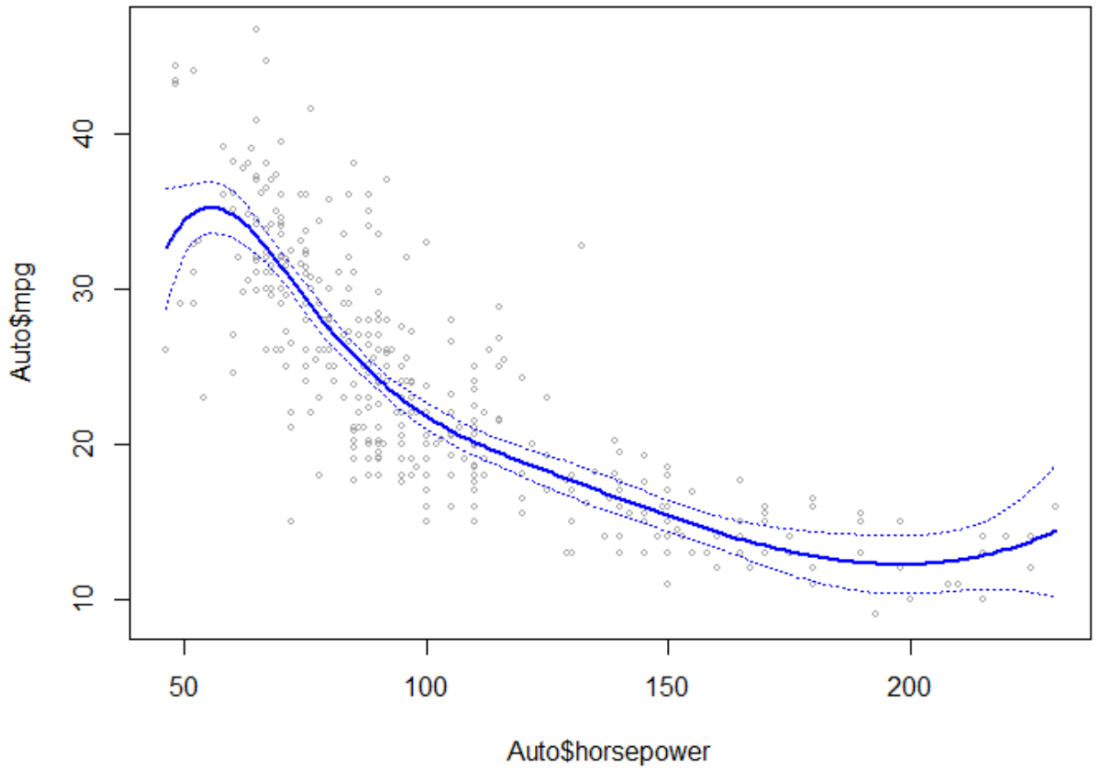
For the Homework 5, we will use the Auto data set. Load the ISLR package to get the Auto data set. Fit below non-linear models to the Auto data set. We will treat horsepower as the predictor and mpg as the response.

1. **Fit the cubic spline with 3 knots (25th percentile, 50th percentile, and 75th percentile of horsepower)**

```r
library(ISLR)
library(splines)
head(Auto)
attach(Auto)

horsepower.limits <- range(horsepower) #range of all values of age from
#smallest to largest
horsepower.limits # 18 - 80
horsepower.grid <- seq(horsepower.limits[1], horsepower.limits[2], 1)
horsepower.grid

splines.fit <- lm(mpg ~ bs(horsepower, df = 6), data = Auto)
pred <- predict(splines.fit, newdata = list(horsepower = horsepower.grid),
                se = TRUE)
dim(bs(horsepower, df = 6))
attr(bs(horsepower, df = 6), "knots")

plot(Auto$horsepower, Auto$mpg, xlim = horsepower.limits, cex = 0.5,
     col = "darkgrey")
lines(horsepower.grid, pred$fit, lwd = 2, col = "blue")

se.bands <- cbind(pred$fit + 2*pred$se.fit, pred$fit - 2*pred$se.fit)
matlines(horsepower.grid, se.bands, lwd = 1, col = "blue", lty = 3)
```

```r
library(ISLR)
library(splines)
head(Auto)
attach(Auto)

horsepower.limits <- range(horsepower) #range of all values of age from
#smallest to largest
horsepower.limits # 18 - 80
horsepower.grid <- seq(horsepower.limits[1], horsepower.limits[2], 1)
horsepower.grid

##1. CUBIC SPLINE:
splines.fit <- lm(mpg ~ bs(horsepower, df = 6), data = Auto)
pred <- predict(splines.fit, newdata = list(horsepower = horsepower.grid),
            se = TRUE)
dim(bs(horsepower, df = 6))
attr(bs(horsepower, df = 6), "knots")

plot(Auto$horsepower, Auto$mpg, xlim = horsepower.limits, cex = 0.5,
    col = "darkgrey")
lines(horsepower.grid, pred$fit, lwd = 2, col = "blue")

se.bands <- cbind(pred$fit + 2*pred$se.fit, pred$fit - 2*pred$se.fit)
```

