KNN Regression

Gregoire Gasparini, Aurora Hofman, Sarah Musiol, Beatriu Tort

14/02/20

KNN Regression

Preliminaries

Reading in the data.

```
library(MASS)
data(Boston)
help(Boston)
```

Set variables x and y.

```
x <- Boston$lstat
y <- Boston$medv
#plot(x,y)
```

1. Task

Write a function for the knn regression estimator.

```
knn.reg <- function(t, x, y, k = 50) {
  d_t_x <- as.matrix(dist(c(t, x)))[1, -1]
  d_t_x_k <- sort(d_t_x, partial = k)[k]
  N_t_k <-  which(d_t_x <= d_t_x_k)
  m_hat <- sum(y[N_t_k]) / k
  return(m_hat)
}
```

2. and 3. Task

Define t as a sequence from 1 to 40. Estimation of the regression values for multiple values for k.

```
t <- 1:40
m_hat_500 <- rep(0, length(t))
m_hat_100 <- rep(0, length(t))
m_hat_50 <- rep(0, length(t))
m_hat_30 <- rep(0, length(t))
m_hat_10 <- rep(0, length(t))
for (i in 1:length(t)) {
  m_hat_500[i] <- knn.reg(t[i], x, y, 500)
  m_hat_100[i] <- knn.reg(t[i], x, y, 100)
  m_hat_50[i] <- knn.reg(t[i], x, y, 50)
  m_hat_30[i] <- knn.reg(t[i], x, y, 30)
  m_hat_10[i] <- knn.reg(t[i], x, y, 10)
}
```

4. Task

Plot x against y and represent the regression function.

```
plot(x,y, pch = 20, main = "50 nearest neighbours")
lines(m_hat_50, col = "red", lwd = 2)
```

