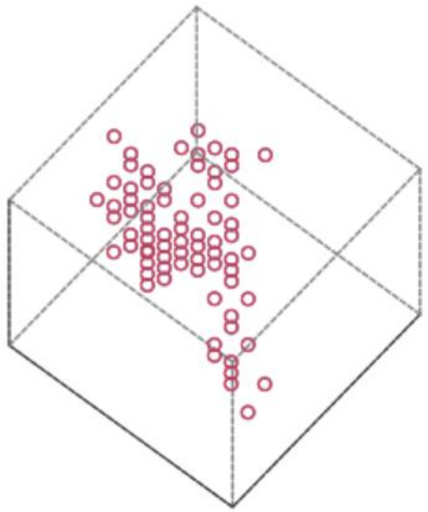
Multiple Linear Regression

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon$$

ผู้ช่วยศาสตราจารย์กัญญารัตน์ บุษบรรณ

Credit : Krin Chinprasatsak

What is Multiple Linear Regression ?



Multiple Linear Regression คือ สมการระนาบ ที่ใช้ตัวนำเข้าเป็นตัวแปรต้น D ตัวเพื่อพยากรณ์ ตัวแปรตาม 1 ตัว

D ตัวแปรต้น	=>	1 ตัวแปรตาม
$x_1, x_2, x_{3,\dots,} x_D$	=>	y

Multiple Linear Regression : How it works

State Data

คือ ตัวแปร(ปัจจัย,คุณลักษณะ)ที่ D ของตัวอย่างที่ 2

$x1$	$x2$	$\cdots$	xD	Y
x_1^1	x_2^1	$\cdots$	x_D^1	y_1
x_1^2	x_2^2	$\cdots$	x_D^2	y_2
x_1^3	x_2^3	$\cdots$	x_D^3	y_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_1^N	x_2^N	$\cdots$	x_D^N	y_N

Data

$$X = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_D^1 \\ x_1^2 & x_2^2 & \cdots & x_D^2 \\ x_1^3 & x_2^3 & \cdots & x_D^3 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^N & x_2^N & \cdots & x_D^N \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix}$$

เมื่อ N คือ จำนวนตัวอย่างทั้งหมด

Multiple Linear Regression

State Data

#ห้องนอน	#ห้องน้ำ	...	ราคาที่ดิน	ราคาขาย
2	1	...	1 M	3 M
5	5	...	2 M	25.3 M
1	1	...	200 K	2 M
3	2	...	1.2 M	12 M

Data

N=4

$$X = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 5 & 5 & \cdots & 2 \\ 1 & 1 & \cdots & 0.2 \\ 3 & 2 & \cdots & 1.2 \end{bmatrix}$$

$$Y = \begin{bmatrix} 3 \\ 25.3 \\ 2 \\ 12 \end{bmatrix}$$

Multiple Linear Regression

State Model

D ตัวแปรต้น => 1 ตัวแปรตาม

$x_1, x_2, x_{3,\dots,} x_D$ => y

Model จะอยู่ในรูปของ $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_D x_D + \varepsilon$

Multiple Linear Regression

State Prediction

$$\hat{y} = w_o + w_1 x_1 + w_2 x_2 + \cdots + w_D x_D$$

$$Yhat = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} w_0 + w_1 x_1^1 + w_2 x_2^1 + \cdots + w_D x_D^1 \\ w_0 + w_1 x_1^2 + w_2 x_2^2 + \cdots + w_D x_D^2 \\ \vdots \\ w_0 + w_1 x_1^N + w_2 x_2^N + \cdots + w_D x_D^N \end{bmatrix}$$

Multiple Linear Regression

State Prediction

$$X_b = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + X$$

$$\text{Yhat} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} w_0 + w_1 x_1^1 + w_2 x_2^1 + \cdots + w_D x_D^1 \\ w_0 + w_1 x_1^2 + w_2 x_2^2 + \cdots + w_D x_D^2 \\ \vdots \\ w_0 + w_1 x_1^N + w_2 x_2^N + \cdots + w_D x_D^N \end{bmatrix}$$

$$X_b = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_D^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_D^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^N & x_2^N & \cdots & x_D^N \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_D^1 \\ x_1^2 & x_2^2 & \cdots & x_D^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^N & x_2^N & \cdots & x_D^N \end{bmatrix}$$

$$\boxed{\text{Yhat} = X_b W}$$

$$\text{Yhat} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_D^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_D^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^N & x_2^N & \cdots & x_D^N \end{bmatrix} \times \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_D \end{bmatrix}$$

