STATS 413 PROBLEM SET 2 SOLUTION

This problem set is due at **noon ET** on **Sep 23, 2021**. Please upload your solutions to Canvas in two files: a PDF file containing the solutions and a ZIP file containing code that reproduces any computer output in the solutions. You are encouraged to collaborate on problem sets with your classmates, but the final write-up (including any code) **must be your own**.

1. Multi-task linear regression. Thus far, we only considered regression with scalar-valued outcomes. In some applications, the outcome is itself a vector: $\mathbf{y}_i \in \mathbf{R}^K$. We posit the relationship between the features and the vector-valued outcome is linear:

$$\mathbf{y}_i^T \approx \mathbf{x}_i^T \widehat{B}, \tag{1.1}$$

for some matrix of regression coefficients $\widehat{B} \in \mathbf{R}^{p\times K}$.

(a) The sum of squared residuals (SSR) here is

$$\mathrm{SSR}(B) \triangleq \sum_{i=1}^{n}\sum_{k=1}^{K}\frac{1}{2}(\mathbf{y}_{i,k} - \mathbf{x}_i^T b_k)^2,$$

where $b_k \in \mathbf{R}^p$ is the k-th column of B. Express $\mathrm{SSR}(B)$ in matrix notation (*i.e.* without using any explicit summations).

Hint: work out how to express the SSR in terms of

$$\mathbf{X} = \begin{bmatrix} - & \mathbf{x}_1^T & - \\ & \vdots & \\ - & \mathbf{x}_n^T & - \end{bmatrix} \in \mathbf{R}^{n\times p}, \quad \mathbf{y} = \begin{bmatrix} - & \mathbf{y}_1^T & - \\ & \vdots & \\ - & \mathbf{y}_n^T & - \end{bmatrix} \in \mathbf{R}^{n\times K}.$$

Solution:

$$\mathrm{SSR}(B) \triangleq \sum_{i=1}^{n}\sum_{k=1}^{K}\frac{1}{2}(\mathbf{y}_{i,k} - \mathbf{x}_i^T b_k)^2 = \frac{1}{2}\sum_{k=1}^{K}(y_k - XB_k)^T(y_k - XB_k) = \frac{1}{2}tr((y-XB)^T(y-XB))$$

(b) Find a closed-form expression for the matrix of regression coefficients that minimizes the SSR; *i.e.* find a (closed-form) expression for $\widehat{B} \in \arg\min_{B\in\mathbf{R}^{p\times K}}\mathrm{SSR}(B)$.

Solution: From (a), we get

$$\mathrm{SSR}(B) \triangleq \sum_{i=1}^{n}\sum_{k=1}^{K}\frac{1}{2}(\mathbf{y}_{i,k} - \mathbf{x}_i^T b_k)^2 = \frac{1}{2}\sum_{k=1}^{K}(y_k - XB_k)^T(y_k - XB_k) = \frac{1}{2}tr((y-XB)^T(y-XB))$$

Let's take derivative of $\mathrm{SSR}(B)$ towards B: $\frac{\partial SSR(B)}{\partial B} = \frac{1}{2}(-2X^Ty + 2X^TXB)$

So let the derivative = 0, we get $\widehat{B} = (X^TX)^{-1}X^TY$

(c) Instead of minimizing the SSR, we break up the problem into K separate regression problems with scalar-valued responses. That is, we fit K linear models of the form

$$\mathbf{y}_{i,k} \approx \mathbf{x}_i^T \widehat{\boldsymbol{\beta}}_k,$$

where $\mathbf{y}_{i,k}$ is k-th outcome of the i-th sample (i.e. the k-th entry of $\mathbf{y}_i$) and $\widehat{\boldsymbol{\beta}}_k \in \mathbf{R}^p$ are the regression coefficients of the k-th linear model. How are the fitted coefficients from the K separate regressions $\widehat{\boldsymbol{\beta}}_1, \ldots, \widehat{\boldsymbol{\beta}}_K$ related to the matrix of regression coefficients that minimizes the SSR $\widehat{B}$?

