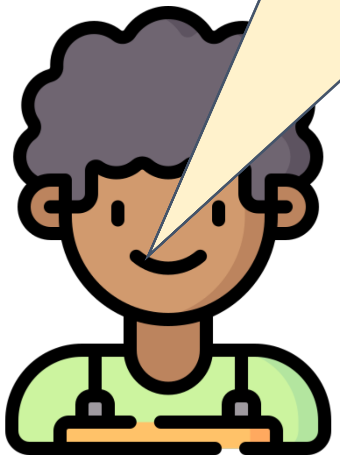
Takeaways Day 1

Before today

Hello, could you please fit a linear regression for me using gradient descent? Also use linear algebra to do it.
K bye thanks!



You: *terrified pear noises*

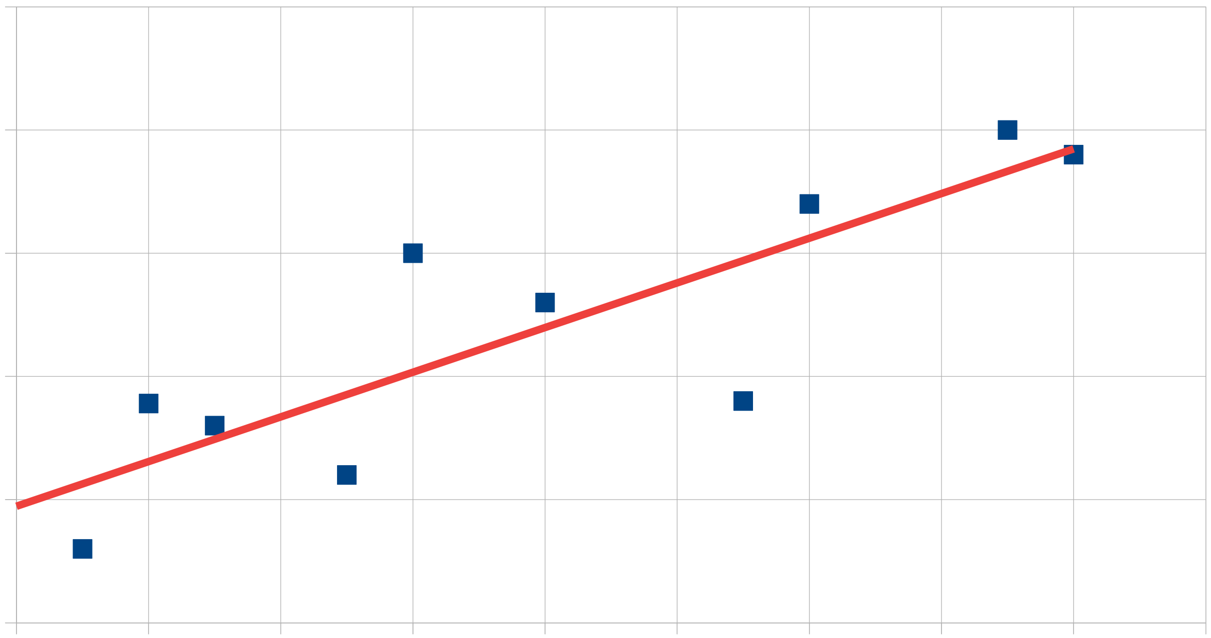
Now:

You: *somewhat understanding but still quite confused pear noises*

Needed

- How wrong are you?

