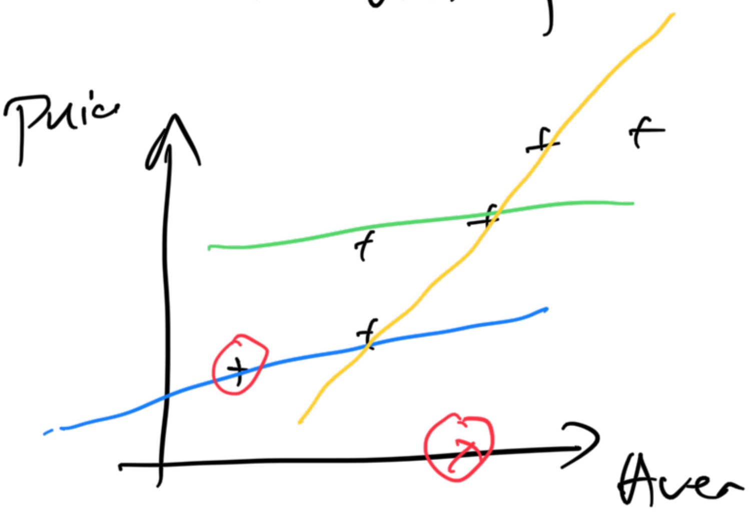
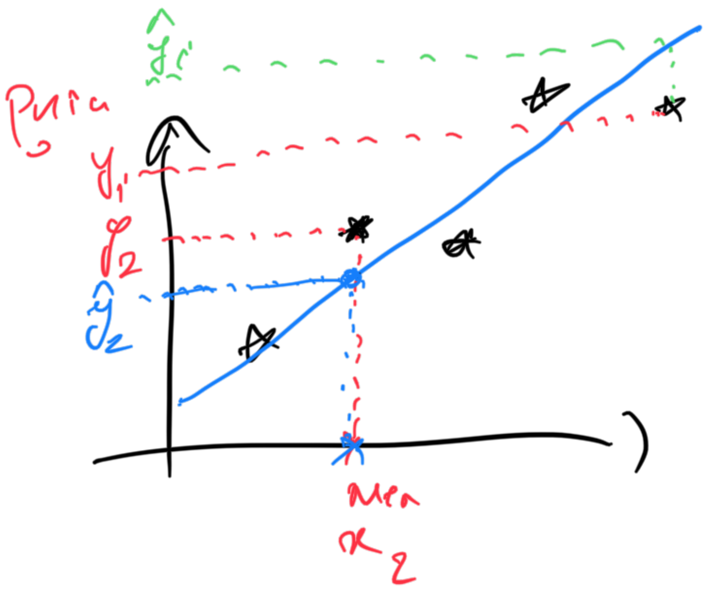
Linear Regression.

Area	Price
x_i	y_i

n data points





$$e_i = (y_i - \hat{y}_i)^2$$

$$\boxed{\mathcal{L} = \sum_{i=1} \left(y_i - \hat{y}_i \right)}$$

$$\hat{y} = mx + b$$

pure

$$\theta = \begin{bmatrix} b \\ m \end{bmatrix}$$

$$\boxed{\mathcal{L}(\theta) = \frac{1}{n} \underbrace{\sum_{i=1}^{n} \left[\left(y_i - (m x_i + b) \right)^2 \right]}}$$

m, b

$f(x, y)$ $\quad$ $f'(x) = 0$

$$\frac{\partial \mathcal{L}}{\partial m} = 0 \qquad \frac{\partial \mathcal{L}}{\partial b} = 0$$

m^* $\quad$ θ^* $\quad$ b^*

$$\boxed{h(x) = f(g(x))}$$

$$f : x \longmapsto x^2$$

$$g : x \longmapsto \exp(x)$$

$$h(x) = (\exp(x))^2$$

$$= \exp(2x)$$

$$= 2\exp(2x)$$

$$\boxed{h'(x) = g'(x) \times f'(g(x))}$$

$$g(m) = y_i - \underline{m}\, x_i - b$$

$$g'(m) = -x_i$$

$$\boxed{\hat{y}_i = m\, x_i + b.1}$$

$$1 \;.\; m\, x_i^2 + b.1$$

$$\hat{y}_i = m_1 x_i + m_2 \ldots$$

$$\hat{y} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_i \\ \vdots \\ \hat{y}_n \end{bmatrix} = \underbrace{\begin{bmatrix} m x_1 + b \cdot 1 \\ \vdots \\ m x_i + b \cdot 1 \\ \vdots \\ m x_n + b \cdot 1 \end{bmatrix}}$$