Multiple Linear Regression :

How to create model

ต้องการสมการระนาบ $(\hat{y})$ ที่ทำให้ Error น้อยที่สุด

$$\hat{y} = w_o + w_1x_1 + w_2x_2 + \cdots + w_Dx_D$$

ต้องการ $w_o, w_1, w_2, \cdots w_D$ ที่ทำให้ Error น้อยที่สุด

ใช้การหาค่าต่ำสุดจากการหาอนุพันธ์อันดับ 1 ของ Error เทียบกับตัวแปรที่ต้องการ ค่าต่ำสุด ซึ่งจะมีค่าอนุพันธ์เป็น 0

Multiple Linear Regression

How to create model

1. $\frac{\partial Error}{\partial w_0} = 0$

2. $\frac{\partial Error}{\partial w_1} = 0$

$\vdots$

D+1. $\frac{\partial Error}{\partial w_D} = 0$

$$\frac{\partial Error}{\partial w_0} = \frac{\partial \sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\partial w_0} = \sum_{i=1}^{N} \frac{\partial (y_i - \hat{y}_i)^2}{\partial w_0}$$

$$\sum_{i=1}^{N} \hat{y}_i = \sum_{i=1}^{N} y_i$$

$$\frac{\partial Error}{\partial w_1} = 0$$

$$\sum_{i=1}^{N} x_1^i \hat{y}_i = \sum_{i=1}^{N} x_1^i y_i$$

$$\frac{\partial Error}{\partial w_2} = 0$$

$$\sum_{i=1}^{N} x_2^i \hat{y}_i = \sum_{i=1}^{N} x_2^i y_i$$

$$\frac{\partial Error}{\partial w_D} = 0$$

$$\sum_{i=1}^{N} x_D^i \hat{y}_i = \sum_{i=1}^{N} x_D^i y_i$$

Multiple Linear Regression

How to create model

1. $\frac{\partial Error}{\partial w_0} = 0$

2. $\frac{\partial Error}{\partial w_1} = 0$

$\vdots$

D+1. $\frac{\partial Error}{\partial w_D} = 0$

$$\sum_{i=1}^{N} \hat{y}_i = \sum_{i=1}^{N} y_i$$

$$\sum_{i=1}^{N} x_1^i \hat{y}_i = \sum_{i=1}^{N} x_1^i y_i$$

$$\vdots$$

$$\sum_{i=1}^{N} x_D^i \hat{y}_i = \sum_{i=1}^{N} x_D^i y_i$$

Multiple Linear Regression

How to create model

$$\sum_{i=1}^{N} \hat{y}_i = \sum_{i=1}^{N} y_i$$

$$\sum_{i=1}^{N} x_1^i \hat{y}_i = \sum_{i=1}^{N} x_1^i y_i$$

$$\vdots$$

$$\sum_{i=1}^{N} x_D^i \hat{y}_i = \sum_{i=1}^{N} x_D^i y_i$$

$$\left(X_b\right)^T \times Yhat = \left(X_b\right)^T \times Y$$

$$\begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_D^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_D^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^N & x_2^N & \cdots & x_D^N \end{bmatrix}^T \times \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_D^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_D^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^N & x_2^N & \cdots & x_D^N \end{bmatrix}^T \times \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

Multiple Linear Regression

How to create model (Global)

เนื่องจาก $Yhat = X_b W$

$$(X_b)^T \times Yhat = (X_b)^T \times Y$$

$$(X_b)^T \times (X_b W) = (X_b)^T \times Y$$

$$\left[(X_b)^T X_b\right] \times W = (X_b)^T \times Y$$

$$W = \left[(X_b)^T X_b\right]^{-1}\left[(X_b)^T Y\right]$$

$$W = \left[(X_b)^T X_b\right]^{-1}\left[(X_b)^T Y\right]$$

ค่าคลาดเคลื่อนมาตรฐาน(Standardized Residual)

