Multiple Linear Regression

Dr. Supaporn Erjongmanee

Department of Computer Engineering Kasetsart University fengspe@ku.ac.th

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering
Slide 1



1

Outline

- Introduction
- Estimating Parameters
- Residual, SSE, SST, R²
- Example

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 2



Multiple Regression Model | Multiple Regression

General model

neral model
$$x_1, x_2, ..., x_n \rightarrow f = ?$$

 $Y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k + \epsilon$

- where
 - X = independent/predictor variables
 - Y = dependent variable
 - k = number of predictors
 - ϵ is normally distributed with $\mu = 0$, var = σ^2
 - When $\sigma^2 \rightarrow 0$, ϵ is close to zero, Y is closer to true regression line
 - When σ^2 is large, ϵ is close to zero, Y is closer to true regression line
 - b_i = rate how y increases according to x_i increases

fengspe@ku.ac.th

Statistics in Computer Engineering



3

Observation Set

- Let x_{ii} = observed jth predictor in ith data set
 - i = 1,2, ..., n, i = 1,2, ..., k
- The data set is composed of n sets:
 - $(x_{11}, x_{12}, x_{13}, ..., x_{1k}, y_1)$ $y_1 = b_0 + b_1 x_{11} + b_2 x_{12} + ... + b_k x_{1k} + \varepsilon$

•
$$(x_{21}, x_{22}, x_{23}, ..., x_{2k}, y_2)$$
 $y_2 = b_0 + b_1 x_{21} + b_2 x_{22} + ... + b_k x_{2k} + \varepsilon$

•
$$(x_{n1}, x_{n2}, x_{n3}, ..., x_{nk}, y_n)$$
 $y_n = b_0 + b_1 x_{n1} + b_2 x_{n2} + ... + b_k x_{nk} + \varepsilon$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 4



Outline

- Introduction
- Estimating Parameters
- Residual, SSE, SST, R²
- Example

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 5



5

Estimating Parameters

$$Y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k + \varepsilon$$

- Parameters to estimate are
 - b₁, b₂, ..., b_k
- Also, apply principle of least squares to minimize sum of errors $[g(\cdot)]$ to find estimated parameters

$$g(b_0, b_1, b_2, b_k) = \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_{i1} + ... + b_k x_{ik})]^2$$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 6



Estimating Parameters (cont.)

$$g(b_0, b_1, b_2, b_k) = \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_{i1} + ... + b_k x_{ik})]^2$$

• Take partial derivatives of $g(\cdot)$ with respect to each b_i and set to zero.

- Help for solving b_i's
 - Use software for help finding values of b_i's
 - Use matrix to help out calculation

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 7



7

Regression with Matrices

$$y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k + \varepsilon$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ & \vdots & & \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} b_0 \\ \vdots \\ b_k \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$y = X\beta + \varepsilon$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ & \vdots & & \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} b_0 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 8



Regression with Matrices (cont.)

$$g(b_0, b_1, b_2, b_k)$$

$$= \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_{i1} + ... + b_k x_{ik})]^2$$

$$= (y - X\beta)'(y - X\beta)$$

$$= ||y - X\beta||^2$$

- ullet Apply principle of least squares to minimize sum of errors (\cdot) to find $oldsymbol{eta}$
- Take partial derivatives of $g(\cdot)$ with respect to each b_i and set to zero.

$$\begin{array}{l} b_0 \sum_{i=1}^n 1 \ + b_1 \sum_{i=1}^n x_{i1} \ + \dots + b_k \sum_{i=1}^n x_{ik} \ = & \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_{i1} + b_1 \sum_{i=1}^n x_{i1} x_{i1} + \dots + b_k \sum_{i=1}^n x_{i1} x_{ik} = & \sum_{i=1}^n x_{i1} y_i \\ \dots \\ b_0 \sum_{i=1}^n x_{ik} + b_1 \sum_{i=1}^n x_{ik} x_{i1} + \dots + b_k \sum_{i=1}^n x_{ik} x_{ik} = & \sum_{i=1}^n x_{ik} y_i \end{array}$$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 9



9

Regression with Matrices (cont.)

• Take partial derivatives of $g(\cdot)$ with respect to each b_i and set to zero.

