Splines

Mats Hansson

12 mars 2016

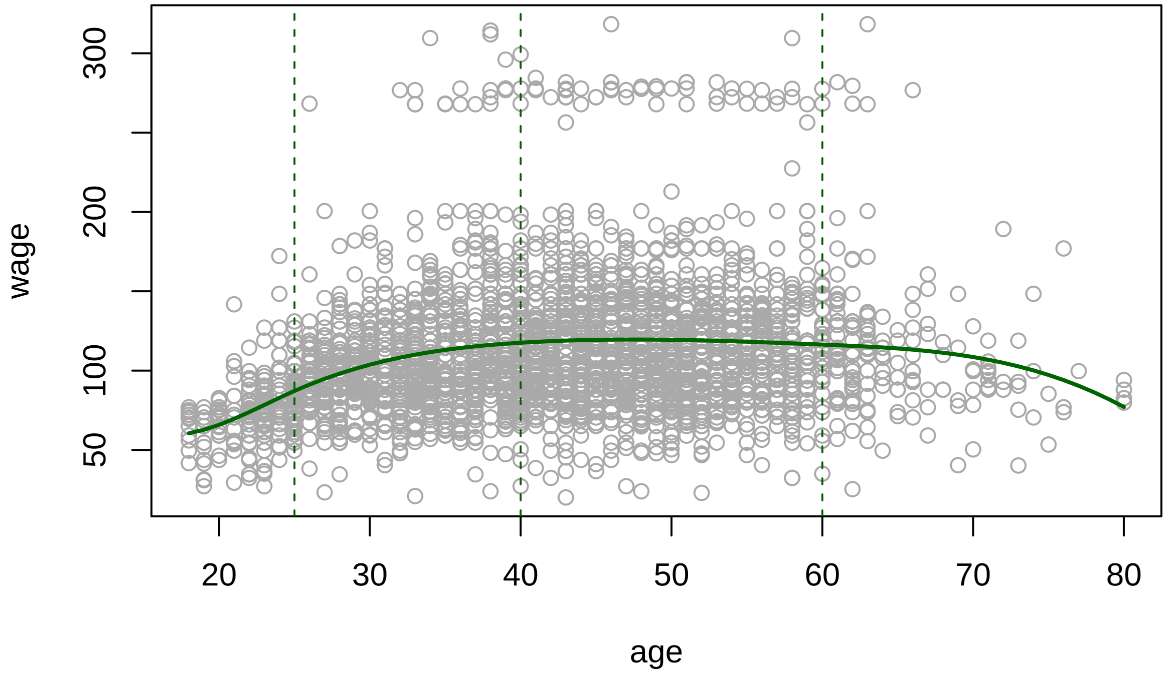
Splines

Splines are more flixible than polynomials, but the idea is rather similar. Here will we explore cubic spline.

```
library(splines)
library(ISLR)
attach(Wage)

fit=lm(wage~bs(age, knots=c(25,40,60)), data=Wage)
plot(age, wage, col="darkgrey")

age.grid=seq(from=min(age), to=max(age))
lines(age.grid, predict(fit, list(age=age.grid)), col="darkgreen", lwd=2)
abline(v=c(25,40,60), col="darkgreen", lty=2)
```

