Chapter 1 Simple Linear Regression (part 6: matrix version)

1 Overview

- Simple linear regression model: response variable Y, a single independent variable X

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

- Multiple linear regression model: response Y, more than one independent variables $X_1, X_2, ..., X_p$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p + \varepsilon$$

- To investigate multiple regression model, we need matrix

2 Review of Matrices

- A matrix: a rectangular array of elements arranged in rows and columns
- an example

$$\begin{array}{c} \\ \\ \text{Row 1} \\ \text{Row 2} \\ \text{Row 3} \end{array} \begin{array}{c} \begin{array}{cc} \text{Column} & \text{Column} \\ 1 & 2 \end{array} \\ \left[\begin{array}{cc} 100 & 22 \\ 300 & 46 \\ 600 & 81 \end{array}\right] \end{array}.$$

2.1 A matrix with r rows and c columns

- $$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2c} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{ic} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rj} & \cdots & a_{rc} \end{bmatrix}.$$

- Sometimes denote it as $\mathbf{A} = [a_{ij}] \quad i = 1, ..., r; \quad j = 1, ..., c$

- r and c are called the dimension of a matrix

2.2 Square matrix and Vector

- Square matrix: equal number of rows and columns

$$\begin{bmatrix} 4 & 7 \\ 3 & 9 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

- vector: matrix with only one row or one column

$$\mathbf{A} = [4 \quad 7 \quad 10] \qquad \mathbf{B} = \begin{bmatrix} 15 \\ 25 \\ 20 \end{bmatrix}.$$

2.3 Transpose of a matrix and equality of matrices

- transpose of a matrix $\mathbf{A}$ is another matrix denoted by $\mathbf{A}'$

$$\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 7 & 10 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} 2 & 7 & 3 \\ 5 & 10 & 4 \end{bmatrix}.$$

- two matrices are equal if they have the same dimension and all the corresponding elements are equal

  Suppose

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 17 & 2 \\ 14 & 5 \\ 13 & 9 \end{bmatrix}.$$

  If $\mathbf{A} = \mathbf{B}$, then $a_{11} = 17, a_{12} = 2, ...$

2.4 Matrix addition and subtraction

- Adding or subtracting of two matrices require that they have the same dimension.

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix},$$

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1+1 & 4+2 \\ 2+2 & 5+3 \\ 3+3 & 6+4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \\ 6 & 10 \end{bmatrix},$$

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} 1-1 & 4-2 \\ 2-2 & 5-3 \\ 3-3 & 6-4 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \end{bmatrix}$$

2.5 Matrix multiplication

- Multiplication of a Matrix by a Scalar

$$\mathbf{A} = \begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix},$$

$$4\mathbf{A} = \mathbf{A}4 = 4\begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 28 \\ 36 & 12 \end{bmatrix}$$

- Multiplication of a Matrix by a Matrix. If $\mathbf{A}$ has dimension $r \times c$ and $\mathbf{B}$ has dimension $c \times s$, the product $\mathbf{AB}$ is a matrix of dimension $r \times s$ with the element in the ith row and jth column:

$$\sum_{k=1}^{c} a_{ik} b_{kj}$$

- $$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 5 & 8 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 4a_1 + 2a_2 \\ 5a_1 + 8a_2 \end{bmatrix}$$

2.6 Regression examples

- It is easy to check

$$\begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_1 \\ \beta_0 + \beta_1 X_2 \\ \vdots \\ \beta_0 + \beta_1 X_n \end{bmatrix}$$

- Let

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

- The regression model

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_1 + \varepsilon_1, \\ Y_2 &= \beta_0 + \beta_1 X_2 + \varepsilon_2, \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_n + \varepsilon_n \end{aligned}$$

can be written as

$$\mathbf{Y} = \mathbf{X}\beta + \mathcal{E}$$

- Other calculations

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix}' \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{bmatrix}$$

$$\mathbf{Y}'\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \sum_{i=1}^n Y_i^2$$

2.7 Special types of matrices

- Symmetric Matrix $\mathbf{A} = \mathbf{A}'$

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{bmatrix}$$

- Diagonal Matrix: a square matrix whose off-diagonal elements are all zeros

$$\begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix}$$

- Identity Matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

facts: for any appropriate matrix $\mathbf{AI} = \mathbf{A}$ and $\mathbf{IB} = \mathbf{B}$

- zero vector and unit vector

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

2.8 Inverse of a square matrix

- the inverse of a square matrix A is another square matrix, denoted by $\mathbf{A^{-1}}$, such that $\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}$$

