MSE Loss f^n in Scalar Notation

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

$\hat{y}_i \Rightarrow$ Predictions from Model

$y_i =$ Actual Values

$\hat{y} - y \Rightarrow$ Error Terms / Residuals.

$$\hat{y} = w_0 + w_1 x$$

$$L = \frac{1}{N} \sum \left[(w_0 + w_1 x) - y \right]^2$$

$$L = \frac{1}{N} \sum \left[w_0^2 + w_1^2 x^2 + 2w_0 w_1 x - y^2 \right]$$

Gradients of Loss f^n w.r.t. model coeff.

$$\frac{\partial L}{\partial w_0} = \frac{1}{N} \sum \left[2w_0 + 2 w_1 x \right] = 0$$

$$\frac{\partial L}{\partial w_1} = \frac{1}{N} \sum \left[2w_1 x^2 + 2w_0 x \right] = 0$$

$$\frac{\partial L}{\partial w_0} = \frac{1}{N} \sum [w_0 + w_1 x] \cdot 1 = 0 \qquad —① $$

$$\frac{\partial L}{\partial w_1} = \frac{1}{N} \sum [w_0 + w_1 x] \cdot x = 0 \qquad —② $$

$$② = x \cdot ①$$

$$w_1 = \frac{(\bar{xy}) - (\bar{x}) \cdot (\bar{y})}{(\bar{x^2}) - (\bar{x})^2}$$

$(\bar{x}) \rightarrow$ mean x value

$\bar{xy} \Rightarrow$ means mean of the products of values of x & y

x	y	xy	x^2
-	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-

Loss f^n In Vector Notation: $\hat{y} = XW$

$$L = \frac{1}{N} (\hat{y} - y)^T (\hat{y} - y)$$

$$L = \frac{1}{N}(Y-y)(Y-y)$$

$$= \frac{1}{N}\left[(xw-y)^T (xw-y)\right]$$

$$L = \frac{1}{N}\left[w^T x^T \cdot xw - y^T xw - w^T x^T y + y^T y\right]$$

$$\frac{\partial L}{\partial w} = \left(\quad \cdots \quad \right] = 0$$

$$= \quad \underline{\quad - \quad -}$$

$$= \quad \overline{\quad - \quad -}.$$

$$\Rightarrow (x^T x)\, w - x^T y = 0$$

Model coeff.
for Linear Reg.
$$\boxed{w = (x^T x)^{-1} \cdot x^T y} \quad \rightarrow SLR \\ \rightarrow MLR$$

Model and **Algo.**

$$\boxed{Model = Algo(X, Y)}$$

$$\begin{cases} Model_1 = Algo1\,(Data) \\ Model_2 = Algo\,2\,(Data) \end{cases}\Bigg] \begin{array}{l} Same \\ data \\ Change\ the \\ Algo \end{array}$$

$$DT_1 \quad vs \quad DT_2 \quad vs \quad DT_3 ----$$

Hyperparam.

$$y = f(x)$$
$$= \quad \downarrow$$
$$x = 2 \qquad Some\ operation \\ on\ x$$

$$y = \sin(x)$$
$$y = 2x^2 - 3a$$
$$y = \cos(x) \cdot \log(x)$$