Model: $y_i = w_1 x_i + w_0$

Loss function: $$J = \frac{1}{2N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

$$= \frac{1}{2N}\sum_{i=1}^{N}(y_i - w_1 x_i + w_0)^2$$

$$\nabla_{w_1} J = \frac{1}{N}\sum_{i=1}^{N} x_i \cdot (w_1 x_i + w_0 - y_i)$$

$$\nabla_{w_0} J = \frac{1}{N}\sum_{i=1}^{N}(w_1 x_i + w_0 - y_i)$$

Biểu diễn dưới dạng ma trận.

$$w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}; \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$$

$$\hat{y} = X \cdot w$$

Loss function: $$J = \frac{1}{2N}\|y - Xw\|_2^2$$

$$= \frac{1}{2N}(y - Xw)^T(y - Xw)$$

$$= \frac{1}{2N}(y^T - w^T X^T)(y - Xw)$$

$$= \frac{1}{2N}(y^T y - y^T X w - w^T X^T y + w^T X^T X w)$$

Ta có:

$\Rightarrow x_i w = [1, x_1, \dots x_m]\begin{bmatrix} w_0 \\ w_1 \\ \dots \\ w_m \end{bmatrix} = w_0 + w_1 x_1 + \dots w_m x_m$

$\Rightarrow \nabla_{w_i}(x_i w) = x_i$

Date:..............................No...............

$\Rightarrow \nabla_w (Xw) = X$

+) $(w^T x^T) = (xw)^T$.

$\nabla_w (w^T x^T) = \nabla_w (xw)^T = \nabla_w (xw) = x$.

$\Rightarrow \nabla_w (w^T X^T) = X$.

+) $\dfrac{\nabla(w^T X w)}{\nabla w} = (\nabla w^T) . Xw + \nabla(Xw) w^T$ (X đối xứng)

$= Xw + w^T X$

$= Xw + X^T w = 2 w^T X$.

Vì vậy, đạo hàm Loss function là:

$\nabla J = \dfrac{1}{2N}(0 - y^T X - y^T X + 2 w^T X^T X)$ ($X^T X$ đối xứng)

Để hàm loss cực tiểu thì w thoả mãn $\nabla J = 0$.

$\Rightarrow 2 w^T X^T X = 2 y^T X$

$\Leftrightarrow X^T X w = X^T y$

$\Leftrightarrow w = (X^T X)^{-1} . X^T y$.

Nếu $X^T X$ không khả nghịch, ta sẽ dùng giả nghịch đảo.

$\nabla J = \dfrac{1}{2N} (2 w^T x^T X - 2 y^T X)$

$= \dfrac{1}{N} X^T (Xw - y) = \dfrac{1}{N} X^T (\hat{Y} - Y)$.