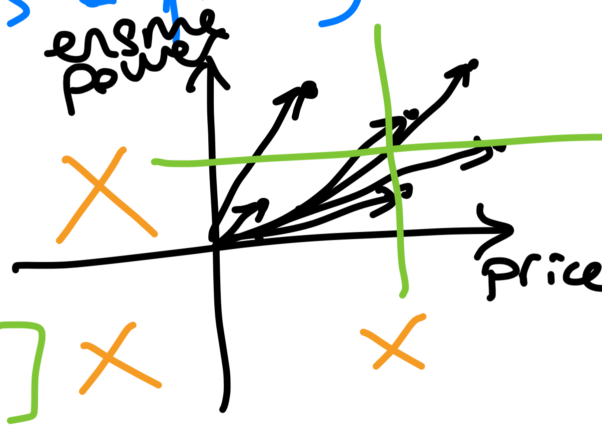
Supervised Learning

$\mathcal{X} = \{(x_i, y_i)\}_{i=1}^{N}$ } training set

Task: predicting whether a car is a family car or not.

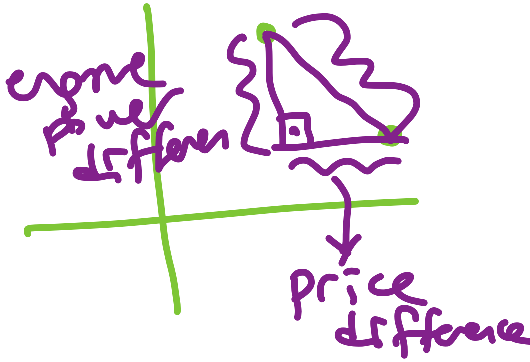
$x_i \in \mathbb{R}^2$



$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix}_{2\times 1}$$

$x_{i1} \rightarrow$ price [10000\$ - 1000000\$]

$x_{i2} \rightarrow$ engine power [1500 - 4000 cc]



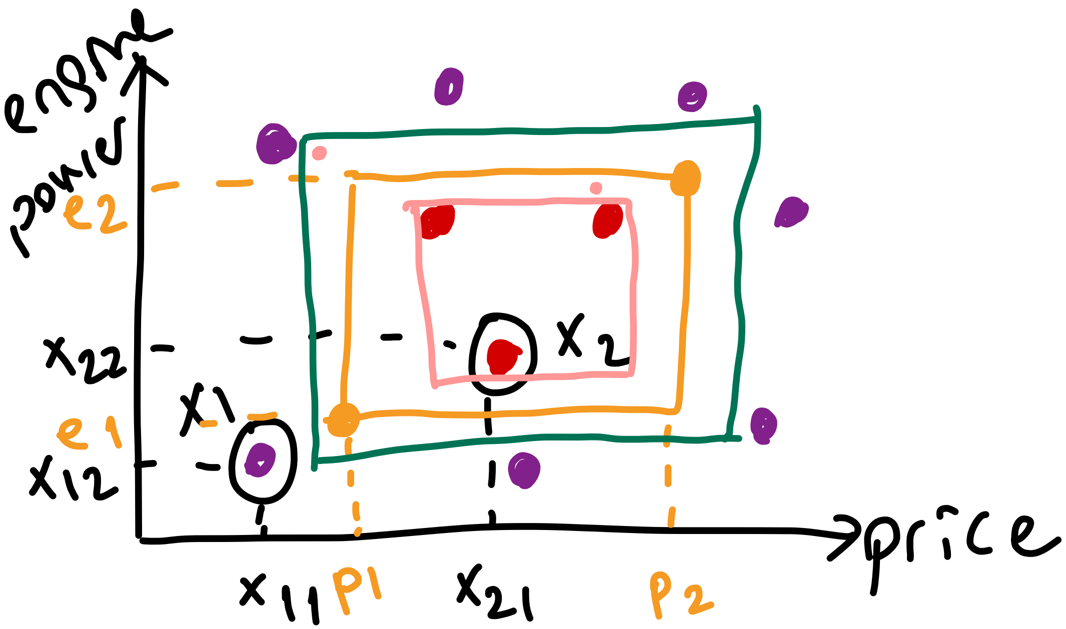
$$y_i = \begin{cases} +1 \quad 1 \quad + & \text{if } x_i \text{ is a family car} \\ -1 \quad 0 \quad - & 0 \text{ otherwise} \end{cases}$$