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{X} \cdot \underbrace{\begin{bmatrix} b \\ m \end{bmatrix}}_{\theta}$$

$$\boxed{X\Theta = \hat{y}}$$

$$\uparrow$$

$$M \cdot M^{-1} = I$$

$$X\theta = \hat{y}$$

X square and invertible

$$X^{-1}X\theta = X^{-1}y$$

$$I \cdot \theta = X^{-1}y$$

$$\boxed{\theta = X^{-1}y}$$

$$X\theta = y$$

$$\boxed{X^{T}X\theta = X^{T}y}$$

Normal Equation

X size [n m] $X^{T}X$ size [m m]

X^{T} size [m n]

$$\boxed{\theta^{*} = (X^{T}X)^{-1}\underset{\substack{A \\ B}}{X^{T}}\,y}$$

$$\theta^{*} = \begin{bmatrix} b^{*} \\ m^{*} \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_1 \end{bmatrix} \leftarrow \text{House 2}$$

$$X = \begin{bmatrix} \vdots & & \vdots \\ 1 & \cdots & x_n \end{bmatrix} \leftarrow \text{House } n$$

$$\hat{y} = X\theta^*$$

$$X' = \overbrace{\begin{bmatrix} 1 & x_1^1 & \cdots & x_1^k \\ \vdots & \vdots & & \vdots \\ 1 & x_n^1 & \cdots & x_n^k \end{bmatrix}}^{k} \quad \underline{(n, k+1)}$$

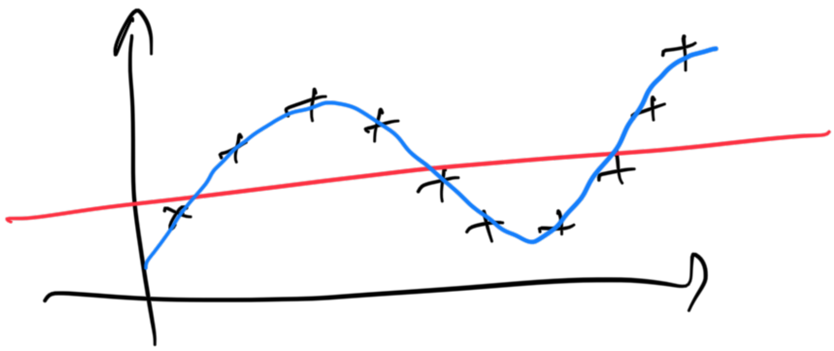
$$\theta = \begin{bmatrix} b \\ m \end{bmatrix} \rightsquigarrow \theta' = \begin{bmatrix} w_0 \\ \vdots \\ w_k \end{bmatrix} \; k+1$$