50 nearest neighbours
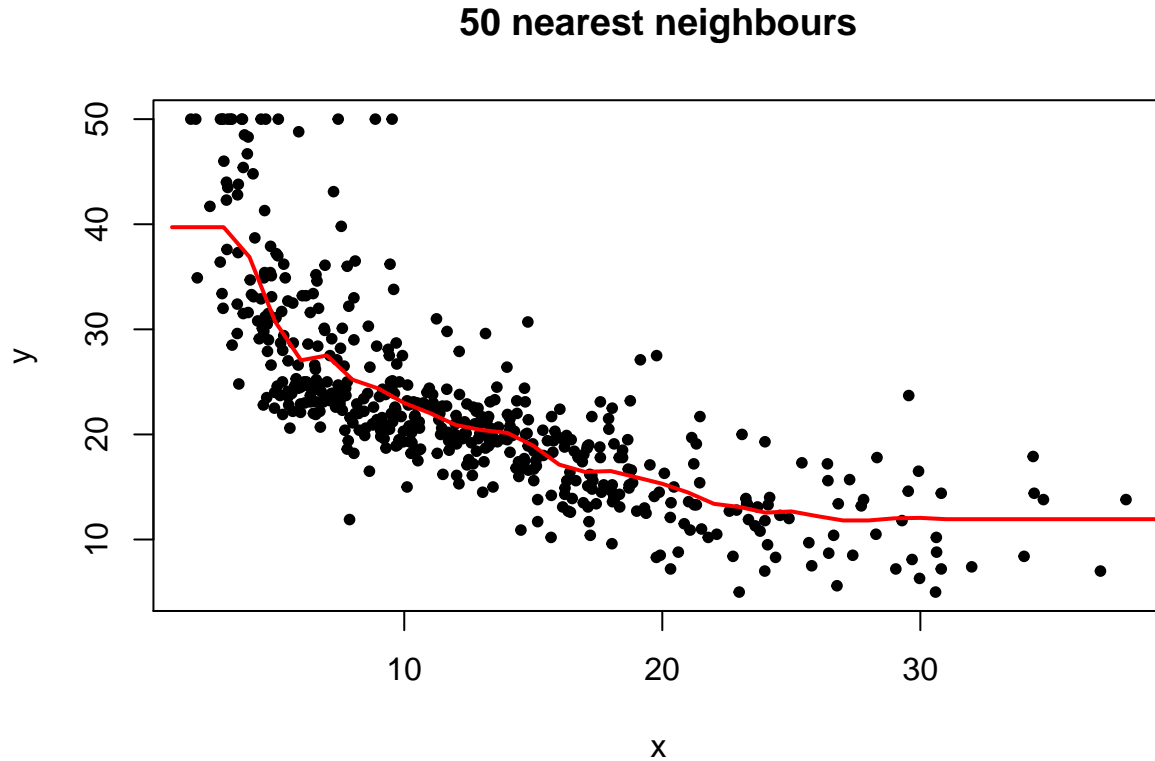
5. Task

Represent Regression function for different values of k.

```
par(mfrow = c(2,2))
plot(x,y, pch = 20, main = "500 nearest neighbours")
lines(m_hat_500, col = "red", lwd = 2)
plot(x,y, pch = 20, main = "100 nearest neighbours")
lines(m_hat_100, col = "red", lwd = 2)
plot(x,y, pch = 20, main = "30 nearest neighbours")
lines(m_hat_30, col = "red", lwd = 2)
plot(x,y, pch = 20, main = "10 nearest neighbours")
lines(m_hat_10, col = "red", lwd = 2)
```

500 nearest neighbours

100 nearest neighbours

30 nearest neighbours

10 nearest neighbours
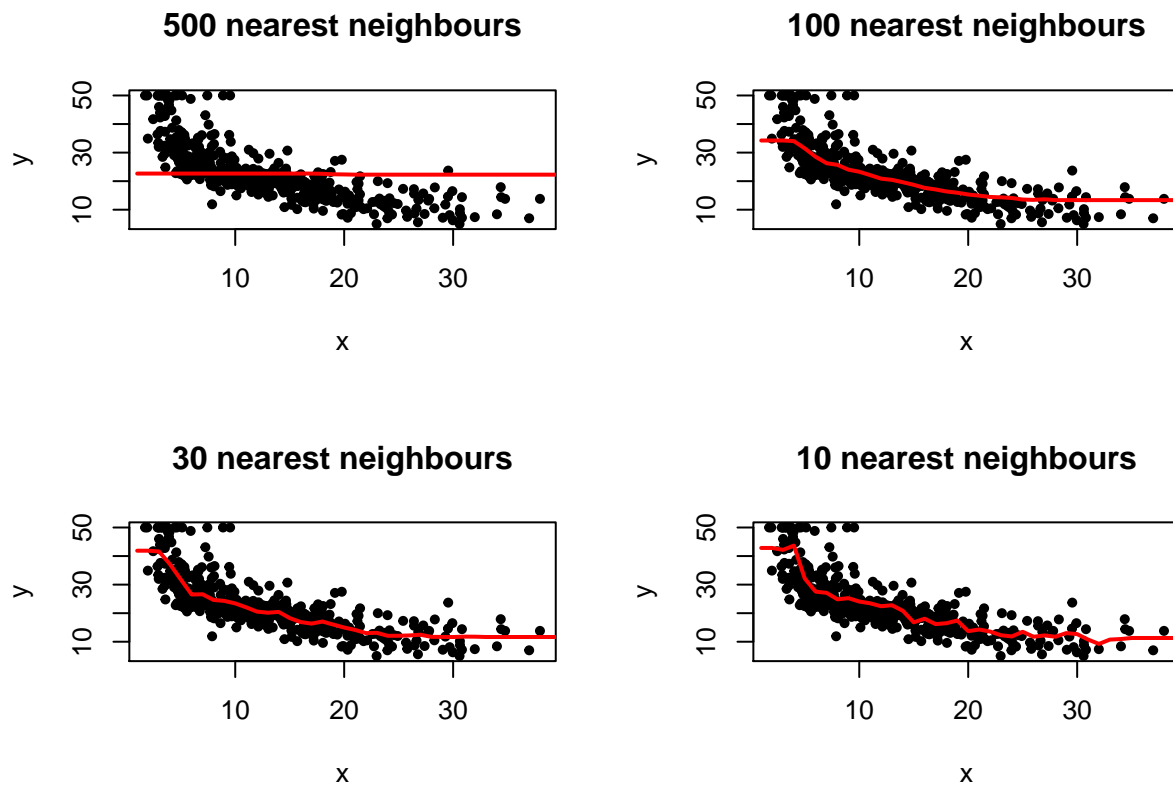
```r
par(mfrow = c(1,1))
```

Conclusion

Based on the plots, one can see that using a large amount of nearest neighbours the regression line tends to a constant line. While using a small amount of nearest neighbour the results are not as smooth.