In a matrix, we can write:

$$\begin{bmatrix} \sum_{i=1}^{n} 1 & \sum_{i=1}^{n} x_{i1} & \dots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1} x_{i1} & \dots & \sum_{i=1}^{n} x_{i1} x_{ik} \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{ik} x_{i1} & \dots & \sum_{i=1}^{n} x_{ik} x_{ik} \end{bmatrix} \begin{bmatrix} b_0 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \vdots \\ \sum_{i=1}^{n} x_{i1} y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik} y_i \end{bmatrix}$$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 10





• In a matrix, we can write:

$$\begin{bmatrix} \sum_{i=1}^{n} 1 & \sum_{i=1}^{n} x_{i1} & \dots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1} x_{i1} & \dots & \sum_{i=1}^{n} x_{i1} x_{ik} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{ik} x_{i1} & \dots & \sum_{i=1}^{n} x_{ik} x_{ik} \end{bmatrix} \begin{bmatrix} b_0 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \vdots \\ \sum_{i=1}^{n} x_{i1} y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik} y_i \end{bmatrix}$$

$$X'X\beta = X'y$$

Note: X' = transpose of X

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1}X'y$$



 $\hat{y} = X\hat{\beta}$

After finding b_0 , b_1 , b_2 , ..., b_k , we can compute SSE, SST, R^2 ,

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 11



11

Example

 Use data of 6 cars, build a model to predict horsepower (hp) using two inputs: engine size (liters) and fuel type

Make	hp	Engine size	Fuel
Ford	132	2.0	Regular
Mazda	167	2.0	Premium
Subaru	170	2.5	Regular
Lexus	204	2.5	Premium
Mitsubishi	230	3.0	Regular
BMW	260	3.0	Premium

- Let
 - x_1 = Engine size
 - x_2 = Fuel (replace 0 for regular, 1 for premium)
 - Y = hp

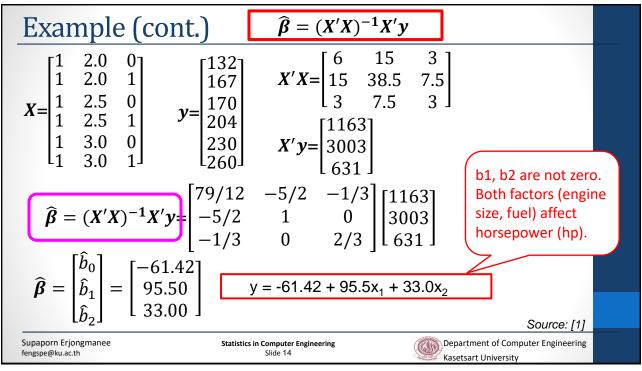
Source: [1]

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 12



Example (cont.)		Make	hp	Engine size	Fuel
		d	132	2.0	Regular
	Maz	da	167	2.0	Premium
• n = 6	Suba	aru	170	2.5	Regular
• k = 2	Lexu	ıs	204	2.5	Premium
· –	Mits	subishi	230	3.0	Regular
	BMV	N	260	3.0	Premium
$ X = \begin{bmatrix} 1 & 2.0 & 0 \\ 1 & 2.0 & 1 \\ 1 & 2.5 & 0 \\ 1 & 2.5 & 1 \\ 1 & 3.0 & 0 \\ 1 & 3.0 & 1 \end{bmatrix} y = \begin{bmatrix} 132 \\ 167 \\ 170 \\ 204 \\ 230 \\ 260 \end{bmatrix} X'X = \begin{bmatrix} 6 & 15 & 3 \\ 15 & 38.5 & 7.5 \\ 3 & 7.5 & 3 \end{bmatrix} X'y = \begin{bmatrix} 1163 \\ 3003 \\ 631 \end{bmatrix} $ Source: [1]					
Supaporn Erjongmanee fengspe@ku.ac.th	Statistics in Computer Engineering Slide 13		(-CAD) 3-)	ent of Computer Eng	ineering



Example (cont.)

 $y = -61.42 + 95.5x_1 + 33.0x_2$

What is SSE? SST? R²?

- Let
 - x_1 = Engine size
 - x_2 = Fuel (replace 0 for regular, 1 for premium)
 - Y = hp
- If fuel has no effect, we increase engine size by one, horsepower is increased by 95.5
- If engine size has no effect, we increase fuel by one (change from regular to premium), horsepower is increased by 33.

Source: [1]

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 15



15

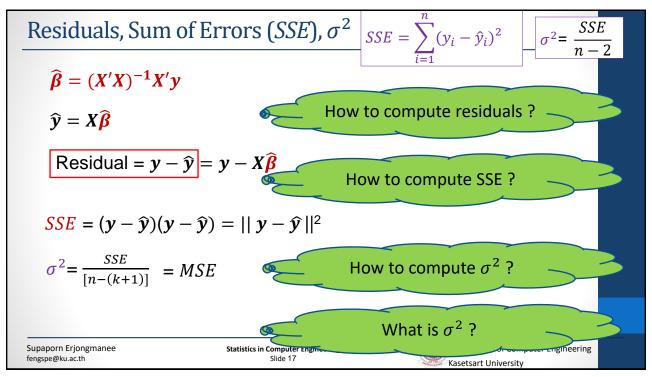
Outline

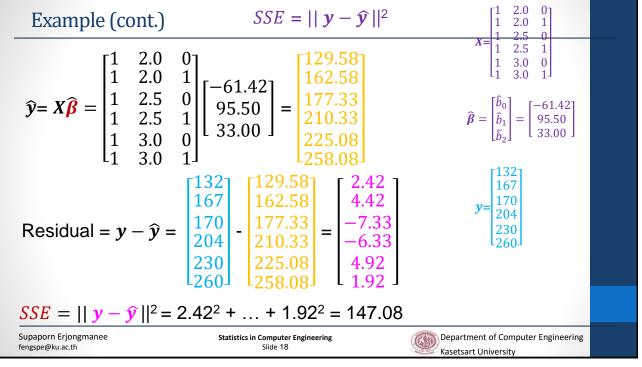
- Introduction
- Estimating Parameters
- Residual, SSE, SST, R²
- Example

