

linear reg \rightarrow OLS [this week]
 \searrow gradient descent (next week)

mlr \rightarrow mistakes \rightarrow live

↳ five sections

linear reg
gradient desc
PCA

rest algo

Recap

19 April 2023 19:24

Simple Linear reg → multiple

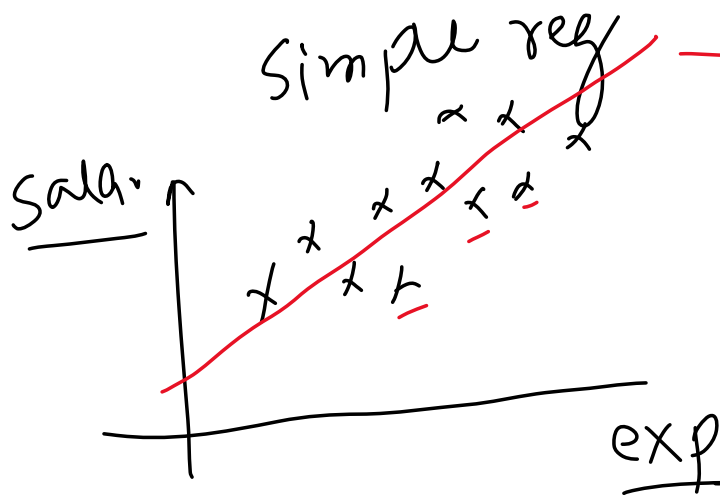
↓
X (input) y (out) → exp → salary

exp (years)	salary (Rs)
10	2500000
3	70000
1	25000

1000

(m)

(b)



$$y = mx + b$$

↑ ↑ ↑
Salary exp

What is Multiple Linear Regression

19 April 2023 19:24

1 input

exp | salary

city | age | gender | edu

more than 1 predictors

→ SLR X

→ MLR


→ model

[simple] { multiple }
(linear reg) linear reg

1000 students

low

<u>(p1)</u>	x_1	x_2 (p2)	placement
cgpa	iq		
8	80		8
9	90		9
5	120		15
...			

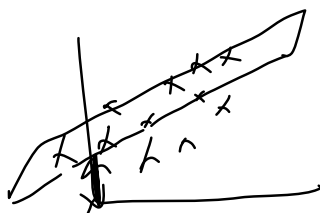
cgpa iq →  → lpa

line

x x x x x

$$y = mx + b$$

$$y = \beta_0 + \beta_1 x_1$$



$x_1 \rightarrow$ cgpa
 $x_2 \rightarrow$ iq

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$\beta_0 \rightarrow b$

$\beta_1 \rightarrow m$

$\beta_2 \rightarrow$

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

hyperplane in n-dim coordinate

cgpa | iq | 12th mark

m, b

coefficient

$$\beta_0, \beta_1, \dots, \beta_n$$

3d data → 2 input → 1 output

Python Code

19 April 2023 19:24

(3) m cols $\rightarrow n$ students

x_1	x_2	y
cgpa	iq	placement
β_0	β_1	β_2
8	80	8
7	70	7
5	120	15

predict

$$\begin{aligned}\hat{y}_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + \beta_4 x_{14} + \dots + \beta_m x_{1m} \\ \hat{y}_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_m x_{2m} \\ \hat{y}_3 &= \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \dots + \beta_m x_{3m} \\ &\vdots \\ \hat{y}_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_m x_{nm}\end{aligned}$$

$$y_1 = 8 \quad x_2 = 7 \quad y_3 = 15$$

$$\hat{y}_1 = 7 \quad \hat{y}_2 = 2 \quad \hat{y}_3 = 2$$

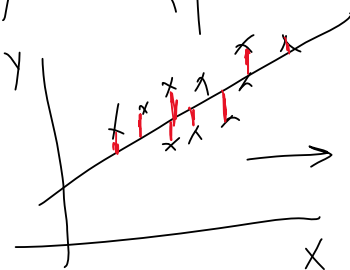
$$\begin{aligned}\hat{y}_1 &= \beta_0 + \beta_1 8 + \beta_2 80 \\ \hat{y}_2 &= \beta_0 + \beta_1 7 + \beta_2 70 \\ \hat{y}_3 &= \beta_0 + \beta_1 5 + \beta_2 120\end{aligned}$$

$$\hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + \beta_4 x_{14} + \dots + \beta_m x_{1m} \\ \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_m x_{2m} \\ \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \dots + \beta_m x_{3m} \\ \vdots \\ \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_m x_{nm} \end{bmatrix}$$

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

Diagram illustrating the matrix multiplication $\hat{y} = X \beta$. The matrix X is of size $n \times (m+1)$, where the first column is all ones (representing β_0) and the subsequent columns are the features $x_{11}, x_{12}, \dots, x_{1m}$ for the first student, and so on. The vector β is of size $(m+1) \times 1$, containing the coefficients $\beta_0, \beta_1, \dots, \beta_m$. The resulting vector \hat{y} is of size $n \times 1$.

