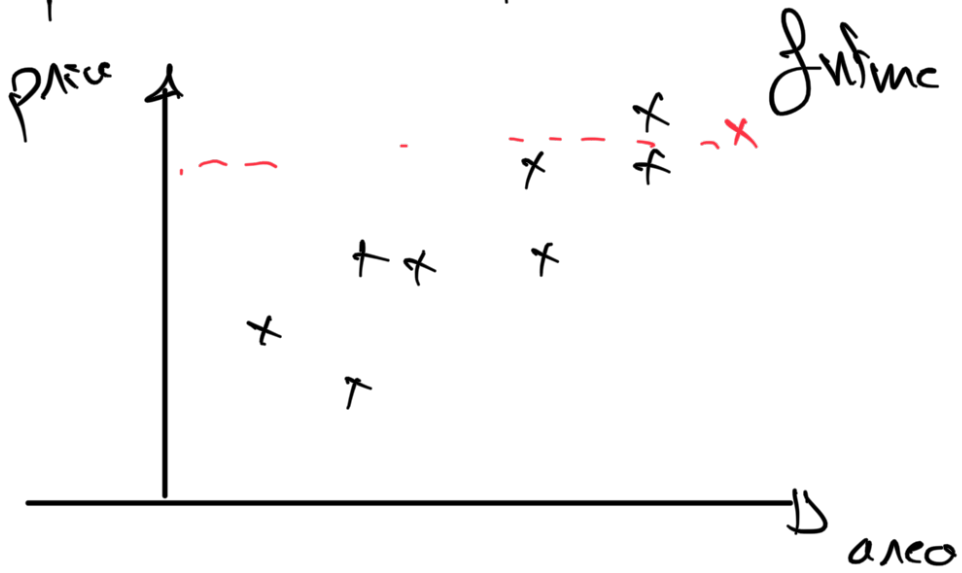
Recap: linear regression

Given some points

→ predict new points in the future



How to predict

→ Machine Learning (build model)

Regression task:

$$\hat{y}_i = m x_i + b$$

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

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

We want to solve θ

⇒ find best θ that minimizes loss function

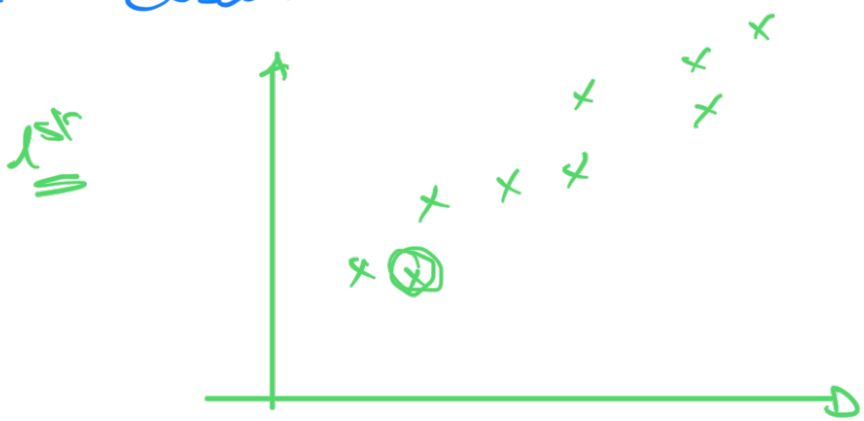
cost

MSE: Mean Squared Error

$$J(\theta) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \underbrace{\hat{y}_i}_{X_i\theta})^2$$

$$\mathcal{L}(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2$$

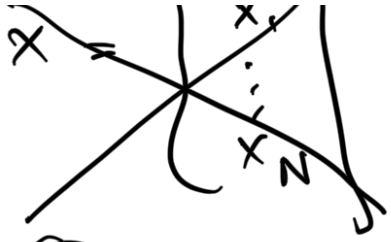
Toy Case:



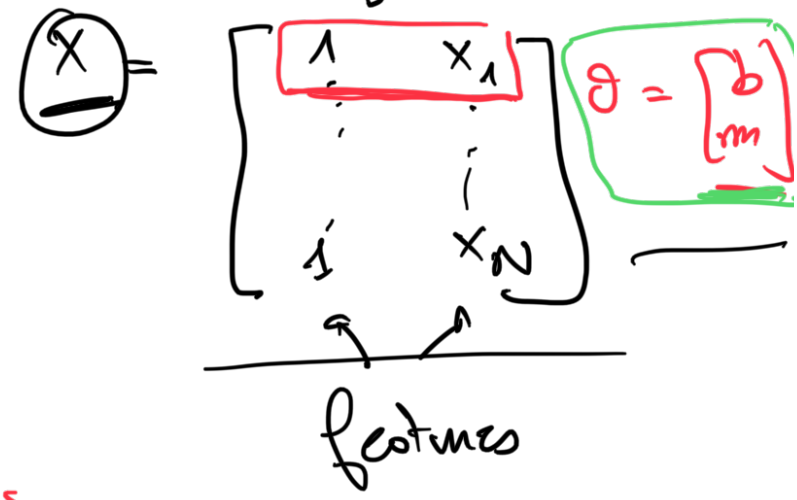
X-points = $\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$ $y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$

linear relation between X_points and y

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

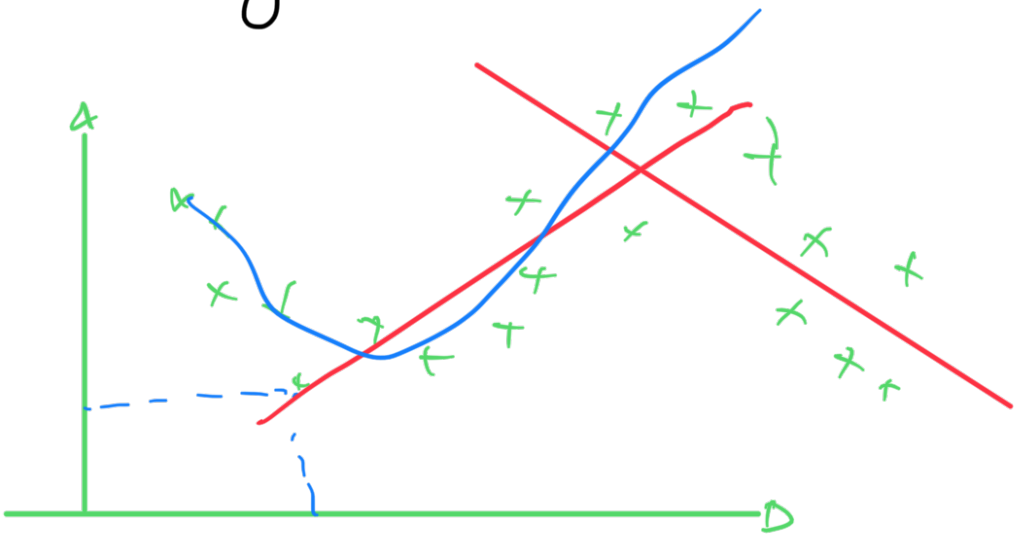


$\theta = \begin{bmatrix} m \\ b \end{bmatrix}$



learnable parameters θ

2nd case



Non linear relation

X_points $= \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$ $y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$

↑ label (True label)

(x_i, y_i)