Since

$$\begin{bmatrix} -0.1 & 0.4 \\ 0.3 & -0.2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

or

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -0.1 & 0.4 \\ 0.3 & -0.2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

So

$$A^{-1} = \begin{bmatrix} -0.1 & 0.4 \\ 0.3 & -0.2 \end{bmatrix}$$

2.9 Finding the Inverse of a matrix

- If

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then

$$\mathbf{A^{-1}} = \begin{bmatrix} \frac{d}{D} & \frac{-b}{D} \\ \frac{-c}{D} & \frac{a}{D} \end{bmatrix}$$

where $D = ad - bc$

- For high dimensional matrix, its inverse is not easy to calculate by hand

2.10 Regression example (continue)

- the inverse of matrix

$$\mathbf{X'X} = \begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix}$$

$D = n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2 = n[\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}] = n\sum_{i=1}^n (X_i - \bar{X})^2$

So

$$\begin{aligned} (\mathbf{X'X})^{-1} &= \begin{bmatrix} \frac{\sum_{i=1}^n X_i^2}{n\sum_{i=1}^n (X_i-\bar{X})^2} & \frac{-\sum_{i=1}^n X_i}{n\sum_{i=1}^n (X_i-\bar{X})^2} \\ \frac{-\sum_{i=1}^n X_i}{n\sum_{i=1}^n (X_i-\bar{X})^2} & \frac{n}{n\sum_{i=1}^n (X_i-\bar{X})^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2} & \frac{-\bar{X}}{\sum_{i=1}^n (X_i-\bar{X})^2} \\ \frac{-\bar{X}}{\sum_{i=1}^n (X_i-\bar{X})^2} & \frac{1}{\sum_{i=1}^n (X_i-\bar{X})^2} \end{bmatrix} \end{aligned}$$

2.11 Use of Inverse Matrix

- Suppose we want to solve two equations:

$$\begin{aligned} 2y_1 + 4y_2 &= 20 \\ 3y_1 + y_2 &= 10 \end{aligned}$$

  Rewrite the equations in matrix notation:

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \end{bmatrix}$$

  So the solution to the equations

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 20 \\ 10 \end{bmatrix}$$

$$= \begin{bmatrix} -0.1 & 0.4 \\ 0.3 & -0.2 \end{bmatrix} \begin{bmatrix} 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

- Estimation a regression model need to solve linear equations, and inverse matri is very useful.

2.12 Other basic facts for matrices

- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- $\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$
- $(\mathbf{A}')' = \mathbf{A}$
- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$
- $(\mathbf{A^{-1}})^{-1} = \mathbf{A}$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$

3 Random vector and matrices

- Random vector

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

- Expectation of random vector

$$\mathbf{E}\{\mathbf{Y}\} = \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ E\{Y_3\} \end{bmatrix}$$

- Random vector

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix}$$

Then

$$E(\mathbf{Y} + \mathbf{Z}) = \mathbf{EY} + \mathbf{EZ}$$

- Variance-covariance Matrix of random vector

$$\mathbf{Var}\{\mathbf{Y}\} = \mathbf{E}\{[\mathbf{Y} - \mathbf{E}\{\mathbf{Y}\}][\mathbf{Y} - \mathbf{E}\{\mathbf{Y}\}]'\}$$

$$= \begin{bmatrix} Var\{Y_1\} & Cov\{Y_1, Y_2\} & Cov\{Y_1, Y_3\} \\ Cov\{Y_2, Y_1\} & Var\{Y_2\} & Cov\{Y_2, Y_3\} \\ Cov\{Y_3, Y_1\} & Cov\{Y_3, Y_2\} & Var\{Y_3\} \end{bmatrix}$$

- In simple linear regression model, errors are uncorrelated, so $Var\{\mathcal{E}\} = \sigma^2 I$

[Proof: for example consider $n = 3$.

$$\mathbf{Var}\{\mathcal{E}\} = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$$

]

3.1 Some basic facts

- If we have a random vector $\mathbf{W}$ equal to a random vector $\mathbf{Y}$ multiplied by a constant matrix $\mathbf{A}$

$$\mathbf{W} = \mathbf{AY}$$

we have

$$\mathbf{E}\{\mathbf{W}\} = \mathbf{AE}\{\mathbf{Y}\}$$

$$\mathbf{Var}\{\mathbf{W}\} = \mathbf{Var}\{\mathbf{AY}\} = \mathbf{AVar}\{\mathbf{Y}\}\mathbf{A}'$$