matlines(horsepower.grid, se.bands, lwd = 1, col = "blue", lty = 3)



2. Fit the natural spline with 3 knots (25th percentile, 50th percentile, and 75th percentile of horsepower)

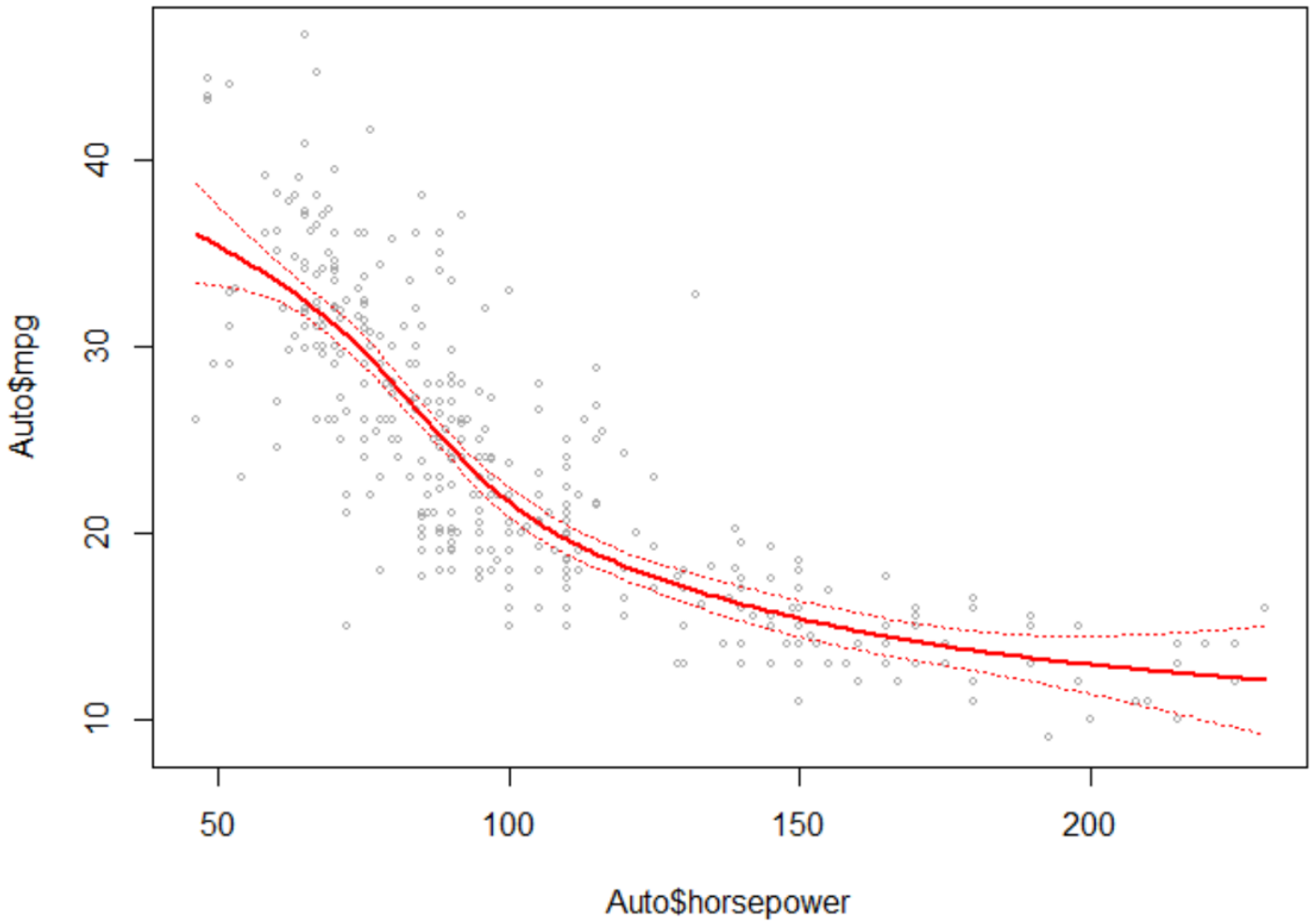
```
##2. NATURAL CUBIC SPLINE:

natural.splines.fit <- lm(mpg ~ ns(horsepower, df = 4), data = Auto)
pred2 <- predict(natural.splines.fit,
                 newdata = list(horsepower = horsepower.grid), se = T)

dim(ns(Auto$horsepower, df = 4))
attr(ns(Auto$mpg, df = 4), "knots")

plot(Auto$horsepower, Auto$mpg, xlim = horsepower.limits,
     cex = 0.5, col = "darkgrey")
lines(horsepower.grid, pred2$fit, lwd = 2, col = "red")
se.bands2 <- cbind(pred2$fit + 2*pred2$se.fit, pred2$fit -
                     2*pred2$se.fit)
matlines(horsepower.grid, se.bands2, lwd = 1, col = "red", lty = 3)
```