$\hat{y} = X\theta$

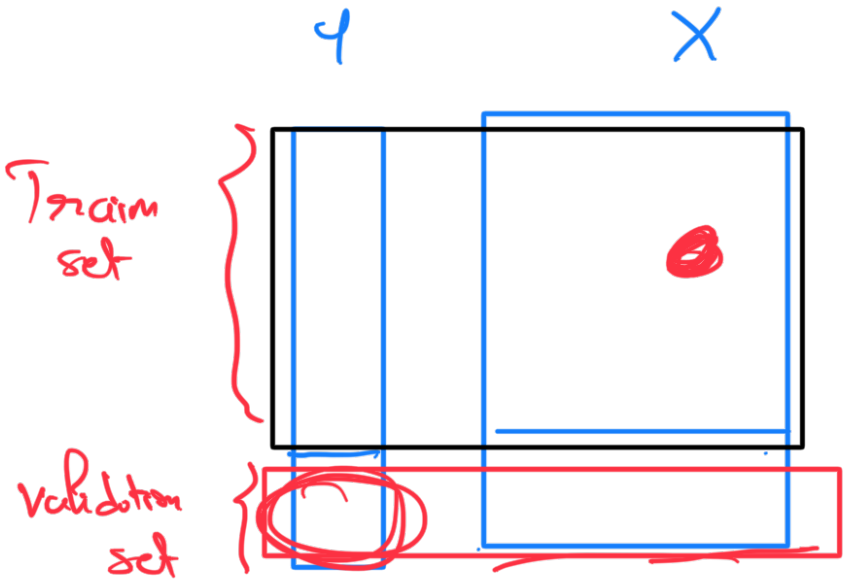
$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^m \end{bmatrix} \qquad \theta = \begin{bmatrix} b \\ m_1 \\ \vdots \\ m_M \end{bmatrix}$$

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

$$J(\theta)$$

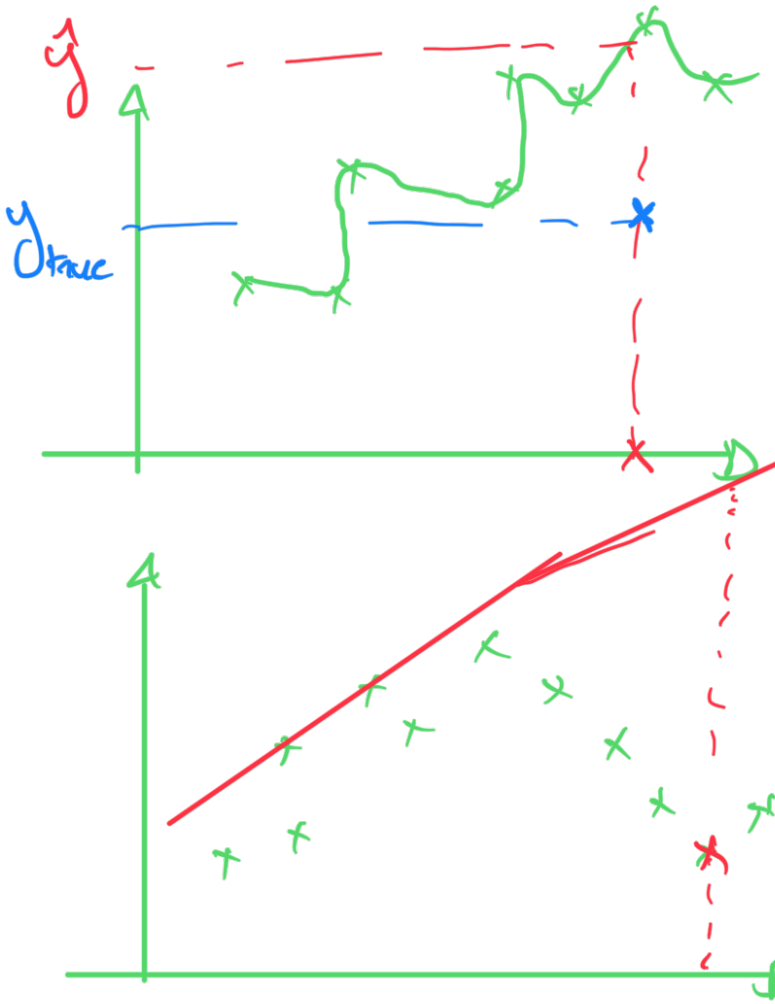
$$\frac{\partial J}{\partial \theta} = 0$$

$$\Rightarrow \text{closed form: } \underline{\theta} = (X^T X)^{-1} X^T y$$



X : X_{train} , X_{val}

Y : Y_{train} , Y_{val}



overfitting

Underfitting