$W^T X = y^T$ $\quad X \in \mathbb{R}^{3\times 2} \quad y \in \mathbb{R}^{2\times 1}$

$$[w_1 \ \ w_2 \ \ w_3] \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} = [y_1 \ \ y_2] \qquad y^T \in \mathbb{R}^{1\times 2}$$

$1\times 3 \qquad 3\times 2$

$W^T = [w_1 \ \ w_2 \ \ w_3] \qquad W = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}.$

$$[\underline{w_1 x_{11} + w_2 x_{21} + w_3 x_{31}} \quad \underline{w_1 x_{12} + w_2 x_{22} + w_3 x_{32}}] = [\underline{y_1} \ \ \underline{y_2}]$$

1×2

$y = ax + b$

$$\nabla_x (x^T A x) = (A + A^T)x = 2Ax.$$

A symmetric

$$\nabla_x (\underline{x}^T \underline{b}) = \underline{b}$$

$$\nabla_x \left(\sum_{i=1}^{d} x_i b_i \right) = \begin{pmatrix} \frac{\partial}{\partial x_1} \sum_{i=1}^{d} x_i b_i \\ \vdots \\ \frac{\partial}{\partial x_d} \sum_{i=1}^{d} x_i b_i \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{pmatrix} = \underline{b}.$$

$\boxed{\underline{x}^T \underline{b} = \underline{b}^T \underline{x}}$ ← 1-dim/scalar/real #

$x^T b \in \mathbb{R}$ $\qquad (x^T b) = (x^T b)^T = b^T (x^T)^T$

$= b^T x$

$(AB)^T = B^T A^T$

Even, Under, Over

$m = d$ $\qquad m < d$ $\qquad m > d.$

$X = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$

# of equations $\nearrow$ $m = d$ $\searrow$ # unknown

$Xw = y$ $\qquad X \in \mathbb{R}^{m \times d}$

$X = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & & \vdots \\ x_{m1} & & x_{md} \end{bmatrix}$ without offset/bias

$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1d} \\ \vdots & \vdots & & \vdots \\ 1 & x_{m1} & & x_{md} \end{bmatrix}$ with offset/bias

Predict the value of y_{new} given a new test sample x_{new} when you use an affine model (bias).

Predict the value of y_{new} given a new test sample X_{new} when you use a linear model (no bias).

· Polynomial_Features

$\underbrace{w_0}+\underbrace{w_1 x}+\underbrace{w_2 x^2}$: polynomial of order 2.

monomial.

2 variables polynomial order 2.

$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2 + w_{12} x_1^1 x_2^1$

$+ w_{11} x_1^2 + w_{22} x_2^2$

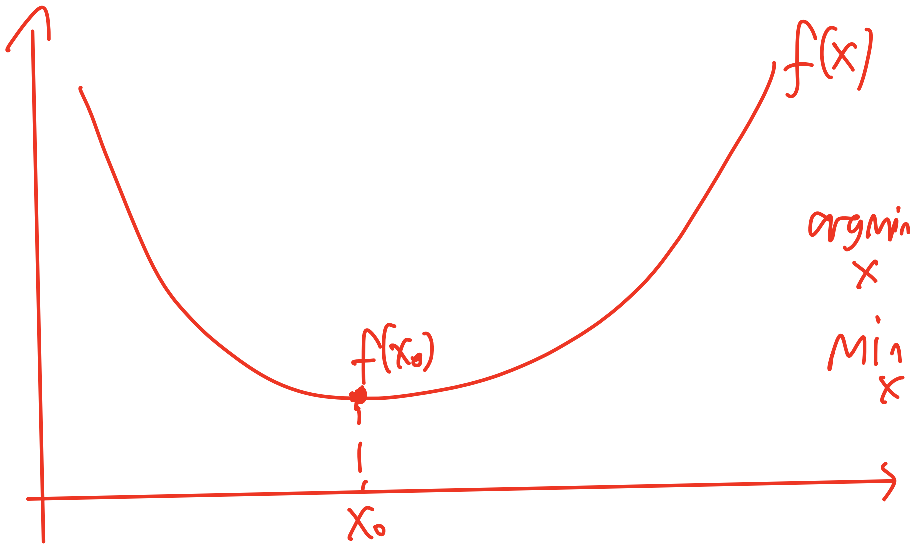
$w_{11} x_1^2 x_2^0$ $\quad$ $w_{22} x_1^0 x_2^2$

Primal $\quad$ Dual

$d+1 < m$ $\quad$ $m < d+1$

$$(X^T X + \lambda I)^{-1} X^T y = X^T (X X^T + \lambda I)^{-1} y$$

$\lambda > 0.$



$$\underset{x}{\operatorname{argmin}} f(x) = x_0$$

$$\min_x f(x) = f(x_0)$$

$\binom{d+p}{p}$ = # of monomial / coefficients of a d-variable order-p polynomial.