Solution: $\widehat{B} = \begin{bmatrix} \widehat{\boldsymbol{\beta}}_1^T & \cdots & \widehat{\boldsymbol{\beta}}_K^T \end{bmatrix}$ Since each $SSR(B_k)$ is independent, as long as for each B_k, the $SSR(B_k)$ is minimum, then the $SSR(B)$ will be minimun. So $\widehat{B}$ is the combination of all the $\widehat{\boldsymbol{\beta}}_i$

Lab 1: Linear regression. In this problem, we will predict the per capita crime using the other variables in the `Boston` dataset.

(a) For each predictor, fit a simple linear regression model to predict the response. In which of the simple linear models is there a statistically significant association between the predictor and the response.

(b) Fit a (multiple) regression model to predict the response using all the other features in the dataset. For which features can we reject the null $H_0 : \beta_j = 0$.

(c) How do the results from (a) and (b) compare. Create a scatterplot displaying the simple regression coefficient of each predictor from (a) on the x-axis, and the multiple regression coefficient from (b) on the y-axis. That is, each predictor is displayed as a point on the plot.

(d) Is there evidence of non-linear relationship between any of the features and response? For each predictor $\mathbf{x}_j$, look at the fit of the cubic model

$$\mathbf{y} \sim \beta_0 + \beta_1 \mathbf{x}_j + \beta_2 \mathbf{x}_j^2 + \beta_3 \mathbf{x}_j^3.$$

PS2 Solution

Runxia Zhao

R Markdown

```
library(MASS)
data("Boston")
attach(Boston)
```

```
lmfit.zn = lm(crim~zn)
summary(lmfit.zn)
```

```
##
## Call:
## lm(formula = crim ~ zn)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.429 -4.222 -2.620  1.250 84.523
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.45369    0.41722  10.675  < 2e-16 ***
## zn          -0.07393    0.01609  -4.594 5.51e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.435 on 504 degrees of freedom
## Multiple R-squared:  0.04019,    Adjusted R-squared:  0.03828
## F-statistic:  21.1 on 1 and 504 DF,  p-value: 5.506e-06
```

Since the p-value of the crim vs zn model is 0.05, meaning the chance of having a null hypothesis (β0) is very low. Therefore we conclude that there is a statistically significant association between crim and zn.

```
lmfit.indus = lm(crim~indus)
summary(lmfit.indus)
```

```
##
## Call:
## lm(formula = crim ~ indus)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -11.972  -2.698  -0.736   0.712  81.813
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.06374    0.66723  -3.093  0.00209 **
## indus        0.50978    0.05102   9.991  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.866 on 504 degrees of freedom
## Multiple R-squared:  0.1653, Adjusted R-squared:  0.1637
## F-statistic: 99.82 on 1 and 504 DF,  p-value: < 2.2e-16
```

#There is a statistically significant relationship between the per capita and Indus. This is because the p-value of the model is 2e-16 which is far less than 0.05

```
lmfit.chas = lm(crim~chas)
summary(lmfit.chas)
```

```
##
## Call:
## lm(formula = crim ~ chas)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.738 -3.661 -3.435  0.018 85.232
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.7444     0.3961   9.453   <2e-16 ***
## chas         -1.8928     1.5061  -1.257    0.209
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.597 on 504 degrees of freedom
## Multiple R-squared:  0.003124,   Adjusted R-squared:  0.001146
## F-statistic: 1.579 on 1 and 504 DF,  p-value: 0.2094
```

#The p-value of the model is 0.209 which is must great than 0.05 and this means that the chances of having a null hypothesis are high and therefore chas is not statistically significant

```
lmfit.nox = lm(crim~nox)
summary(lmfit.nox)
```

```
##
## Call:
## lm(formula = crim ~ nox)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.371  -2.738  -0.974   0.559  81.728
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)  -13.720      1.699  -8.073 5.08e-15 ***
## nox           31.249      2.999  10.419  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.81 on 504 degrees of freedom
## Multiple R-squared:  0.1772, Adjusted R-squared:  0.1756 
## F-statistic: 108.6 on 1 and 504 DF,  p-value: < 2.2e-16
```