data matrix

$$X = \begin{array}{l} \\ \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \\ x_{N1} & x_{N2} \end{bmatrix}_{N\times 2} \end{array} \begin{array}{l} 1\times 2 \\ \text{1st car} \\ \text{2nd car} \\ \\ N\text{th car} \end{array}$$

label vector

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_{N\times 1}$$



- family car
- other type of car

$x_1 = \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix}$ $\quad y_1 = 0$

$x_2 = \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix}$ $\quad y_2 = 1$

Model family

FAMILY OF RECTANGLES

$\theta = \{p1, p2, e1, e2\}$

model parameters

if $p_1^* \le x_{N+1,1} \le p_2^*$ & $e_1^* \le x_{N+1,2} \le e_2^*$

$\hat{y}_{N+1} = 1$

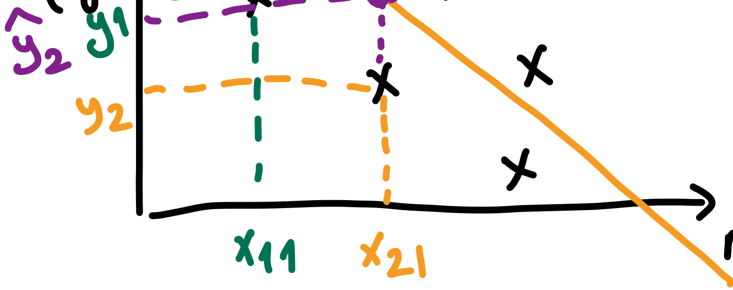
LEARNING $\Rightarrow$ finding the best θ

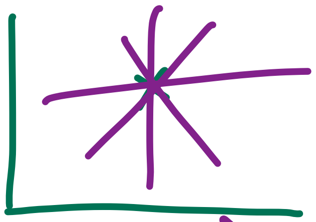
$\theta^* = \{p_1^*, p_2^*, e_1^*, e_2^*\}$

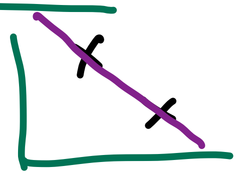
else $\hat{y}_{N+1} = 0$

$f(x_{N+1} \mid p_1^*, p_2^*, e_1^*, e_2^*) = \hat{y}_{N+1}$

Regression







data matrix $X = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{41} \\ x_{51} \\ x_{61} \end{bmatrix}$ 6×1

$N \times D$

N → # of data points

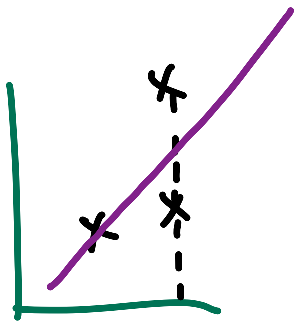
D → dimensionality, # of features

target/output vector $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix}$ 6×1

MODEL FAMILY ⇒ FAMILY OF LINES

$w_1 . x + w_0$

$$\hat{y}_1 = w_1^* x_{11} + w_0^*$$
$$\hat{y}_2 = w_1^* . x_{21} + w_0^*$$
$$\vdots$$
$$\hat{y}_6 = w_1^* . x_{61} + w_0^*$$