natural.splines.fit <- lm(mpg ~ ns(horsepower, df = 4), data = Auto)
pred2 <- predict(natural.splines.fit,
          newdata = list(horsepower = horsepower.grid), se = T)

dim(ns(Auto$horsepower, df = 4))
attr(ns(Auto$mpg, df = 4), "knots")

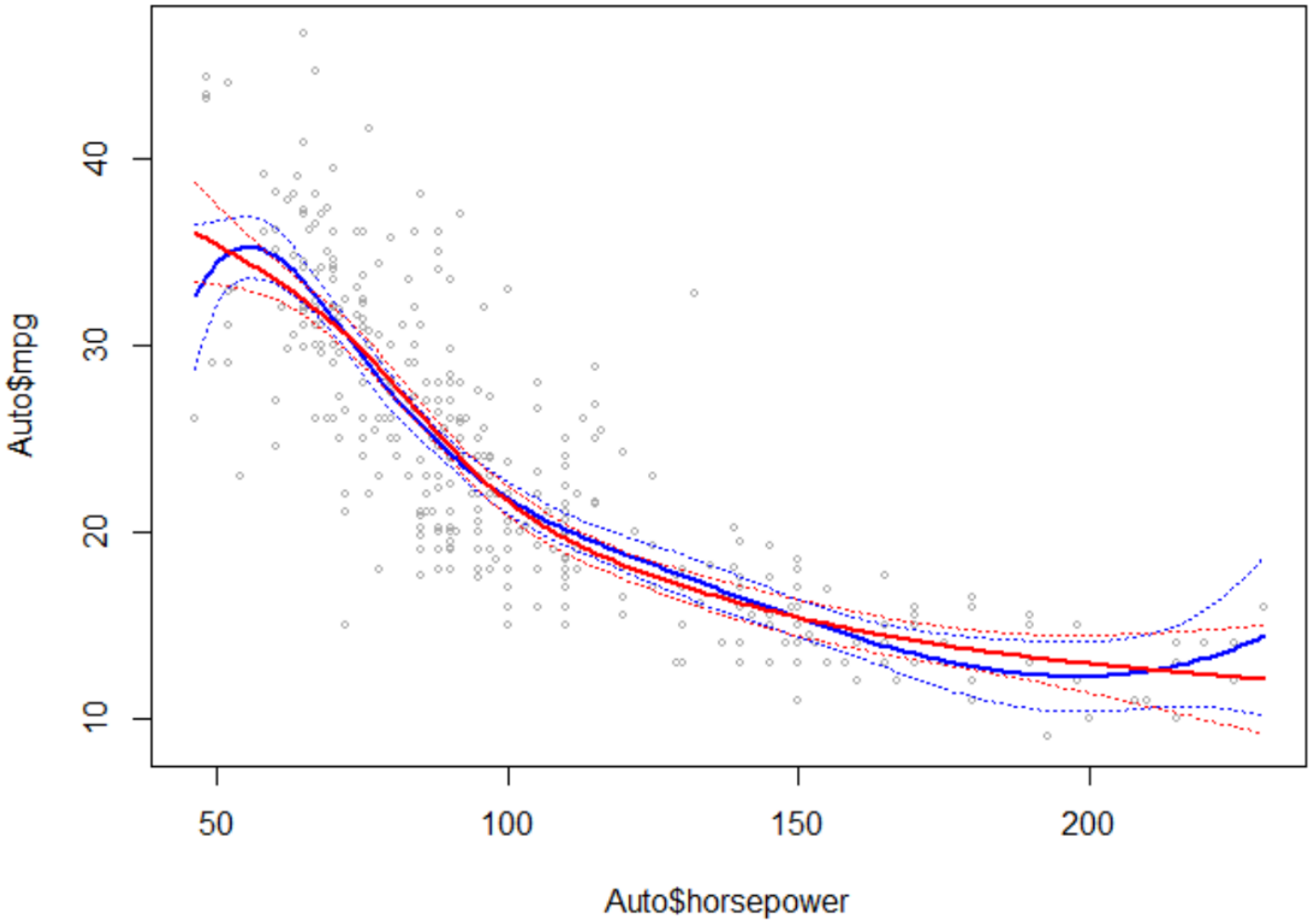
plot(Auto$horsepower, Auto$mpg, xlim = horsepower.limits,
    cex = 0.5, col = "darkgrey")

```
lines(horsepower.grid, pred2$fit, lwd = 2, col = "red")
se.bands2 <- cbind(pred2$fit + 2*pred2$se.fit, pred2$fit -
                   2*pred2$se.fit)
matlines(horsepower.grid, se.bands2, lwd = 1, col = "red", lty = 3)
```




```
plot(Auto$horsepower, Auto$mpg, xlim = horsepower.limits,
     cex = 0.5, col = "darkgrey")
lines(horsepower.grid, pred$fit, lwd = 2, col = "blue")
matlines(horsepower.grid, se.bands, lwd = 1, col = "blue", lty = 3)
lines(horsepower.grid, pred2$fit, lwd = 2, col = "red")
matlines(horsepower.grid, se.bands2, lwd = 1, col = "red", lty = 3)
```


```
plot(Auto$horsepower, Auto$mpg, xlim = horsepower.limits,
    cex = 0.5, col = "darkgrey")
lines(horsepower.grid, pred$fit, lwd = 2, col = "blue")
matlines(horsepower.grid, se.bands, lwd = 1, col = "blue", lty = 3)