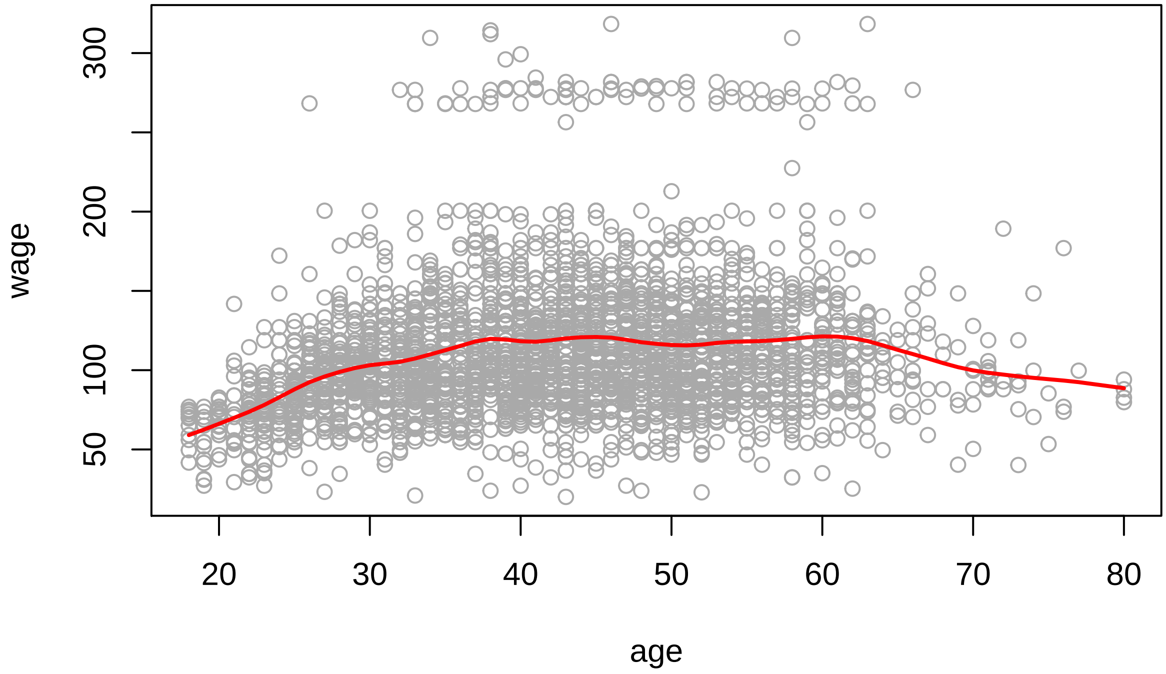


```
### How to make it in matrix algebra ###
########################################
```

```
### m=model.matrix( ~ bs(age.grid,knots=c(25, 40, 60)))
### p<-(m%*%as.matrix(coef(fit)))
```

The smoothing splines doesnot require knot selection, but it does have a smoothing parameters, which can conveniently be specified via the effective degrees of freedom or `df`.

```
plot(age, wage, col="darkgrey")
fit.smooth=smooth.spline(age, wage, df=16)
lines(fit.smooth, col="red", lwd=2)
```