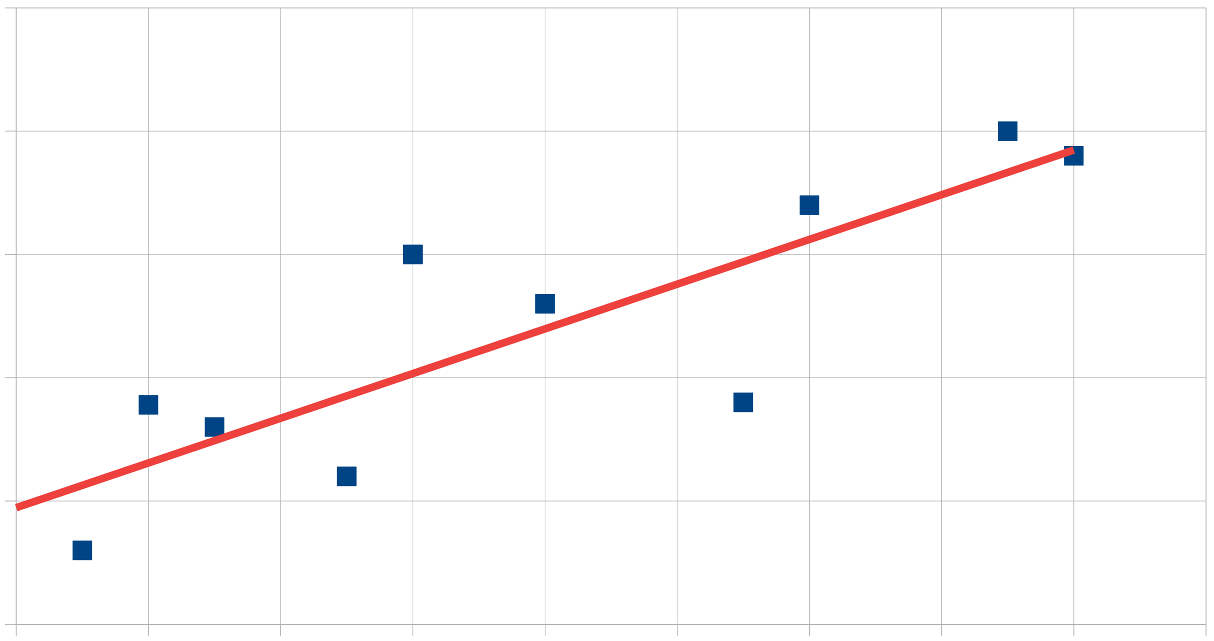
$$h_\theta(x) = \theta_0 + \theta_1 x$$



Needed

- How wrong are you?

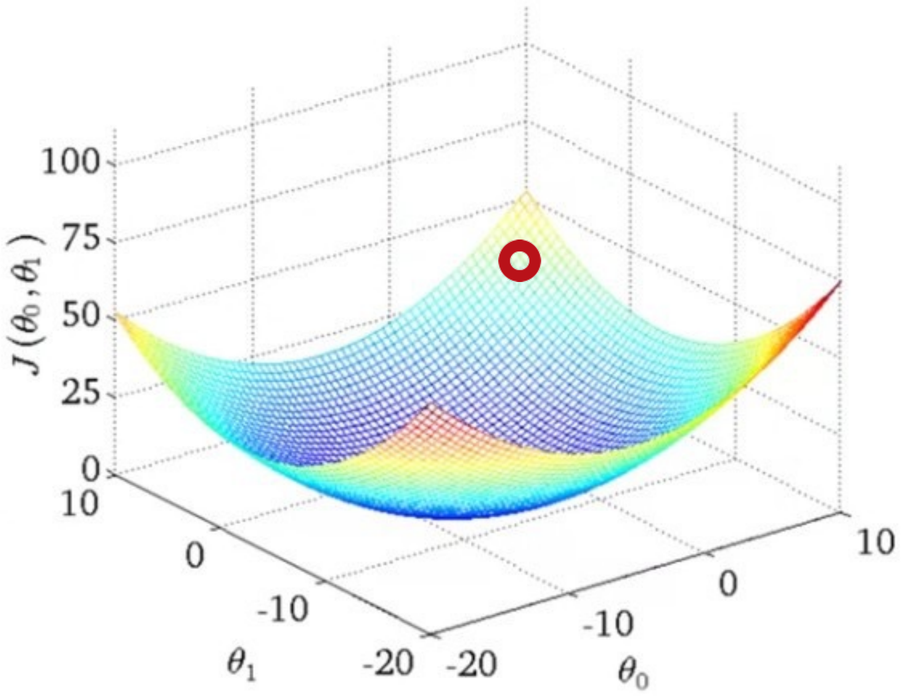
$$J(\theta_0, \theta_1) = \underbrace{\frac{1}{2m} \sum_{i}^{m} \underbrace{(\theta_0 + \theta_1 \cdot x^{(i)} - y^{(i)})^2}_{\text{Per sample}}}_{\text{On average}}$$



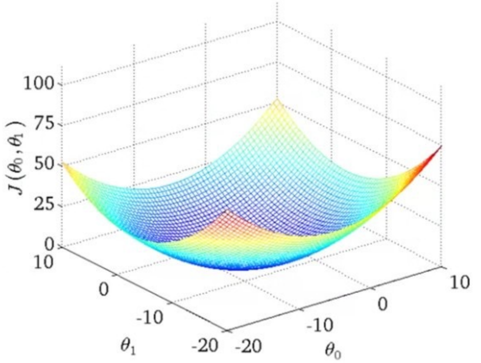
Needed

- Updating

$$J(\theta_0,\theta_1)=\frac{1}{2m}\sum_{i}^{m}(\theta_0+\theta_1\cdot x^{(i)}-y^{(i)})^2$$



Cost function and gradient descent



$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x) - y^{(i)})^2$$

$$\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})$$

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x^{(i)}) - y^{(i)}) \cdot x^{(i)})$$

Procedure

- Start with random theta values
- Calculate your hypothesis function: the linear regression prediction
- Calculate the errors
- Calculate the partial derivatives for each theta
- Take a small step

1	x1	x2	x3
1	0.5	0.2	0.8
1	0.9	-0.4	0.45

@

Theta_0
Theta_1
Theta_2
Theta_3

Errors =

pred_1
Pred_2
Pred_3

-

y_1
y_2
y_3

Take a step

$$\frac{\partial}{\partial\theta_0} J(\theta_0,\theta_1) = \frac{1}{m}\sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})$$

$$\frac{\partial}{\partial\theta_1} J(\theta_0,\theta_1) = \frac{1}{m}\sum_{i=1}^{m} ((h_\theta(x^{(i)}) - y^{(i)})\cdot x^{(i)})$$

Theta old – alpha * partial derivative = theta new