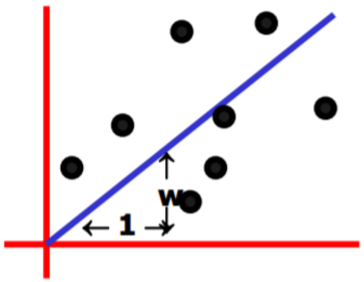
Regression

- Linear Regression $y_n = w_1 x_{1,n} + w_2 x_{2,n} + \ldots + w_M x_{M,n} + w_0, \forall n = 1 \ldots N$ N=5, M=1



$$y_1 = w_1 x_1 + w_0,$$
$$y_2 = w_1 x_2 + w_0,$$
$$\ldots$$
$$y_5 = w_1 x_5 + w_0,$$

$$\begin{bmatrix} y_1 \\ \ldots \\ y_5 \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ \ldots & \ldots \\ x_5 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_0 \end{bmatrix}$$

$$y_n = \phi(x_n)\mathrm{w}$$
$$\mathrm{y} = \phi\mathrm{w}$$

inputs	outputs
$x_1 = 1$	$y_1 = 1$
$x_2 = 3$	$y_2 = 2.2$
$x_3 = 2$	$y_3 = 2$
$x_4 = 1.5$	$y_4 = 1.9$
$x_5 = 4$	$y_5 = 3.1$

$x+y=1$
$x+2y=5$

$$\rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$$

1. Regression analysis describes the relationship between two (or more) variables
2. Linear regression assumes that the expected value of the output given an input, E[y|x], is linear
3. Simplest case: **y** = **ϕw** for some unknown a

   If $x+y=1$, 해가 무수히 많다.

   - Given the data, we can estimate **w**

- Linear regression

$y = ax + b$

$\vec{y} = X\vec{w}$

$$\text{ex)}\quad \underset{1\times 2}{\begin{bmatrix} x & 1 \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} w_0 \\ w_1 \end{bmatrix}} = \underset{1\times 1}{y}$$

$$\underset{\vec{w}}{\operatorname{argmin}} \|X\vec{w} - \vec{y}\|^2$$

X란 행렬

$$\underset{X:\ n\times 2}{\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} w_0 \\ w_1 \end{bmatrix}} = \underset{n\times 1}{\begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}}$$

$$\Rightarrow (X\vec{w} - \vec{y})^T (X\vec{w} - \vec{y})$$

$$\Rightarrow (\vec{w}^T X^T - \vec{y}^T)(X\vec{w} - \vec{y})$$

$(2\times 1)^T (n\times 2)^T (n\times 1) \to 1\times 1$

$$\Rightarrow \vec{w}^T X^T X \vec{w} + \vec{y}^T \vec{y} - \vec{w}^T X^T \vec{y} - (\vec{y}^T X \vec{w})^T$$

$(n\times 1)^T (n\times 2)(2\times 1) \to 1\times 1$

$$\frac{d}{d\vec{w}} \Rightarrow 2X^T X \vec{w} - 2X^T \vec{y} = 0$$

$$\Rightarrow X^T X \vec{w} = \vec{y}^T X$$

$$\Rightarrow \vec{w} = (X^T X)^{-1} X^T \vec{y}$$

Regression

- Linear Regression

$$\hat{\mathrm{w}} = \underset{\mathrm{w}}{\operatorname{argmin}} \|\phi \mathrm{w} - \mathrm{y}\|_2^2$$

$$\|\phi \mathrm{x} - \mathrm{y}\|_2^2 = (\phi \mathrm{w} - \mathrm{y})^{\mathrm{T}}(\phi \mathrm{w} - \mathrm{y})$$

w

To find minimum

W

$$(\mathrm{w}^{\mathrm{T}}\phi^{\mathrm{T}} - \mathrm{y}^{\mathrm{T}})(\phi \mathrm{x} - \mathrm{y})$$

$$= \left(\mathrm{w}^{\mathrm{T}}\phi^{\mathrm{T}}\phi \mathrm{w} + \mathrm{y}^{\mathrm{T}}\mathrm{y} - 2\mathrm{w}^{\mathrm{T}}\phi^{\mathrm{T}}\mathrm{y}\right)$$

$$= \frac{\partial}{\partial \mathrm{w}}\left(\mathrm{w}^{\mathrm{T}}\phi^{\mathrm{T}}\phi \mathrm{w} + \mathrm{y}^{\mathrm{T}}\mathrm{y} - 2\mathrm{w}^{\mathrm{T}}\phi^{\mathrm{T}}\mathrm{y}\right)$$

T

$$= \left(2\phi^{\mathrm{T}}\phi \mathrm{w} - 2\mathrm{w}\cancel{\phi}^{\mathrm{T}}\mathrm{y}\right) = 0$$

$$\hat{\mathrm{w}} = \left(\phi^{\mathrm{T}}\phi\right)^{-1}\phi^{\mathrm{T}}\mathrm{y}$$

input output

Ref.1

$$\hat{\mathrm{x}} = \underset{\mathrm{x}}{\operatorname{argmin}} \|\mathrm{Ax} - \mathrm{b}\|_2^2$$

$$\hat{\mathrm{x}} = \left(\mathrm{A}^{\mathrm{T}}\mathrm{A}\right)^{-1}\mathrm{A}^{\mathrm{T}}\mathrm{b}$$

Ref.2

$$\phi = \begin{bmatrix} x_1 & 1 \\ \cdots & \cdots \\ x_5 & 1 \end{bmatrix}$$

$$\mathrm{w} = \begin{bmatrix} w_1 \\ w_0 \end{bmatrix}$$

Ref.3

$$\frac{d}{\partial \mathbf{w}}(\mathbf{w}^{\mathrm{T}}\mathbf{x}) = \begin{bmatrix} \dfrac{dw^{\mathrm{T}}x}{\partial w_1} \\ \dfrac{dw^{\mathrm{T}}x}{\partial w_2} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}$$

$$\mathbf{w}^{\mathbf{T}}\mathbf{x} = w_1 x_1 + w_2 x_2$$

Ref.4

W^2이니까

$$\frac{d}{\partial \mathbf{w}}(\mathbf{w}^{\mathrm{T}}\mathbf{A}\mathbf{w}) = 2\mathbf{A}\mathbf{x}, \ \text{(A: symmetric matrix)}$$