- If c is a constant vector, then

$$E(\mathbf{c} + \mathbf{AY}) = \mathbf{c} + \mathbf{AEY}$$

and

$$Var(\mathbf{c} + \mathbf{AY}) = \mathbf{Var}(\mathbf{AY}) = \mathbf{A}\ \mathbf{Var}\{\mathbf{Y}\}\mathbf{A}'$$

- In simple linear regression model, it follows from above $Var\{\mathbf{Y}\} = \sigma^2 I$

3.2 An illustration

- $$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} Y_1 - Y_2 \\ Y_1 + Y_2 \end{bmatrix}$$
  $$\mathbf{E} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \end{bmatrix} = \begin{bmatrix} E\{Y_1\} - E\{Y_2\} \\ E\{Y_1\} + E\{Y_2\} \end{bmatrix}$$

- $$Var\{\begin{bmatrix} W_1 \\ W_2 \end{bmatrix}\} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} Var\{Y_1\} & Cov\{Y_1, Y_2\} \\ Cov\{Y_2, Y_1\} & Var\{Y_2\} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

4 Simple linear regression model (matrix version)

The model

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_1 + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 X_2 + \varepsilon_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_n + \varepsilon_n \end{aligned}$$

with assumption

1. $E(\varepsilon_i) = 0$,
2. $Var(\varepsilon_i) = \sigma^2, Cov(\varepsilon_i, \varepsilon_j) = 0$ for all $1 \le i \ne j \le n$.
3. $\varepsilon_i \sim N(0, \sigma^2), i = 1, ..., n$ are independent

Recall, the model can be written as

$$\mathbf{Y} = \mathbf{X}\beta + \mathcal{E}$$

Note that

$$\mathbf{E}\{\mathcal{E}\} = \mathbf{0}, \qquad \mathbf{Var}\{\mathcal{E}\} = \begin{bmatrix} \sigma^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^{\mathbf{2}}\mathbf{I}$$

The assumptions can be rewritten as

1. $E(\mathcal{E}) = 0$,
2. $Var(\mathcal{E}) = \sigma^2 I$
3. $\mathcal{E} \sim N(0, \sigma^2 I)$

Thus $E(\mathbf{Y}) = \mathbf{X}\beta$ and $Var(\mathbf{Y}) = \sigma^{\mathbf{2}}\mathbf{I}$. The model (with assumptions 1, 2, and 3.) can also be written as

$$\mathbf{Y} \sim \mathbf{N}(\mathbf{X}\beta, \sigma^{\mathbf{2}}\mathbf{I})$$

or

$$\mathbf{Y} = \mathbf{X}\beta + \mathcal{E}, \qquad \mathcal{E} \sim \mathbf{N}(\mathbf{0}, \sigma^{\mathbf{2}}\mathbf{I})$$

4.1 Least squares estimator b_0, b_1

- The normal equations can be written as

$$\begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} nb_0 + b_1 \sum_{i=1}^n X_i \\ b_0 \sum_{i=1}^n X_i + b_1 \sum_{i=1}^n X_i^2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{bmatrix}$$

- $$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{bmatrix}$$

- let

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

  Then the normal equation is

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{Y}$$

- we can find b by

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}$$

4.2 An example

- $$\mathbf{Y} = \begin{bmatrix} 16 \\ 5 \\ 10 \\ 15 \\ 13 \\ 22 \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} 1 & 4 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

- $$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 6 & 17 \\ 17 & 55 \end{bmatrix}, \mathbf{X}'\mathbf{Y} = \begin{bmatrix} 81 \\ 261 \end{bmatrix}$$

- $$\mathbf{b} = \begin{bmatrix} 6 & 17 \\ 17 & 55 \end{bmatrix}^{-1} \begin{bmatrix} 81 \\ 261 \end{bmatrix}$$

4.3 Fitted values and residuals in matrix form

- $$\hat{\mathbf{Y}} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix} = \begin{bmatrix} b_0 + b_1 X_1 \\ b_0 + b_1 X_2 \\ \vdots \\ b_0 + b_1 X_n \end{bmatrix} = \mathbf{X}\mathbf{b}$$

- $$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

- $$\hat{\mathbf{Y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}$$

- Denote $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'$ by $\mathbf{H}$, we have

$$\hat{\mathbf{Y}} = \mathbf{H}\mathbf{Y}, \quad \mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{Y}$$

