Week1

Li-Hsin Chien

2026-02-26

1. Data

```
library(faraway)
```

```
## Warning: 套件 'faraway' 是用 R 版本 4.4.3 來建造的
```

```
#https://cran.r-project.org/web/packages/faraway/refman/faraway.html

data(gala, package="faraway")
#?gala
head(gala[,-2])
```

```
##              Species  Area Elevation Nearest Scruz Adjacent
## Baltra            58 25.09       346     0.6   0.6     1.84
## Bartolome         31  1.24       109     0.6  26.3   572.33
## Caldwell           3  0.21       114     2.8  58.7     0.78
## Champion          25  0.10        46     1.9  47.4     0.18
## Coamano            2  0.05        77     1.9   1.9   903.82
## Daphne.Major      18  0.34       119     8.0   8.0     1.84
```

2. Find OLS estimate:

2.1. R commend *lm*

```
lmod <- lm(Species ~ Area + Elevation + Nearest + Scruz  + Adjacent, data=gala)
summary(lmod)
```

```
##
## Call:
## lm(formula = Species ~ Area + Elevation + Nearest + Scruz + Adjacent,
##     data = gala)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -111.679  -34.898   -7.862   33.460  182.584
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.068221  19.154198   0.369 0.715351
## Area        -0.023938   0.022422  -1.068 0.296318
## Elevation    0.319465   0.053663   5.953 3.82e-06 ***
## Nearest      0.009144   1.054136   0.009 0.993151
## Scruz       -0.240524   0.215402  -1.117 0.275208
## Adjacent    -0.074805   0.017700  -4.226 0.000297 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 60.98 on 24 degrees of freedom
## Multiple R-squared:  0.7658, Adjusted R-squared:  0.7171
## F-statistic:  15.7 on 5 and 24 DF,  p-value: 6.838e-07
```

2.2. Matrix Operations

Suppose we have a response variable y_i for $i = 1, \ldots, n$. This response may be related to a set of $p-1$ predictors $(x_{i1}, \ldots, x_{i(p-1)})$.

Following the linear regression model approach:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{p-1} x_{i(p-1)} + \varepsilon_i, \quad i = 1, \ldots n$$

Let

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{p \times 1}, X = \begin{pmatrix} 1 & x_{11} & \ldots & x_{1(p-1)} \\ \vdots & & & \\ 1 & x_{n1} & \ldots & x_{n(p-1)} \end{pmatrix}_{n \times p}$$

Then the OLS estimates:

$$\hat{\beta}_{\text{OLS}} = (X'X)^{-1} X'Y$$

```
y<-gala$Species

x<-as.matrix(cbind(1,gala[,-c(1,2)]))
#head(x)
#x<-model.matrix(lmod)

beta.hat<-solve(t(x)%*%x)%*%t(x)%*%y

beta.hat
```

```
##                  [,1]
## 1          7.068220709
## Area      -0.023938338
## Elevation  0.319464761
## Nearest    0.009143961
## Scruz     -0.240524230
## Adjacent  -0.074804832
```

```
lmod$coefficients
```

```
##  (Intercept)         Area    Elevation      Nearest        Scruz     Adjacent
##  7.068220709 -0.023938338  0.319464761  0.009143961 -0.240524230 -0.074804832
```