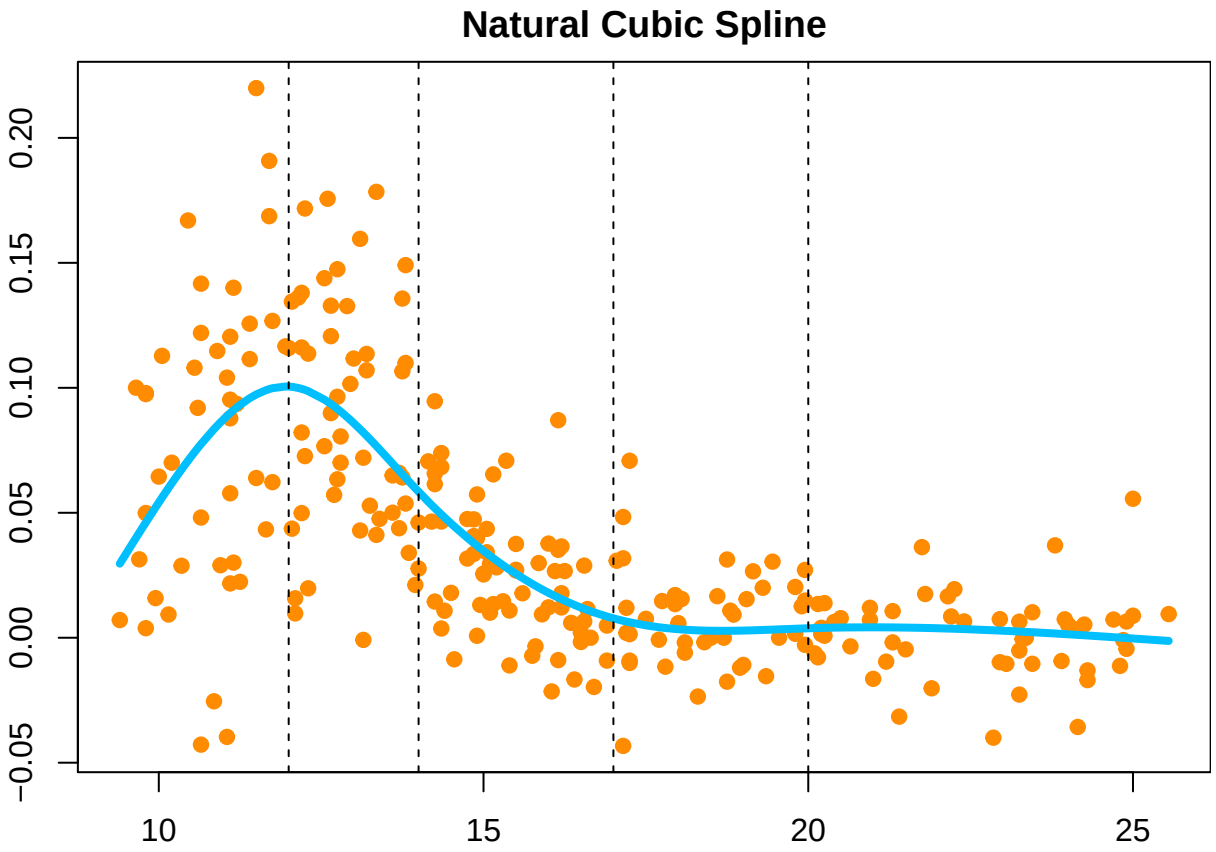
```
d <- function(x=traindata$age,k){
  ((x-myknots[k])^3*((x-myknots[k])>0)-(x-max(x))^3*((x-max(x))>0))/(max(x)-myknots[k])
}
x=traindata$age
mybasis = cbind("N_1" = 1, "N_2" = traindata$age,
                "N_3"=(((x-min(x))^3*((x-min(x))>0)-(x-max(x))^3*((x-max(x))>0))/(max(x)-min(x)))-d(k=4
                "N_4"=d(k=1)-d(k=4),
                        "N_5" = d(k=2)-d(k=4),
                "N_6"=d(k=3)-d(k=4))
df=5
```

Degree of freedom is 5.

```
lmfit <- lm(traindata$spnbmd ~ .-1, data = data.frame(mybasis))
par(mar = c(2,3,2,0))
plot(traindata$spnbmd~traindata$age, pch = 19, col = "darkorange")
lines(traindata$age, lmfit$fitted.values, lty = 1, col = "deepskyblue", lwd = 4)
abline(v = myknots, lty = 2)
title("Natural Cubic Spline")
```




```
library(splines)
lmfit=lm(spnbmd~ns(age,knots = myknots),data=traindata)
par(mar = c(2,3,2,0))
plot(traindata$spnbmd~traindata$age, pch = 19, col = "darkorange")
lines(traindata$age, lmfit$fitted.values, lty = 1, col = "deepskyblue", lwd = 4)
abline(v = myknots, lty = 2)
title("Natural Cubic Spline")
```



Question2

```
data(ozone)
head(ozone)
```

```
##   ozone radiation temperature wind
## 1    41       190          67  7.4
## 2    36       118          72  8.0
## 3    12       149          74 12.6
## 4    18       313          62 11.5
## 5    23       299          65  8.6
## 6    19        99          59 13.8
```

```
    library(ElemStatLearn)
    library(gam)
```

```
## Loading required package: foreach
```

```
## Loaded gam 1.16
```

```
    form = formula("ozone ~ ns(radiation,df=4) + ns(temperature,df=4) + ns(wind,df=4)")

    m = gam(form, data=ozone)
    summary(m)
```

```
##
## Call: gam(formula = form, data = ozone)
```

```
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -49.311 -10.484  -1.915   8.745  71.341
##
## (Dispersion Parameter for gaussian family taken to be 308.8443)
##
##     Null Deviance: 121801.9 on 110 degrees of freedom
## Residual Deviance: 30266.74 on 98 degrees of freedom
## AIC: 965.5228
##
## Number of Local Scoring Iterations: 2
##
## Anova for Parametric Effects
##                           Df Sum Sq Mean Sq F value    Pr(>F)
## ns(radiation, df = 4)      4  30645  7661.3  24.806 3.357e-14 ***
## ns(temperature, df = 4)    4  44920 11229.9  36.361 < 2.2e-16 ***
## ns(wind, df = 4)           4  15970  3992.6  12.927 1.723e-08 ***
## Residuals                 98  30267   308.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
par(mfrow = c(1, 3), mar = c(5, 5, 2, 0))
plot(m, se = TRUE, residuals = TRUE)
```