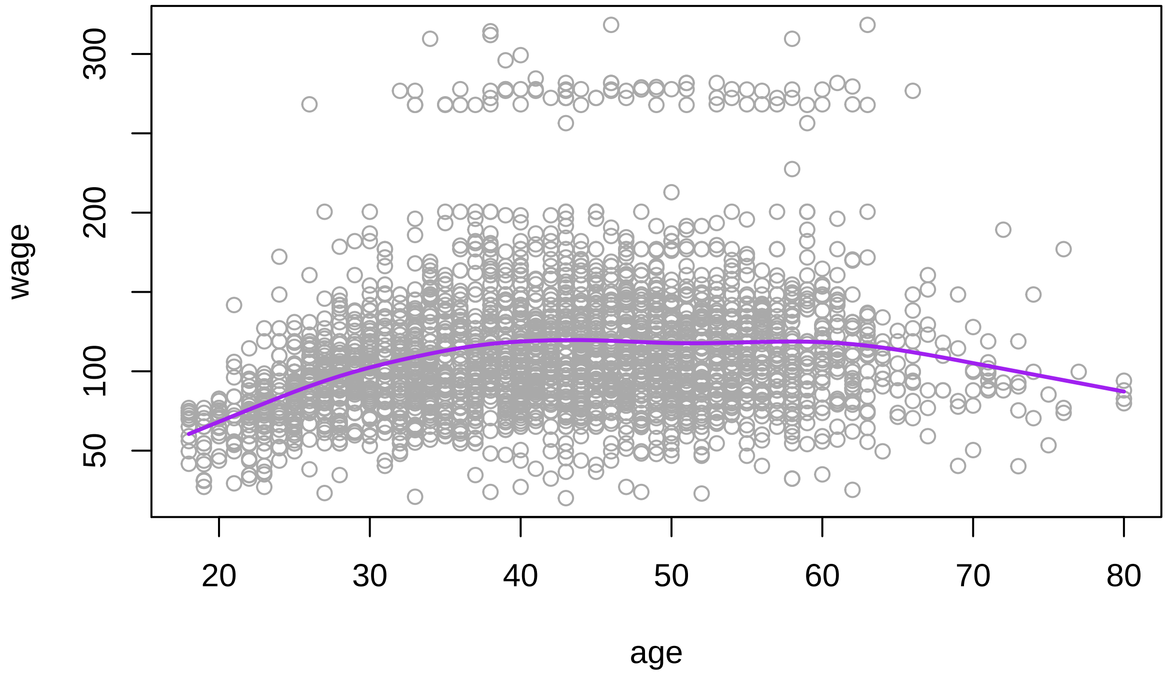


Or we can use leave one out cross-validation to select the smoothing parameters automatically:

```
plot(age, wage, col="darkgrey")
fit.CV=smooth.spline(age, wage, cv=TRUE)
```

```
## Warning in smooth.spline(age, wage, cv = TRUE): cross-validation with non-
## unique 'x' values seems doubtful
```

```
lines(fit.CV, col="purple", lwd=2)
```

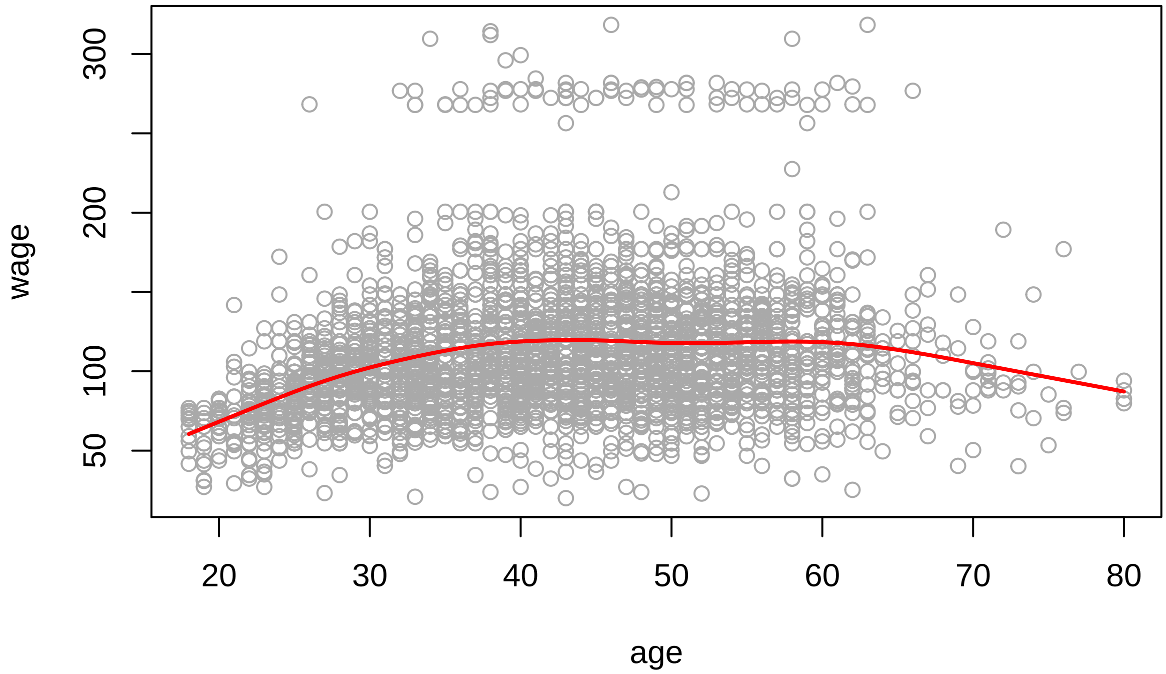


```
fit.CV
```

```
## Call:
## smooth.spline(x = age, y = wage, cv = TRUE)
##
## Smoothing Parameter  spar= 0.6988943  lambda= 0.02792303 (12 iterations)
## Equivalent Degrees of Freedom (Df): 6.794596
## Penalized Criterion: 75215.9
## PRESS: 1593.383
```

Where the Degress of freesom is 6.794596 which com compare to our first model that have 6 degree of freedom. Therefore we seet the lambda to 6.794596 in the model above

```
fit.smooth_1=smooth.spline(age, wage, df=6.794596)
plot(age, wage, col="darkgrey")
lines(fit.smooth_1, col="red", lwd=2)
```


wage
300
200
100
50
20
30
40
50
60
70
80
age