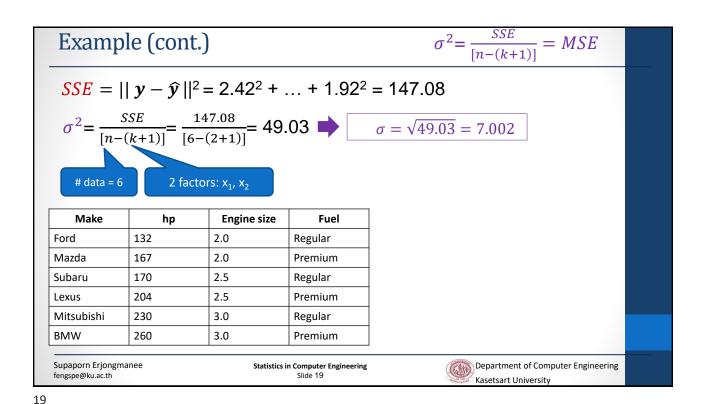
Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 16

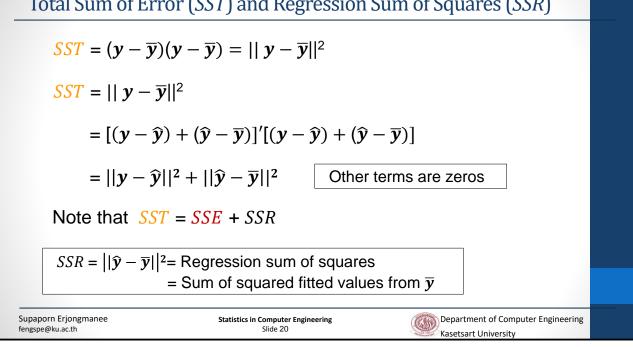








Total Sum of Error (SST) and Regression Sum of Squares (SSR)



Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{SSE}{SST}$$

Proportion of fitted values that can be explained by multiple linear model

Supaporn Erjongmanee

Statistics in Computer Engineering



21

Example (cont.)

$$SST = ||\mathbf{y} - \overline{\mathbf{y}}||^2$$

$$SST = ||\mathbf{y} - \overline{\mathbf{y}}||^2 \qquad R^2 = 1 - \frac{SSE}{SST}$$

$$SST = ||y - \overline{y}||^2 = \sum_{i=1}^{n} (y_i - 193.83)^2 = 10,900.83$$

$$SSE = ||y - \hat{y}||^2 = 2.42^2 + ... + 1.92^2 = 147.08$$

$$SSR = SST - SSE = 10,573.75$$

Make	hp		
Ford	132		
Mazda	167		
Subaru	170		
Lexus	204		
Mitsubishi	230		
BMW	260		

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{147.08}{10900.83} = 0.9865$$

 $\bar{y} = 193.83$

98.65% of fitted values that can be explained by multiple linear model

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 22



Outline

- Introduction
- Estimating Parameters
- Residual, SSE, SST, R²
- Example

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 23

Department of Computer Engineering
Kasetsart University

23

Example 2

- What affects human wingspan? Height? Foot length?
- Use the given data to find the followings:
 - σ, σ²
 - SSE
 - R²
 - SST, SSR
 - MSR, MSE

$$\sigma^2 = \frac{SSE}{[n - (k+1)]}$$

$$R^2 = 1 - \frac{SSE}{SST}$$

SSR = SST - SSE

$$n = 16, k = 2$$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 24



Example 2 (cont.)

- What affects human wingspan? Height? Foot length?
- Use the given data to find the followings:
 - At α = 0.05 , does height affect wingspan?
 - At α = 0.05 , does foot length affect wingspan?

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 25



25

Example 2 (cont.)

- Use the given data to find the followings:
 - At α = 0.05, does <u>height</u> affect wingspan (not consider foot length)?

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 26



Example 2 (cont.)

- Compute correlation between
 - Wingspan and height (r₁)
 - Wingspan and foot length (r₂)
 - Height and foot length (r₃)

 $r = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}$ $= \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 27 Department of Computer Engineering Kasetsart University

27

References

1. J.L. Devore and K.N.Berk, Modern Mathematical Statistics with Applications, Springer, 2012.

Supaporn Erjongmanee fengspe@ku.ac.th

Statistics in Computer Engineering Slide 28 Department of Computer Engineering
Kasetsart University