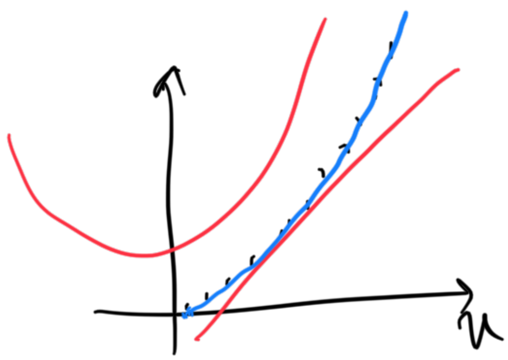
$$X'\theta' = \hat{y}$$

$$X'^T X'\theta' = X'^T y$$

$$\boxed{\theta'^* = (X'^T X')^{-1} X'^T \cancel{y}}$$

<u>Non Linear Problem</u>





x^2

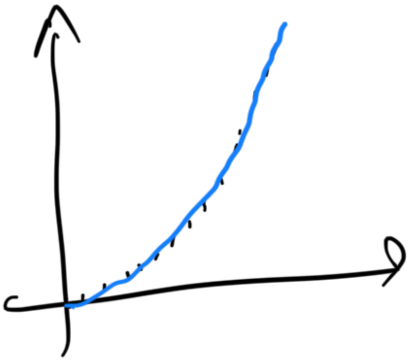
$$\hat{y} = mx + b$$

$$\hat{y} = \underline{m}\, x^2 + \underline{b}$$

$1 \quad x \quad x^2 , \ldots , \quad x^n$

$x^0 \quad x^1$

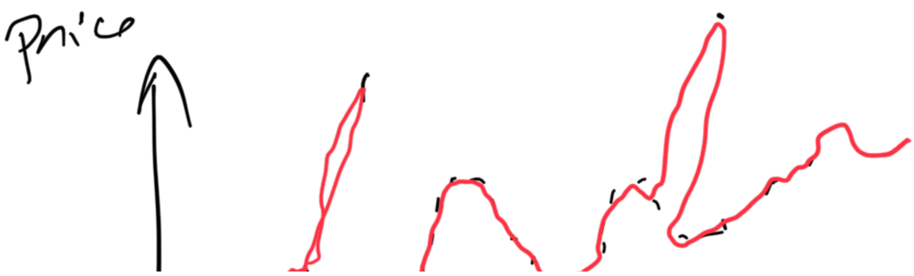
$$Y = w_2 x^2$$

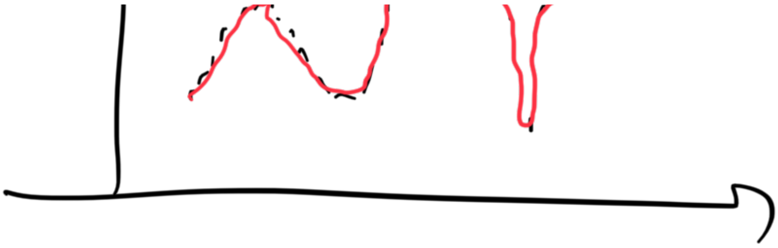


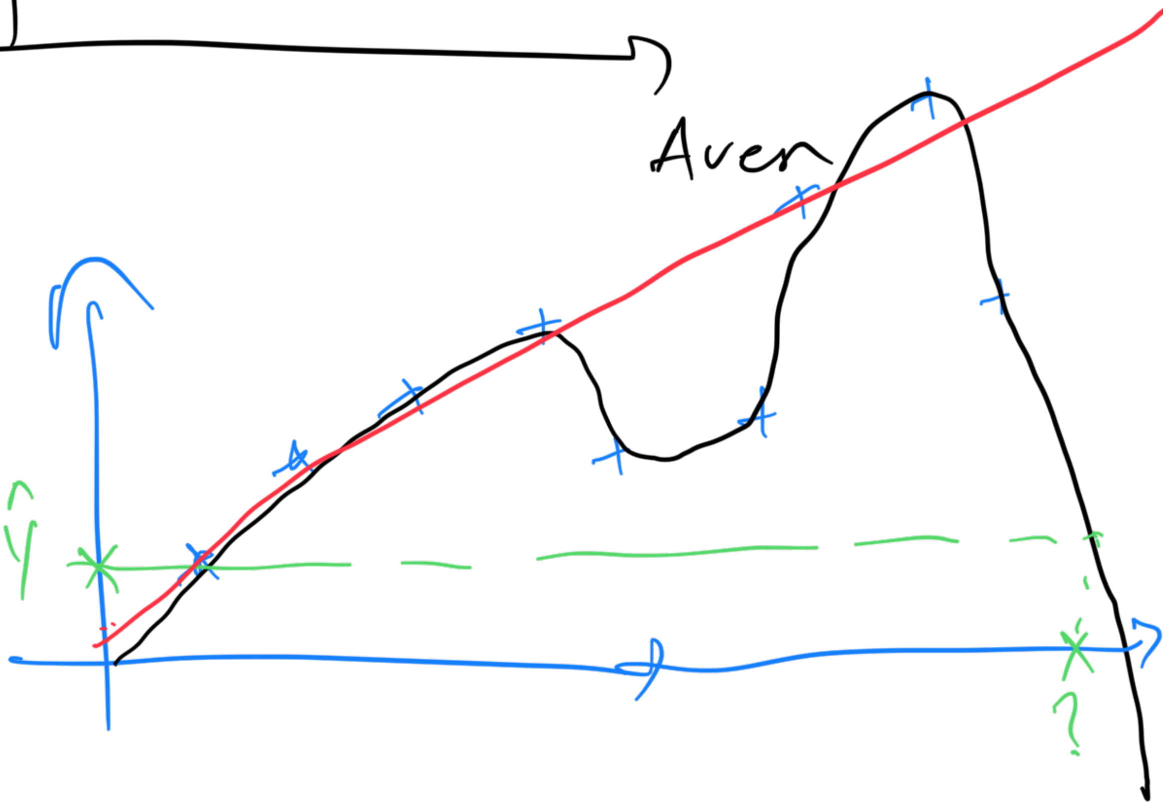
$$X = \begin{bmatrix} 1 & x_1^1 & \cdots & x_1^k \\ \vdots & \vdots & & \\ 1 & x_n^1 & \cdots & x_n^k \end{bmatrix}$$

$$\begin{array}{cc} \uparrow & \uparrow \\ \text{Area} & \text{place} \end{array}$$

$$X^* = \begin{bmatrix} 1 & x_1^1 & (x_1^1)^2 & (x_1^1)^3 & \cdots \\ \vdots & & & & \\ 1 & & & & \end{bmatrix}$$







Overfitting:

very good on training data

very bad on validation data
unseen

underfitting.

Bad even on the training.