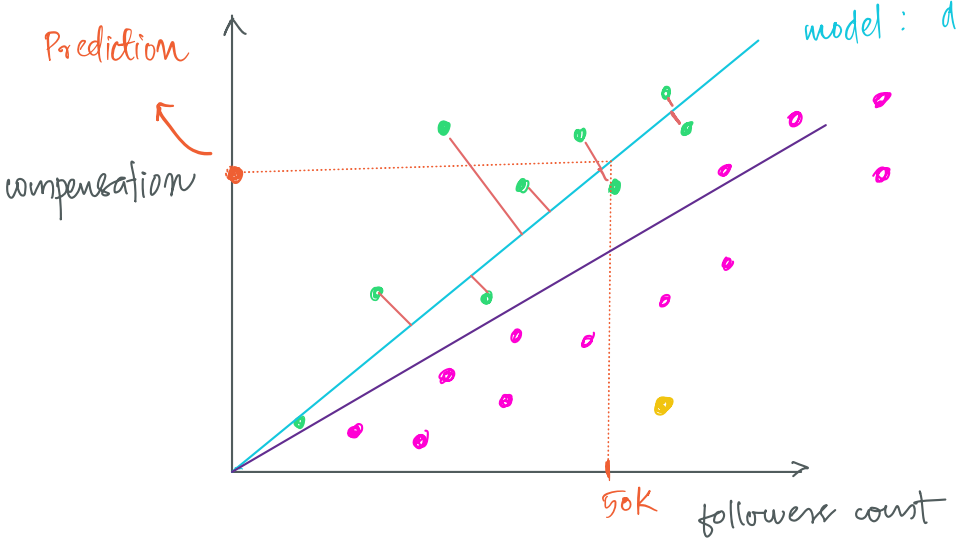
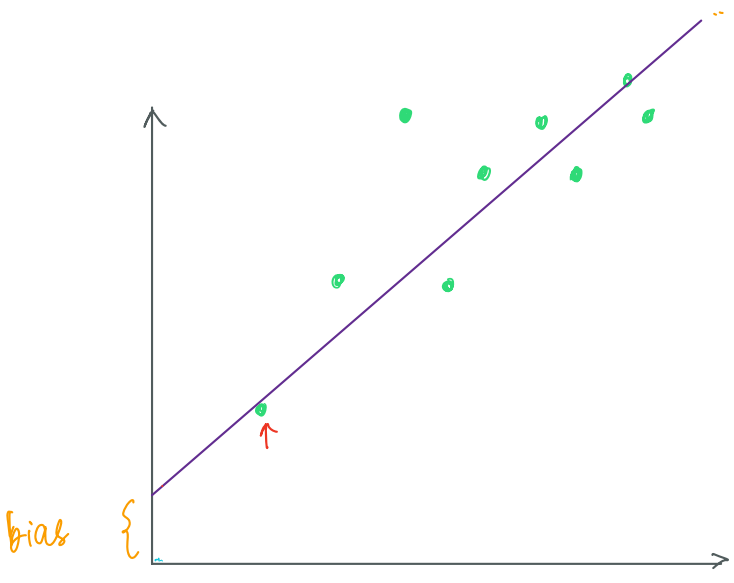
followers count → compensation for posting a brand's content



