$$\hat{w} = \arg\min_{w} \ (w-1)^2+3 \qquad \hat{w} = ?\ 1 \qquad \hat{x} = \arg\min_{x} f(x)$$

$$\hat{w} = \arg\min_{w} \ \frac{(w-1)^2+3}{2} \qquad \hat{w} = \underbrace{f'(x) = 0}$$

Training a simple LR model $\iff$ $\hat{w} = \arg\min_{w} \dfrac{1}{2N}\displaystyle\sum_{n=1}^{N}\left(\underbrace{h_w(x^{(n)})} - t_n\right)^2$

$$= \arg\min_{W} \frac{1}{2N}\sum_{n=1}^{N}\left(W^T X^{(n)} - t_n\right)^2$$

$$= \arg\min_{\underbrace{W}_{[w_0, w_1]}} \frac{1}{2N}\sum_{n=1}^{N}\left(\underbrace{w_0 + w_1 x_n} - t_n\right)^2 = J(W) = J(w_0, w_1).$$

$$\hat{w}_0, \hat{w}_1 = \arg\min_{w_0, w_1} J(w_0, w_1)$$ Gradient of $J = 0$

$$\nabla J = \frac{\partial J}{\partial W} = \left[\frac{\partial J}{\partial w_0}, \frac{\partial J}{\partial w_1}\right] = 0 \Rightarrow w_0 = ?, \quad w_1 = ?$$

$$J(\vec{w}) = J(w_0, w_1) = \frac{1}{2N} \sum_{n=1}^{N} \big(\underbrace{w_0 + w_1 x_n} - t_n \big)^2$$

$$\frac{\partial f(x)^2}{\partial x} = 2 f(x) \cdot \frac{\partial f}{\partial x}$$

$$\frac{\partial J}{\partial w_0} = \frac{1}{\cancel{2}N} \sum_{n=1}^{N} \cancel{2} (w_0 + w_1 x_n - t_n) \cdot 1 = 0$$

$$\sum_{n=1}^{N} w_0 + \sum_{n=1}^{N} w_1 x_n - \sum_{n=1}^{N} t_n = 0$$

$$\boxed{w_0 \underbrace{N}_{a_{11}} + w_1 \underbrace{\sum_{n=1}^{N} x_n}_{a_{12}} = \underbrace{\sum_{n=1}^{N} t_n}_{T_1}}$$

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

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$$

$$\frac{\partial J}{\partial w_1} = \frac{1}{\cancel{2}N} \sum_{n=1}^{N} \cancel{2} (w_0 + w_1 x_n - t_n) x_n = 0$$

$$\sum_{n=1}^{N} w_0 x_n + \sum_{n=1}^{N} w_1 \cdot x_n^2 - \sum_{n=1}^{N} t_n x_n = 0$$

$$\boxed{w_0 \underbrace{\sum_{n=1}^{N} x_n}_{a_{21}} + w_1 \underbrace{\sum_{n=1}^{N} x_n^2}_{a_{22}} = \underbrace{\sum_{n=1}^{N} t_n x_n}_{T_2}}$$

$$\boxed{Aw = T} \Rightarrow \boxed{w = A^{-1} T}$$

$$J(w) = \frac{1}{2N}\sum_{n=1}^{N}\left(h_w(\vec{x}_n) - t_n\right)^2$$

$$W^T\vec{x}_n = w_0 + w_1 x_n + \dots + w_M x_n^N$$

$$[w_0, w_1, \dots, w_M] \qquad [1, x_n, x_n^2, \dots, x_n^M]$$

$$\nabla J = \frac{\partial J}{\partial W} = 0 \iff \left[\frac{\partial J}{\partial w_0}, \frac{\partial J}{\partial w_1}, \frac{\partial J}{\partial w_2}, \dots, \frac{\partial J}{\partial w_M}\right] = 0$$