LEARNING ⇒ Finding the best $\theta^* = \{w_1^*, w_0^*\}$

observed, predicted/estimated

$$e_1 = y_1 - \hat{y}_1$$
$$e_2 = y_2 - \hat{y}_2$$
$$\vdots$$
$$e_6 = y_6 - \hat{y}_6$$

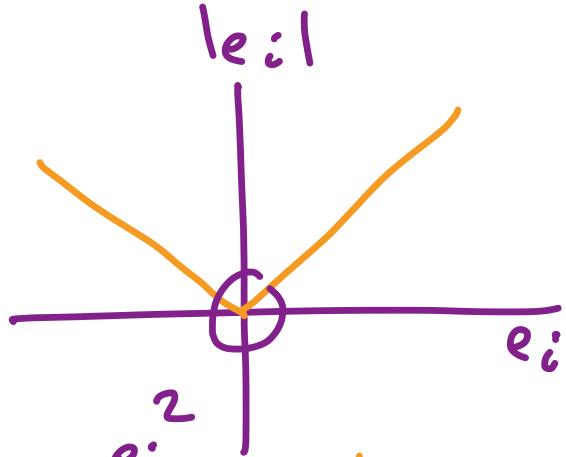
x_{N+1}

⇓

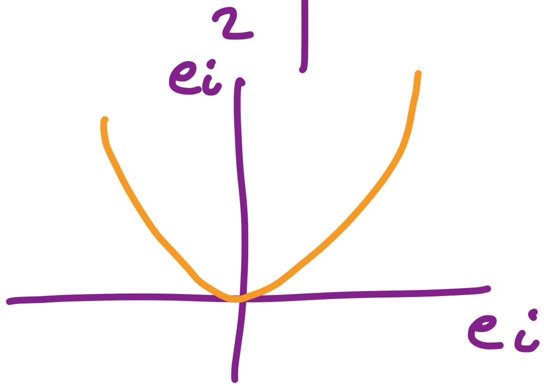
$$\hat{y}_{N+1} = w_1^* x_{N+1,1} + w_0^*$$

✗ minimize $\sum_{i=1}^{N} (y_i - \hat{y}_i) = \sum_{i=1}^{N} e_i$

✗ minimize $\sum_{i=1}^{N} |y_i - \hat{y}_i| = \sum_{i=1}^{N} |e_i|$



✓ minimize $\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{N} e_i^2$



minimize $\sum_{i=1}^{N} \left(y_i - (w_1 \cdot x_{i1} + w_0)\right)^2$

with respect to: w_1 and w_0.

$$Error(w_0, w_1 \mid \mathcal{X}) = \sum_{i=1}^{N} \left(y_i - (w_1 x_{i1} + w_0)\right)^2$$

$$\frac{\partial Error}{\partial w_0} = \sum_{i=1}^{N} \left[\frac{\partial \left(y_i - (w_1 x_{i1} + w_0)\right)^2}{\partial w_0}\right]$$

$$= \sum_{i=1}^{N} 2 \cdot \left(y_i - (w_1 x_{i1} + w_0)\right)(-1) = 0$$

$$\frac{\partial Error}{\partial w_1} = \sum_{i=1}^{N} 2 \cdot \left(y_i - (w_1 x_{i1} + w_0)\right) \cdot (-x_{i1}) = 0$$

Exercise

Solve for w_1^* and w_0^*

$$w_1^* = \frac{\sum_{i=1}^{N} (x_{i1} \cdot y_i) - \left(\sum_{i=1}^{N} x_{i1}/N\right)\left(\sum_{i=1}^{N} y_i/N\right) \cdot N}{\sum_{i=1}^{N} x_{i1}^2 - N\left[\sum_{i=1}^{N} x_{i1}/N\right]^2}$$

$$w_0^* = \underbrace{\left(\sum_{i=1}^{N} y_i/N\right)}_{\bar{y}} - w_1^* \cdot \underbrace{\left(\sum_{i=1}^{N} x_{i1}/N\right)}_{\bar{x}}$$