model : describe relationship btw x and y



$y = mx + c$

y → output, x → input, c → constant

$y = wx + b$

y → o/p, x → i/p, w → weight, b → bias

Linear Regression : $y = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b$

$n \rightarrow$ no of features

Loss

1. Sum of errors :

$$\sum_{(x,y) \in D} y - y'$$

y – observation

y' – prediction

2. Sum of absolute Error

$$\sum_{(x,y) \in D} |y - y'|$$

3. Mean of sum of absolute error (MAE)

$$L_1 \text{ Loss} = \frac{1}{N} \sum_{(x,y) \in D} |y - y'|$$

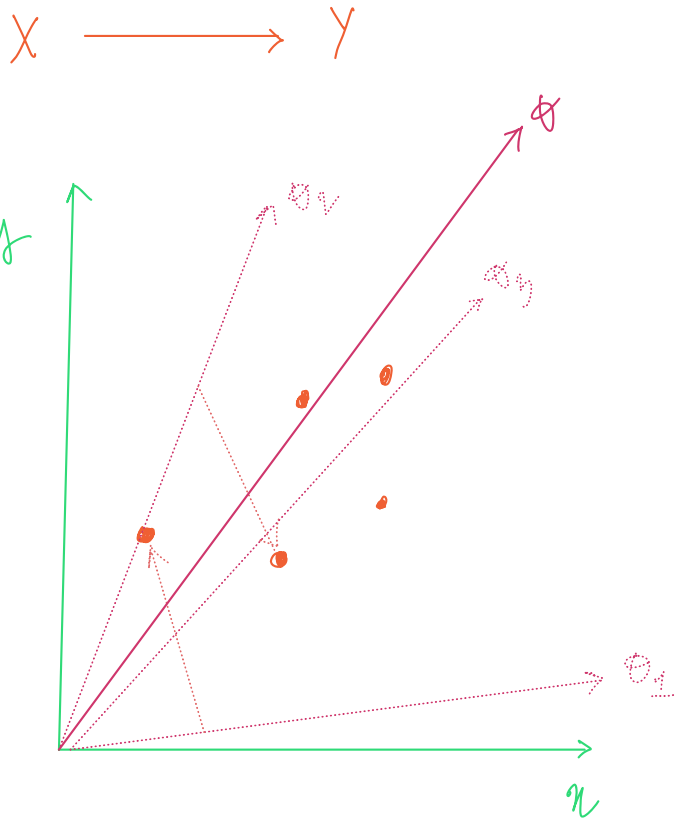
$N \rightarrow$ no. of examples

4. Sum of squared error

$$= \sum (y - y')^2$$