#The 2e-16 p-value of the model is far less than 0.05 which makes the relationship between per capita crime rate and nitrogen oxide concentration statistically significant.

```
lmfit.rm = lm(crim~rm)
summary(lmfit.rm)
```

```
## 
## Call:
## lm(formula = crim ~ rm)
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -6.604 -3.952 -2.654  0.989 87.197 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   20.482      3.365   6.088 2.27e-09 ***
## rm            -2.684      0.532  -5.045 6.35e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 8.401 on 504 degrees of freedom
## Multiple R-squared:  0.04807,    Adjusted R-squared:  0.04618 
## F-statistic: 25.45 on 1 and 504 DF,  p-value: 6.347e-07
```

#There is a statistically significant association between the per capita crime rate and the average number of rooms per dwelling(rm) because of the p- of 6.35e-7 is smaller than 0.05

```
lmfit.age = lm(crim~age)
summary(lmfit.age)
```

```
## 
## Call:
## lm(formula = crim ~ age)
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -6.789 -4.257 -1.230  1.527 82.849 
## 
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -3.77791    0.94398  -4.002 7.22e-05 ***
## age          0.10779    0.01274   8.463 2.85e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 8.057 on 504 degrees of freedom
## Multiple R-squared:  0.1244, Adjusted R-squared:  0.1227 
## F-statistic: 71.62 on 1 and 504 DF,  p-value: 2.855e-16
```

#There exists a statistically significant relationship between per capita crime and age because of the low p-value of 2.85e-16.

```
lmfit.dis = lm(crim~dis)
summary(lmfit.dis)
```

```
## 
## Call:
## lm(formula = crim ~ dis)
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -6.708 -4.134 -1.527  1.516 81.674 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   9.4993     0.7304  13.006   <2e-16 ***
## dis          -1.5509     0.1683  -9.213   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.965 on 504 degrees of freedom
## Multiple R-squared:  0.1441, Adjusted R-squared:  0.1425 
## F-statistic: 84.89 on 1 and 504 DF,  p-value: < 2.2e-16
```

#Because of the small p-value of 2e-16, there exists a statistically significant association between per capita crime rate and dis variable

```
linear_fit.rad <- lm(crim ~ rad)
summary(linear_fit.rad)
```

```
## 
## Call:
## lm(formula = crim ~ rad)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -10.164  -1.381  -0.141   0.660  76.433 
## 
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.28716    0.44348  -5.157 3.61e-07 ***
## rad          0.61791    0.03433  17.998  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.718 on 504 degrees of freedom
## Multiple R-squared:  0.3913, Adjusted R-squared:   0.39
## F-statistic: 323.9 on 1 and 504 DF,  p-value: < 2.2e-16
```

#There is a statistically significant association between the per capita crime rate and rad because of the low p-value of 2e-16.

```
linear_fit.tax <- lm(crim ~ tax)
summary(linear_fit.tax)
```

```
##
## Call:
## lm(formula = crim ~ tax)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.513  -2.738  -0.194   1.065  77.696
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -8.528369   0.815809  -10.45   <2e-16 ***
## tax          0.029742   0.001847   16.10   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.997 on 504 degrees of freedom
## Multiple R-squared:  0.3396, Adjusted R-squared:  0.3383
## F-statistic: 259.2 on 1 and 504 DF,  p-value: < 2.2e-16
```

#Between the per capita crime rate and tax, there is a statistically significant association, and this due to the small p-value of 2e-16.

```
linear_fit.ptratio <- lm(crim ~ ptratio)
summary(linear_fit.ptratio)
```

```
##
## Call:
## lm(formula = crim ~ ptratio)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.654 -3.985 -1.912  1.825 83.353
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -17.6469     3.1473  -5.607 3.40e-08 ***
## ptratio       1.1520     0.1694   6.801 2.94e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.24 on 504 degrees of freedom
## Multiple R-squared:  0.08407,    Adjusted R-squared:  0.08225
## F-statistic: 46.26 on 1 and 504 DF,  p-value: 2.943e-11
```

