Multiple Linear Regression

Dr. Kiah Wah Ong

ILLINOIS TECH

The Least Squares Method

From

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i, \quad i = 1, \cdots, n$$

The least square function is

$$S(\beta_0, \beta_1, \cdots, \beta_k) = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2$$

We want to minimize S with respect to $\beta_0, \beta_1, \cdots, \beta_k$.

The Least Squares Method

To do that, the least squares estimators of $\beta_0, \beta_1, \cdots, \beta_k$ must satisfy

$$\left.\frac{\partial S}{\partial \beta_0}\right|_{\hat{\beta}_0, \hat{\beta}_1, \cdots \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) = 0$$

and

$$\left.\frac{\partial S}{\partial \beta_j}\right|_{\hat{\beta}_0, \hat{\beta}_1, \cdots \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) x_{ij} = 0 \;\; j = 1, 2, \cdots, k$$

The Least Squares Method

Simplify to obtain the **least-squares normal equations**

$$\begin{aligned}
\sum_{i=1}^{n} y_i &= n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} \\
\sum_{i=1}^{n} x_{i1} y_i &= \hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{i1} x_{ik} \\
\vdots &= \vdots \\
\sum_{i=1}^{n} x_{ik} y_i &= \hat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_1 \sum_{i=1}^{n} x_{ik} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ik} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2
\end{aligned}$$

Multiple Linear Regression Model in Matrix Form

Recall if we write

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

then

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i, \quad i = 1, \cdots, n$$

can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

The Least Squares Method

The normal equations in matrix form can then be written as

$$\mathbf{X}^T\mathbf{X}\hat{\beta} = \mathbf{X}^T\mathbf{y}$$

where

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum_i x_{i1} & \sum_i x_{i2} & \cdots & \sum_i x_{ik} \\ \sum_i x_{i1} & \sum_i x_{i1}^2 & \sum_i x_{i1}x_{i2} & \cdots & \sum_i x_{i1}x_{ik} \\ \vdots & \vdots & & \vdots & \\ \sum_i x_{ik} & \sum_i x_{ik}x_{i1} & \sum_i x_{ik}x_{i2} & \cdots & \sum_i x_{ik}^2 \end{bmatrix}$$

The Least Squares Method

with

$$\mathbf{X}^T\mathbf{y} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_{i1} y_i \\ \vdots \\ \sum_i x_{ik} y_i \end{bmatrix}$$

It is now easy to see that the matrix equation

$$\mathbf{X}^T\mathbf{X}\hat{\beta} = \mathbf{X}^T\mathbf{y}$$

is equivalent to the set of normal equations we derived.

The Least Squares Method

Assuming that $\mathbf{X}^T\mathbf{X}$ is invertible, then we can write

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

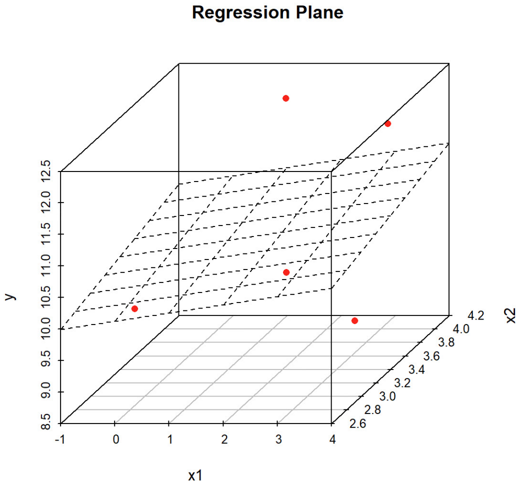
The Least Squares Method

Example

	x1	x2	y
1	1.02189053	4.173966	11.971058
2	3.65738557	3.618790	12.167818
3	-0.02483153	2.887266	10.008135
4	2.64307838	3.917028	8.719754
5	2.93230765	2.776741	10.708046

The Least Squares Method

Example



The Least Squares Method

	x1	x2	y
1	1.02189053	4.173966	11.971058
2	3.65738557	3.618790	12.167818
3	-0.02483153	2.887266	10.008135
4	2.64307838	3.917028	8.719754
5	2.93230765	2.776741	10.708046

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i, \quad i = 1, 2, 3, 4, 5$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \\ 1 & x_{51} & x_{52} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

The Least Squares Method in R

```
#Setting up the matrices y and X
y_tilde=matrix(y, 5,1)
column1=rep(c(1), each=5)
column2=x1
column3=x2
X=cbind(column1, column2, column3)
```

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \\ 1 & x_{51} & x_{52} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

The Least Squares Method in R

```
#Compute the transpose of matrix X
X_t=t(X)
#multiply X transpose with X
X_t.X=X_t %*% X
#find the inverse of the matrix (X transpose times X)
X_t.X.inv=solve(X_t.X)
#then multiply the matrix above with X transpose
X_t.X.inv.X_t=X_t.X.inv %*% X_t
#finally multiply (X^TX)^{-1}X with the vector y to obtain the vector beta
b=X_t.X.inv.X_t %*% y_tilde
```

	V1
column1	9.1472743
column2	0.1299502
column3	0.3746489

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

The Least Squares Method in R

```
model=lm(y~x1+x2)
show(model)
```

```
Call:
lm(formula = y ~ x1 + x2)

Coefficients:
(Intercept)           x1           x2
     9.1473       0.1300       0.3746
```

Hence the least square fit is

$$\hat{y} = 9.1473 + 0.1300x_1 + 0.3746x_2$$