lines(horsepower.grid, pred2$fit, lwd = 2, col = "red")
matlines(horsepower.grid, se.bands2, lwd = 1, col = "red", lty = 3)
```



3. **Fit the local regression with span = 0.75. Also, plot the data and add the non-linear fits to the plot.**

```
49  ##2. LOCAL REGRESSION:
50
51  local.fit <- loess(mpg ~ horsepower, span = 0.75, data = Auto)
52
53  pred <- predict(local.fit, newdata =
54                    data.frame(horsepower = horsepower.grid))
55
56  plot(Auto$horsepower, Auto$mpg, xlim =
57         range(Auto$horsepower), cex = 0.5, col = "darkgrey")
58  lines(horsepower.grid, pred, col = "blue", lwd = 2)
59  legend("topright", legend = c("Span = 0.75"),
60         col = c("blue"), lty = 1, lwd = 2, cex = 0.8)
```

```
local.fit <- loess(mpg ~ horsepower, span = 0.75, data = Auto)

pred <- predict(local.fit, newdata =
          data.frame(horsepower = horsepower.grid))
```

```
plot(Auto$horsepower, Auto$mpg, xlim =
      range(Auto$horsepower), cex = 0.5, col = "darkgrey")
lines(horsepower.grid, pred, col = "blue", lwd = 2)
legend("topright", legend = c("Span = 0.75"),
      col = c("blue"), lty = 1, lwd = 2, cex = 0.8)
```