#There a statistically significant association between per capita crime and ptratio because of the model's small p-value of 2.94e-11

```
linear_fit.black <- lm(crim ~ black)
summary(linear_fit.black)

##
## Call:
## lm(formula = crim ~ black)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.756  -2.299  -2.095  -1.296  86.822
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.553529   1.425903  11.609   <2e-16 ***
## black       -0.036280   0.003873  -9.367   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.946 on 504 degrees of freedom
## Multiple R-squared:  0.1483, Adjusted R-squared:  0.1466
## F-statistic: 87.74 on 1 and 504 DF,  p-value: < 2.2e-16
```

#Basing on the low p-value of 2e-16 which is much less than 0.05, we can state that there is a statistically significant association between the per capita crime rate and the black variable.

```
linear_fit.lstat <- lm(crim ~ lstat)
summary(linear_fit.lstat)

##
## Call:
## lm(formula = crim ~ lstat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.925  -2.822  -0.664   1.079  82.862
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -3.33054    0.69376  -4.801 2.09e-06 ***
## lstat        0.54880    0.04776  11.491  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.664 on 504 degrees of freedom
## Multiple R-squared:  0.2076, Adjusted R-squared:  0.206
## F-statistic:   132 on 1 and 504 DF,  p-value: < 2.2e-16
```

#The p-value of 2e-16 is way below the 0.05 and therefore we can conclude that there is a statistically significant association between the per capita crime rate and lstat

```
linear_fit.medv <- lm(crim ~ medv)
summary(linear_fit.medv)

##
## Call:
## lm(formula = crim ~ medv)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.071 -4.022 -2.343  1.298 80.957
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 11.79654    0.93419   12.63   <2e-16 ***
## medv        -0.36316    0.03839   -9.46   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.934 on 504 degrees of freedom
## Multiple R-squared:  0.1508, Adjusted R-squared:  0.1491
## F-statistic: 89.49 on 1 and 504 DF,  p-value: < 2.2e-16
```

#There is a statistically significant association between the per capita crime rate and medv because of the small p-value of 2e-16

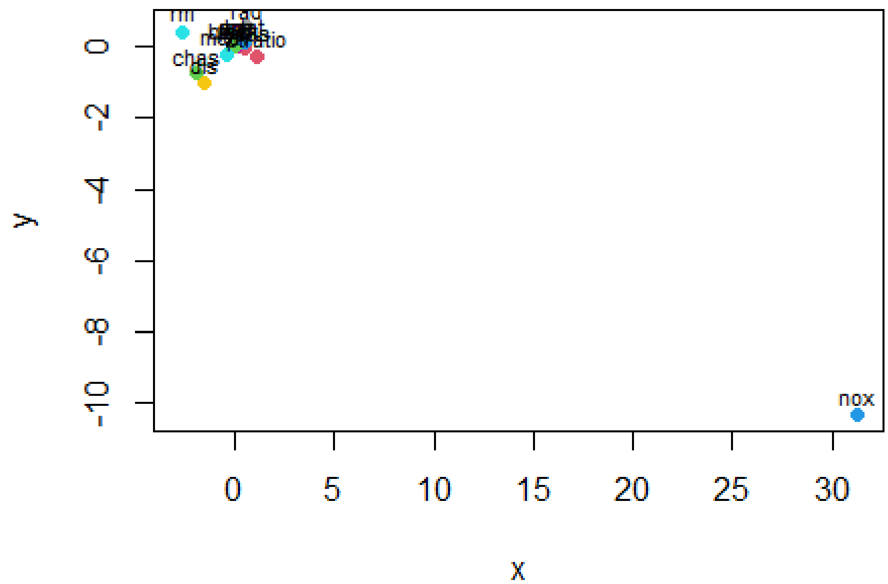
```
multi.lm.fit = lm(crim~., data=Boston)
summary(multi.lm.fit)

##
## Call:
## lm(formula = crim ~ ., data = Boston)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.924 -2.120 -0.353  1.019 75.051
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.033228   7.234903   2.354 0.018949 *
```