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



$$d_1^2 + d_2^2 + d_3^2 + \dots + d_n^2$$

$$E = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

minimize
matrix form
matrix form

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} \quad \hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}_{n \times 1}$$

$$e = y - \hat{y} = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix}$$

$$e = y - \hat{y} = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix}_{n \times 1}$$

$$e^T e = \begin{bmatrix} y_1 - \hat{y}_1 & y_2 - \hat{y}_2 & \dots & y_n - \hat{y}_n \end{bmatrix}_{1 \times n} \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix}_{n \times 1} =$$

$$e^T e = (y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + \dots + (y_n - \hat{y}_n)^2$$

$$= \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$E = e^T e$$

↑ miniz

$$E = (y - \hat{y})^T (y - \hat{y}) = (y^T - \hat{y}^T) (y - \hat{y})$$

$$E = y^T y - y^T \hat{y} - \hat{y}^T y + \hat{y}^T \hat{y}$$

$$E = y^T y - 2 y^T \hat{y} + \hat{y}^T \hat{y} \rightarrow \text{eq (3)} \quad \hat{y} = X\beta$$

$$\textcircled{y} = f(\textcircled{x}) \rightarrow \textcircled{x}$$

$y \rightarrow$ data output

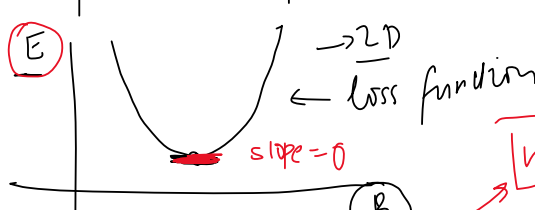
$x \rightarrow$ data input

$$f(x) = x^2 \rightarrow \frac{d}{dx} x^2 = 2x$$

$$E = y^T y - 2 y^T X \beta + (X \beta)^T (X \beta)$$

$$\textcircled{E} = y^T y - 2 y^T X \beta + \beta^T X^T X \beta$$

$E(\beta) =$ find such value of β matrix for which E is min



$$\frac{dE}{d\beta} = 0$$

$$\frac{dE}{d\beta} = 0 - 2 y^T X + 2 \beta^T X^T X = 0$$

$$\beta^T X^T X = y^T X$$

$$\beta^T X^T X = y^T X$$

$$\frac{d}{d\beta} \beta^T (X^T X) \beta$$

A is symmetric

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

$d \beta^T X^T X \beta$ A is symmetric
 $d\beta \downarrow$
 $x^T A x = 2x^T A$
 $= 2\beta^T X^T X$

$(X^T X)^T = X^T X$
 $A^T = A$

$\beta = [(X^T X)^{-1}]^T (Y^T X)^T$

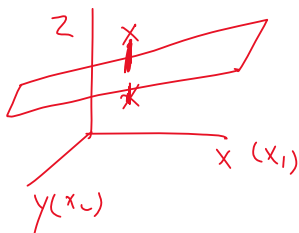
$\beta = [(X^T X)^{-1}]^T X^T Y$

$\beta = (X^T X)^{-1} X^T Y$ OLS

$\beta = \text{values}$
 $(X^T X)^{-1}$
 $(m+1) \times (m+1)$
 $(m+1) \times 1$

$m+1 \times n$ $n \times (m+1)$

$(m+1) \times (m+1)$ $(m+1) \times n$
 $(m+1) \times n$ $n \times 1$



exp | 12th month | salary

$\text{salary} = \exp x \beta_1 + 12^{\text{th}} \times \beta_2 + \beta_0$



$\beta_0 = 0$

$(X^T X)^{-1} X^T Y$

$\beta^T X^T X (X^T X)^{-1} = Y^T X (X^T X)^{-1}$

$\beta^T I = Y^T X (X^T X)^{-1}$

$\beta^T = Y^T X (X^T X)^{-1}$

$(\beta)^T = [Y^T X (X^T X)^{-1}]^T$

$(X^T X)^{-1}$ symmetric

$[(X^T X)^{-1}]^T = (X^T X)^{-1}$
 $X^T X = A$
 $A A^{-1} = I$
 $(A A^{-1})^T = I^T$
 $(A^{-1})^T A^T = I$
 $(A^{-1})^T A = I$
 $(A^{-1})^T A A^{-1} = I A^{-1}$
 $(A^{-1})^T I = A^{-1}$
 $(A^{-1})^T = A^{-1}$

global minimum



$$(X^T X)^{-1} X^T Y$$

$$\begin{array}{c|ccc} & x_1 & x_2 & y \\ \hline & | & | & | \\ & | & | & | \\ & | & | & | \end{array}$$

$$\boxed{Y^T \hat{Y} = \hat{Y}^T Y}$$

$$\underline{A^T B = B^T A}$$

$$(A^T B)^T = B^T A$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \leftarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Y = A \quad \hat{Y} = B$$

$$\underline{A^T B = (A^T + B)^T}$$

$$\boxed{C = C^T}$$

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

$n \times (m+1)$

$$(AB)^T = B^T A^T$$

$$(A^T)^T = A$$

$$\boxed{A^T B = C}$$

symmetrisch
matrix

$A^T B$ is symmetric

$$Y^T \hat{Y} = Y^T X \beta$$

$$(1 \times n) \quad n \times (m+1) \quad (m+1) \times 1$$

$$\begin{matrix} 1 \times n & n \times 1 \\ \hline \end{matrix}$$

$1 \times 1 \rightarrow$ scalar
(1) (2)

$A^T B$ is sym

Code From Scratch

19 April 2023 19:25

Problem with OLS solution

19 April 2023 19:25

linear reg \rightarrow OLS (closed form) \rightarrow gradient descent

10 input- OLS $\rightarrow \sqrt{\underbrace{(X^T X)^{-1}}_{\text{inverse}}}$

$O(n^3)$

$[X] \rightarrow n \text{ rows } (n+1)$
 $n \times (n+1)$

$\underbrace{(n+1) \times n}_{(n+1)} \underbrace{n \times (n+1)}_{(n+1)}$

(100×100)

$X \rightarrow n \times (n+1)$
 $(n+1) \times n \quad n \times (n+1)$
 \downarrow
 $\frac{(n+1) \times (n+1)}{\|X\|}$

$(100)^3$
 $[1000000]$
 $(10000)^3$