

$$\hat{y} = w^T x + b$$

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

$$\frac{\partial E}{\partial w_j} = \sum_{i=1}^n 2(y_i - w^T x_i)(-x_i) = 0$$

$$\sum_{i=1}^n w^T x_i x_{ij} = \sum_{i=1}^n y_i x_{ij}$$

$$a^T b = \sum_{i=1}^n a_i b_i = w^T (X^T X) = y^T X$$

$$(X^T X)w = X^T y$$

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