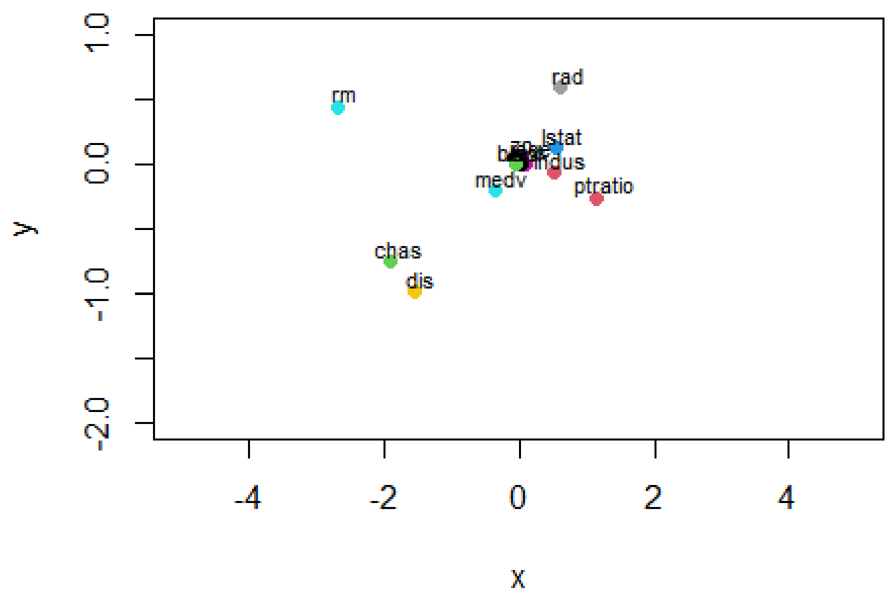
```
## zn            0.044855   0.018734   2.394 0.017025 *
## indus        -0.063855   0.083407  -0.766 0.444294
## chas         -0.749134   1.180147  -0.635 0.525867
## nox         -10.313535   5.275536  -1.955 0.051152 .
## rm            0.430131   0.612830   0.702 0.483089
## age           0.001452   0.017925   0.081 0.935488
## dis          -0.987176   0.281817  -3.503 0.000502 ***
## rad           0.588209   0.088049   6.680 6.46e-11 ***
## tax          -0.003780   0.005156  -0.733 0.463793
## ptratio      -0.271081   0.186450  -1.454 0.146611
## black        -0.007538   0.003673  -2.052 0.040702 *
## lstat         0.126211   0.075725   1.667 0.096208 .
## medv         -0.198887   0.060516  -3.287 0.001087 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.439 on 492 degrees of freedom
## Multiple R-squared:  0.454,  Adjusted R-squared:  0.4396
## F-statistic: 31.47 on 13 and 492 DF,  p-value: < 2.2e-16
```