วัดว่ามีค่าผิดปกติจากการพยากรณ์หรือไม่

$$Z_{\text{residual}} = \frac{\text{residual} - \text{mean}(\text{residual})}{\text{SD}(\text{residual})}$$

$$\text{residual} = y - \hat{y}$$

ค่าคลาดเคลื่อนมาตรฐานสูงสุดไม่ควรเกิน 3

Error Function

ที่นิยมใช้กัน มีอยู่ 4 ฟังก์ชัน ได้แก่

- SSE (Sum Squared Error) $\boldsymbol{SSE = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$
- MSE (Mean Squared Error) $\boldsymbol{MSE = \frac{\sum_{i=1}^{N}(y_i - \widehat{y_i})^2}{N}}$
- MAE (Mean Absolute Error) $MAE = \frac{\sum_{i=1}^{N}|y_i - \widehat{y_i}|}{N}$
- MAPE (Mean Absolute Percentage Error)

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y_i - \widehat{y_i}}{y_i}\right|$$

Multiple Linear Regression : ตัวอย่าง

จากการศึกษาความสัมพันธ์ระหว่างความตั้งใจเรียน(x_1) ความรู้พื้นฐานเดิม(x_2) และผลสัมฤทธิ์ทางการเรียน(y) มีความสัมพันธ์เชิงเส้น ข้อมูลที่เก็บจากกลุ่มตัวอย่าง จำนวน 30 คน (ดังปรากฏในตารางข้างล่าง) (ข้อมูลfile dataMR1)จงทำการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนักเรียนที่มีคะแนนความตั้งใจเรียน 18 คะแนน คะแนนความรู้พื้นฐานเดิม 7 คะแนน

$$\hat{y} = 5.597 + 0.595x_1 + 0.636x_2$$

ค่าพยากรณ์เมื่อ $x_1=18, x_2=7$: $\hat{y} = 5.597 + 0.595(18) + 0.636(7) = 20.759$

คาดว่าคะแนนผลสัมฤทธิ์ทางการเรียนของนักเรียนที่มีคะแนนความตั้งใจเรียน 18 คะแนน คะแนนความรู้พื้นฐานเดิม 7 คะแนน เท่ากับ 20.759 คะแนน

อธิบายค่า

w_0 : เมื่อไม่มีคะแนนความตั้งใจและคะแนนความรู้พื้นฐาน ผลสัมฤทธิ์ทางการเรียนเท่ากับ 5.597 คะแนน

w_1 : เมื่อคะแนนความตั้งใจเพิ่มขึ้น 1 คะแนนโดยกำหนดให้คะแนนความรู้พื้นฐานคงที่ ผลสัมฤทธิ์เพิ่มขึ้น 0.595 คะแนน

w_2 : เมื่อความรู้พื้นฐานเพิ่มขึ้น 1 คะแนนโดยกำหนดให้คะแนนคะแนนความตั้งใจคงที่ ผลสัมฤทธิ์เพิ่มขึ้น 0.636 คะแนน

คนที่	X1	X2	Y
1	14	6	19
2	15	9	20
3	20	10	25
4	13	9	18
5	9	6	16
6	23	9	28
7	13	9	18
8	14	10	19
9	16	10	20
10	9	4	14
11	9	5	14
12	11	6	16
13	8	5	14
14	11	6	16
15	12	8	17
16	8	4	13
17	14	9	19
18	11	7	16
19	12	8	17
20	9	5	14
21	8	4	13
22	7	4	12
23	6	3	11
24	20	9	18
25	13	8	18
26	11	9	20
27	14	9	21
28	14	9	22
29	12	4	14
30	17	10	24

Multiple Linear Regression : ตัวอย่าง

พยากรณ์ยอดขายของพนักงานที่มี
ประสบการณ์ทำงาน 8.5 ปี
ดูแลสาขา 15 สาขา และใช้เงิน
ในการโฆษณา 250000 บาท

Sale(ล้านบาท)	adver(พันบาท)	number(สาขา)	Exp(ปี)
32	249	15	12
47	292	18	15
18	183	14	8
25	201	16	12
49	310	21	16
41	248	20	14
52	246	18	13
38	241	14	10
36	288	13	12
29	191	15	8
43	248	21	17
28	210	18	9
24	256	20	11
36	275	16	10
41	241	19	13

$$\hat{y} = -8.168 + 0.0905(adver) - 0.0713(number) + 1.927(\exp)$$