4.4 Variance-covariance matrix for residuals e

- $\mathbf{Var}\{\mathbf{e}\} = \mathbf{Var}\{(\mathbf{I} - \mathbf{H})\mathbf{Y}\} = (\mathbf{I} - \mathbf{H})\mathbf{Var}\{\mathbf{Y}\}(\mathbf{I} - \mathbf{H})'$

- $\mathbf{Var}\{\mathbf{Y}\} = \mathbf{Var}\{\varepsilon\} = \sigma^{\mathbf{2}}\mathbf{I}$

- $(\mathbf{I} - \mathbf{H})' = \mathbf{I}' - \mathbf{H}' = \mathbf{I} - \mathbf{H}$

- $\mathbf{H}\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}' = \mathbf{H}$

- $(\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H}) = \mathbf{I} - \mathbf{2H} + \mathbf{H}\mathbf{H} = \mathbf{I} - \mathbf{H}$

- $\mathbf{Var}\{\mathbf{e}\} = \sigma^{\mathbf{2}}(\mathbf{I} - \mathbf{H})$ estimated by $\hat{\sigma}^{\mathbf{2}}(\mathbf{I} - \mathbf{H})$

4.5 Analysis of variance in matrix form

- Let $\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ then

$$\mathbf{J} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} = \mathbf{1}\mathbf{1}'.$$

- $$SST = \mathbf{Y}'\mathbf{Y} - \frac{\mathbf{1}}{\mathbf{n}}\mathbf{Y}'\mathbf{J}\mathbf{Y} = \mathbf{Y}'(\mathbf{I} - \frac{\mathbf{1}}{\mathbf{n}}\mathbf{J})\mathbf{Y}$$

  [Proof

$$SST = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n} Y_i^2 - \frac{(\sum_{i=1}^{n} Y_i)^2}{n}$$

$$\mathbf{Y}'\mathbf{Y} = \sum_{i=1}^{n} Y_i^2, \quad \mathbf{1}^{\mathbf{1}}\mathbf{Y} = \mathbf{Y}'\mathbf{1} = \sum_{i=1}^{n} Y_i$$

$$(\sum_{i=1}^{n} Y_i)^2 = \mathbf{Y}'\mathbf{1}\mathbf{1}'\mathbf{Y} = \mathbf{Y}'\mathbf{J}\mathbf{Y}$$

  ]

- $$SSE = \sum_{i=1}^{n} e_i^2 = \mathbf{e}'\mathbf{e} = (\mathbf{Y} - \mathbf{X}\mathbf{b})'(\mathbf{Y} - \mathbf{X}\mathbf{b}) = \mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y}$$

  [Proof

$$\begin{aligned} SSE &= (\mathbf{Y} - \mathbf{X}\mathbf{b})'(\mathbf{Y} - \mathbf{X}\mathbf{b}) = \mathbf{Y}'\mathbf{Y} - \mathbf{2}\mathbf{b}'\mathbf{X}'\mathbf{Y} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} \\ &= \mathbf{Y}'\mathbf{Y} - \mathbf{2}\mathbf{b}'\mathbf{X}'\mathbf{Y} + \mathbf{b}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y} \\ &= \mathbf{Y}'\mathbf{Y} - \mathbf{2}\mathbf{b}'\mathbf{X}'\mathbf{Y} + \mathbf{b}'\mathbf{X}'\mathbf{Y} = \mathbf{Y}'\mathbf{Y} - \mathbf{b}'\mathbf{X}'\mathbf{Y} \\ &= \mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y} \end{aligned}$$

  ]

- $$SSR = SST - SSE = \mathbf{Y}'(\mathbf{H} - \frac{\mathbf{1}}{\mathbf{n}}\mathbf{J})\mathbf{Y}$$

4.6 Variance-covariance matrix for b

- $$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

- $$\begin{aligned} \mathbf{Var}\{\mathbf{b}\} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Var}\{\mathbf{Y}\}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^{\mathbf{2}}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^{\mathbf{2}} \begin{bmatrix} \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} & \frac{-\bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ \frac{-\bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} & \frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{bmatrix} \end{aligned}$$

  where σ^2 can be estimated by $\hat{\sigma}^2$ =MSE

4.7 Variance for the predicted value

- $\hat{Y} = b_0 + b_1 X = \begin{bmatrix} 1 & X \end{bmatrix} b$
- $Var\{\hat{Y}\} = \begin{bmatrix} 1 & X \end{bmatrix} Var\{b\} \begin{bmatrix} 1 \\ X \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & X \end{bmatrix} (\mathbf{X}'\mathbf{X})^{-1} \begin{bmatrix} 1 \\ X \end{bmatrix}$