#A few predictors of the fitted multiple regression model are found to be statistically significant and these include ” zn”, “dis”, “rad”, “black”, and “medv”. dis and rad at the 0.001 level, medv at 0.01 level and zn and black at 0.05 level. Other remaining variables because of their high p-values, we cannot reject the null hypothesis (H0: βj = 0). In conclusion, we can only reject the null hypothesis for “zn”, ”dis”, ”rad”, ”black” and “medv”

```
result <- apply(Boston[, -1], 2, FUN = function(x){
  lm(crim ~ x, data = Boston)
})
x <- c(result$zn$coefficients[2],
       result$indus$coefficients[2],
       result$chas$coefficients[2],
       result$nox$coefficients[2],
       result$rm$coefficients[2],
       result$age$coefficients[2],
       result$dis$coefficients[2],
       result$rad$coefficients[2],
       result$tax$coefficients[2],
       result$ptratio$coefficients[2],
       result$black$coefficients[2],
       result$lstat$coefficients[2],
       result$medv$coefficients[2]
       )
y <- multi.lm.fit$coefficients[-1]
plot(x, y, pch = 16, col=c(1:13))
text(x, y+0.5, labels = colnames(Boston)[-1], cex = 0.7)
```




```r
plot(x, y, pch = 16, col = c(1:13),ylim = c(-2, 1), xlim =c(-5, 5))
text(x+0.1, y+0.1, labels = colnames(Boston)[-1], cex = 0.7)
```



```r
poly_zn <- lm(crim~ zn + I(zn^2) +I(zn^3))
summary(poly_zn)
```

```
##
## Call:
## lm(formula = crim ~ zn + I(zn^2) + I(zn^3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.821 -4.614 -1.294  0.473 84.130
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.846e+00  4.330e-01  11.192  < 2e-16 ***
## zn          -3.322e-01  1.098e-01  -3.025  0.00261 **
## I(zn^2)      6.483e-03  3.861e-03   1.679  0.09375 .
## I(zn^3)     -3.776e-05  3.139e-05  -1.203  0.22954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.372 on 502 degrees of freedom
## Multiple R-squared:  0.05824,    Adjusted R-squared:  0.05261
## F-statistic: 10.35 on 3 and 502 DF,  p-value: 1.281e-06
```

```r
poly_indus <- lm(crim~ indus + I(indus^2) +I(indus^3))
summary(poly_indus)
```

```
##
## Call:
## lm(formula = crim ~ indus + I(indus^2) + I(indus^3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -8.278 -2.514  0.054  0.764 79.713
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.6625683  1.5739833   2.327   0.0204 *
## indus       -1.9652129  0.4819901  -4.077 5.30e-05 ***
## I(indus^2)   0.2519373  0.0393221   6.407 3.42e-10 ***
## I(indus^3)  -0.0069760  0.0009567  -7.292 1.20e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.423 on 502 degrees of freedom
## Multiple R-squared:  0.2597, Adjusted R-squared:  0.2552
## F-statistic: 58.69 on 3 and 502 DF,  p-value: < 2.2e-16
```

```r
poly_age <- lm(crim~ age + I(age^2) +I(age^3))
summary(poly_age)
```

```
##
## Call:
## lm(formula = crim ~ age + I(age^2) + I(age^3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.762 -2.673 -0.516  0.019 82.842
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.549e+00  2.769e+00  -0.920  0.35780
## age          2.737e-01  1.864e-01   1.468  0.14266
## I(age^2)    -7.230e-03  3.637e-03  -1.988  0.04738 *
## I(age^3)     5.745e-05  2.109e-05   2.724  0.00668 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.84 on 502 degrees of freedom
## Multiple R-squared:  0.1742, Adjusted R-squared:  0.1693
## F-statistic: 35.31 on 3 and 502 DF,  p-value: < 2.2e-16
```

```
poly_dis <- lm(crim~ dis + I(dis^2) +I(dis^3))
summary(poly_dis)
```

```
##
## Call:
## lm(formula = crim ~ dis + I(dis^2) + I(dis^3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.757  -2.588   0.031   1.267  76.378
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  30.0476     2.4459  12.285  < 2e-16 ***
## dis         -15.5543     1.7360  -8.960  < 2e-16 ***
## I(dis^2)      2.4521     0.3464   7.078 4.94e-12 ***
## I(dis^3)     -0.1186     0.0204  -5.814 1.09e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.331 on 502 degrees of freedom
## Multiple R-squared:  0.2778, Adjusted R-squared:  0.2735
## F-statistic: 64.37 on 3 and 502 DF,  p-value: < 2.2e-16
```

```
poly_rad <- lm(crim~ rad + I(rad^2) +I(rad^3))
summary(poly_rad)
```

```
##
## Call:
## lm(formula = crim ~ rad + I(rad^2) + I(rad^3))
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.381  -0.412  -0.269   0.179  76.217
