Linear Models

Prof. Sharon McNicholas

STATS 780/CSE 780

Introduction

- Before going any further, please note that there are entire courses devoted to linear models.
- Today I want to give you a flavour — there is a limit to what I can cover in one "lecture".
- I recommend Faraway (2014)[a] for further reading.

[a] Faraway, J.J. (2014). Linear Models with R. 2nd edn. Boca Raton: Chapman & Hall/CRC Press.

Basic Model Description

- Consider $(\mathbf{X}_1, Y_1), (\mathbf{X}_2, Y_2), \ldots, (\mathbf{X}_n, Y_n)$.

- In simple terms, we can think of a linear model with J predictors as

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_J X_{iJ} + \varepsilon_i = \beta_0 + \sum_{j=1}^{J} \beta_j X_{ij} + \varepsilon_i,$$

  for $i = 1, \ldots, n$, where the $\beta_j \in \mathbb{R}$ are (unknown) parameters and $\varepsilon_i \overset{\text{i.i.d.}}{\sim} \mathsf{N}(0, \sigma^2)$.

- Note that the predictors can be transformed, giving tremendous flexibility.

Recall: Fathers & Sons

- Height for 1,078 fathers and sons (in inches).

- Very famous example used by Pearson.

- We looked at some pictures, and the first one is just a (simple) linear model:

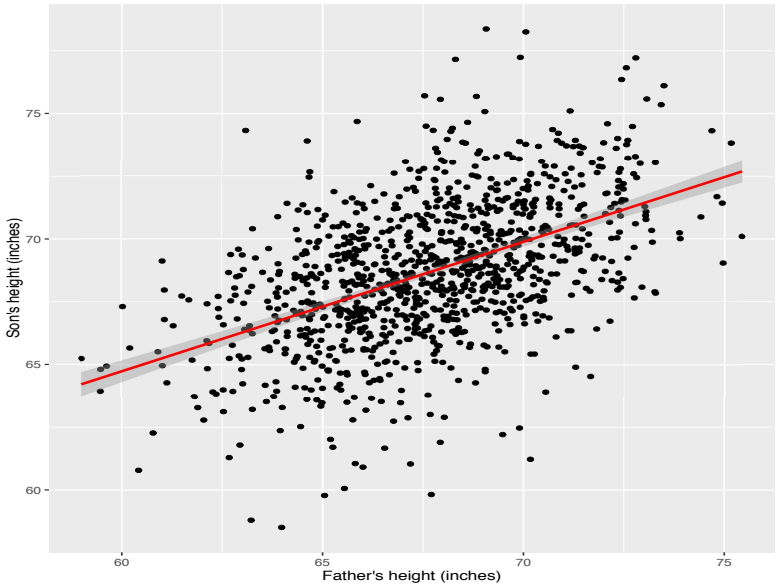
$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,$$

  where Y_i is the height of the ith son and X_i is the height of the ith father.

- The plotted (red) line is defined by

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X.$$

A Linear Model



Output from R

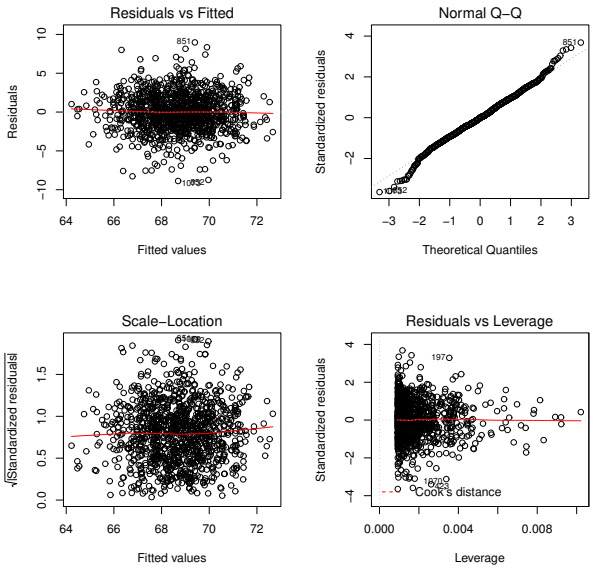
```
>data(father.son, package="UsingR")
>model<-lm(sheight~fheight, father.son)
>summary(model)
Call:
lm(formula = sheight ~ fheight, data = father.son)

Residuals:
    Min      1Q  Median      3Q     Max
-8.8772 -1.5144 -0.0079  1.6285  8.9685

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 33.88660    1.83235   18.49   <2e-16 ***
fheight      0.51409    0.02705   19.01   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.437 on 1076 degrees of freedom
Multiple R-squared:  0.2513,Adjusted R-squared:  0.2506
F-statistic: 361.2 on 1 and 1076 DF,  p-value: < 2.2e-16
```

Residual Plots



Comments

- What can we actually learn from this model?
- Now, we will turn to R and look at some data sets.
- Learning in this way is, I think, effective in this environment.
- However, I strongly recommend that you do some background reading.
- I suggest using Faraway (2014) or the first edition thereof.