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.605545   2.050108  -0.295    0.768
## rad          0.512736   1.043597   0.491    0.623
## I(rad^2)    -0.075177   0.148543  -0.506    0.613
## I(rad^3)     0.003209   0.004564   0.703    0.482
##
## Residual standard error: 6.682 on 502 degrees of freedom
## Multiple R-squared:    0.4,  Adjusted R-squared:  0.3965
## F-statistic: 111.6 on 3 and 502 DF,  p-value: < 2.2e-16
```

```
poly_tax <- lm(crim~ tax + I(tax^2) +I(tax^3))
summary(poly_tax)
```

```
##
## Call:
## lm(formula = crim ~ tax + I(tax^2) + I(tax^3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.273  -1.389   0.046   0.536  76.950
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.918e+01  1.180e+01   1.626    0.105
## tax         -1.533e-01  9.568e-02  -1.602    0.110
## I(tax^2)     3.608e-04  2.425e-04   1.488    0.137
## I(tax^3)    -2.204e-07  1.889e-07  -1.167    0.244
##
## Residual standard error: 6.854 on 502 degrees of freedom
## Multiple R-squared:  0.3689, Adjusted R-squared:  0.3651
## F-statistic:  97.8 on 3 and 502 DF,  p-value: < 2.2e-16
```

```
poly_ptratio <- lm(crim~ ptratio + I(ptratio^2) +I(ptratio^3))
summary(poly_ptratio)
```

```
##
## Call:
## lm(formula = crim ~ ptratio + I(ptratio^2) + I(ptratio^3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.833 -4.146 -1.655  1.408 82.697
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  477.18405  156.79498   3.043  0.00246 **
## ptratio      -82.36054   27.64394  -2.979  0.00303 **
## I(ptratio^2)   4.63535    1.60832   2.882  0.00412 **
## I(ptratio^3)  -0.08476    0.03090  -2.743  0.00630 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.122 on 502 degrees of freedom
## Multiple R-squared:  0.1138, Adjusted R-squared:  0.1085
## F-statistic: 21.48 on 3 and 502 DF,  p-value: 4.171e-13
```

```
poly_black <- lm(crim~ black + I(black^2) +I(black^3))
summary(poly_black)
```

```
##
## Call:
## lm(formula = crim ~ black + I(black^2) + I(black^3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.096  -2.343  -2.128  -1.439  86.790
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.826e+01  2.305e+00   7.924  1.5e-14 ***
## black       -8.356e-02  5.633e-02  -1.483    0.139
## I(black^2)   2.137e-04  2.984e-04   0.716    0.474
## I(black^3)  -2.652e-07  4.364e-07  -0.608    0.544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.955 on 502 degrees of freedom
## Multiple R-squared:  0.1498, Adjusted R-squared:  0.1448
## F-statistic: 29.49 on 3 and 502 DF,  p-value: < 2.2e-16
```

```
poly_lstat <- lm(crim~ lstat + I(lstat^2) +I(lstat^3))
summary(poly_lstat)
```

```
##
## Call:
## lm(formula = crim ~ lstat + I(lstat^2) + I(lstat^3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -15.234  -2.151  -0.486   0.066  83.353
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.2009656  2.0286452   0.592   0.5541
## lstat       -0.4490656  0.4648911  -0.966   0.3345
## I(lstat^2)   0.0557794  0.0301156   1.852   0.0646 .
```

```
## I(lstat^3)  -0.0008574  0.0005652  -1.517   0.1299
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.629 on 502 degrees of freedom
## Multiple R-squared:  0.2179, Adjusted R-squared:  0.2133
## F-statistic: 46.63 on 3 and 502 DF,  p-value: < 2.2e-16
```

```
poly_medv <- lm(crim~ medv + I(medv^2) +I(medv^3))
summary(poly_medv)
```

```
##
## Call:
## lm(formula = crim ~ medv + I(medv^2) + I(medv^3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.427  -1.976  -0.437   0.439  73.655
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 53.1655381  3.3563105  15.840  < 2e-16 ***
## medv        -5.0948305  0.4338321 -11.744  < 2e-16 ***
## I(medv^2)    0.1554965  0.0171904   9.046  < 2e-16 ***
## I(medv^3)   -0.0014901  0.0002038  -7.312 1.05e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.569 on 502 degrees of freedom
## Multiple R-squared:  0.4202, Adjusted R-squared:  0.4167
## F-statistic: 121.3 on 3 and 502 DF,  p-value: < 2.2e-16
```

I**2   I**3

Zn   NS   NS

Indus S   S

#Nox   S   S

#rm   NS   NS

#age   S   S

#dis   S   S

#rad   NS   NS

#tax   NS   NS

#ptratio S   S

#black   NS   NS

#Istat   NS   NS

#medv   S   S

#chas   NA   NA