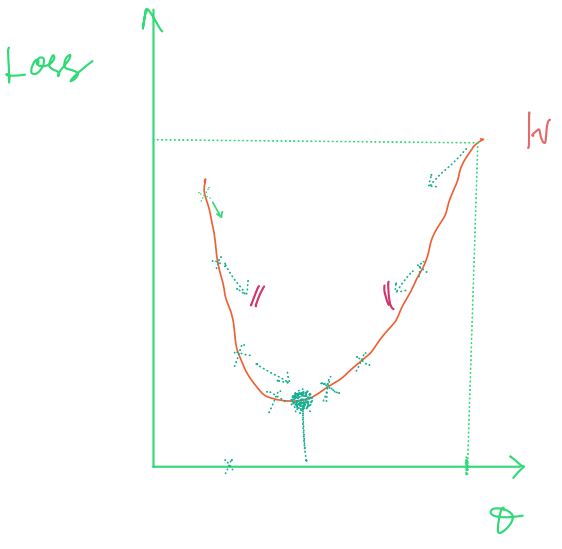
5. Mean of sum of squared error = MSE

$$L_2 \text{ Loss} \quad \frac{1}{N} \sum (y - y')^2$$

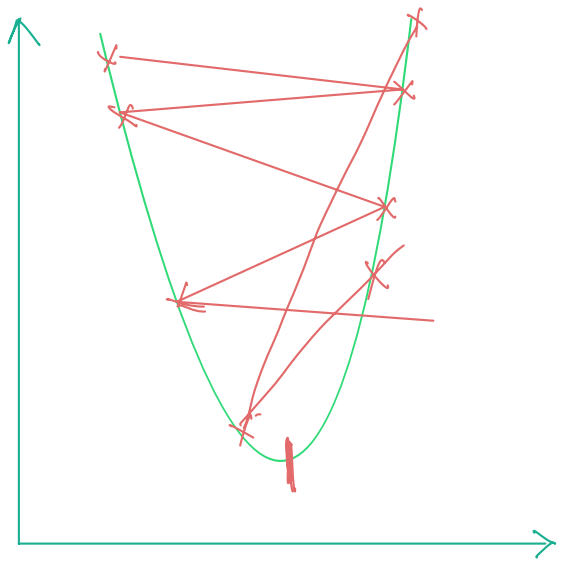


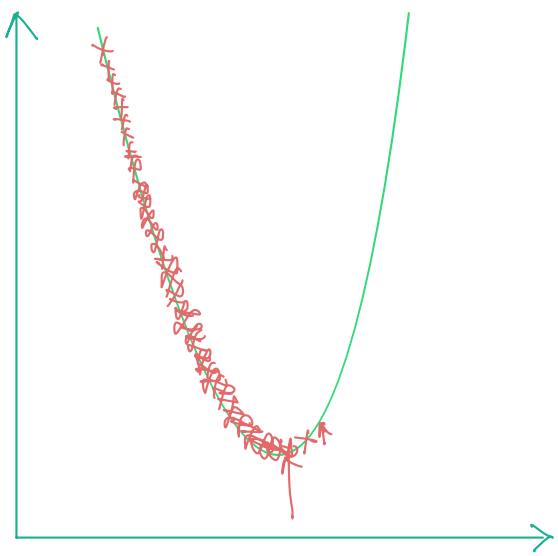
$y = wx + b$

$y = \theta x$ → hyperparameter



Learning Rate → Hyperparameter





Stochastic

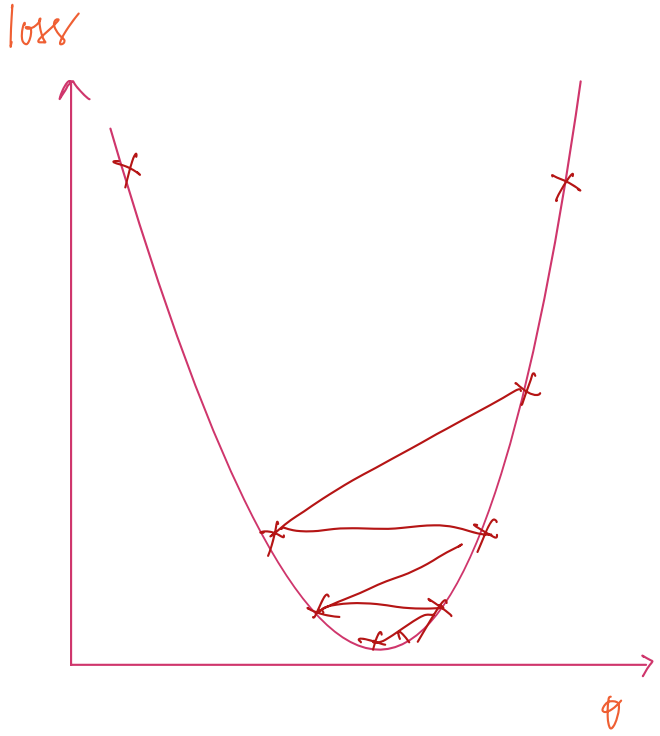
one example at a time

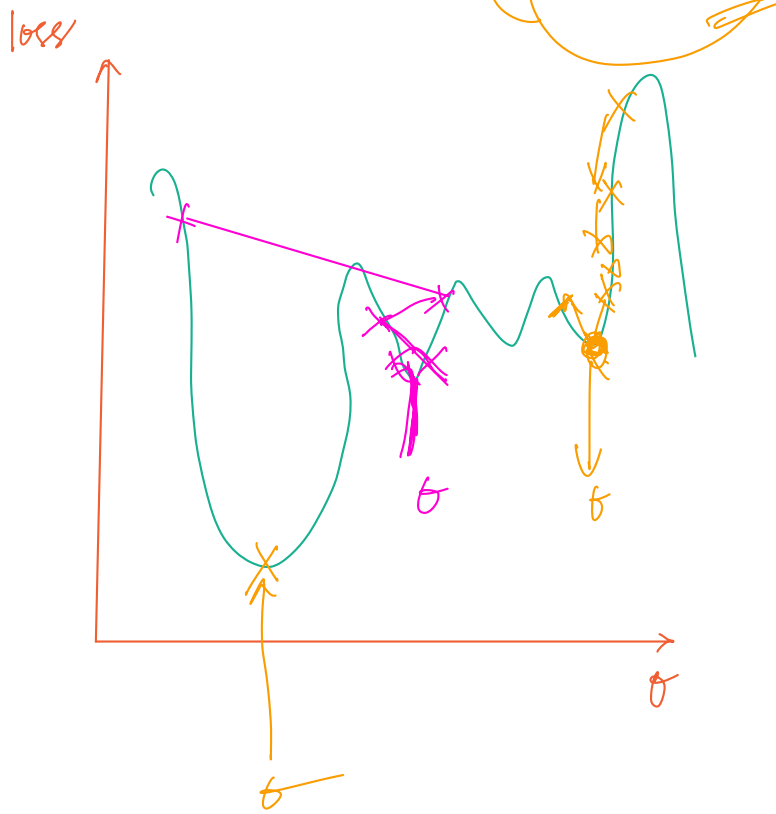
Mini Batch

10 - 100

Stochastic gradient descent

Does it matter where we start?





random seed = 42

Regularisation

Overfitting → High variance

↳ amount that the prediction will change if training data changes.

Regularisation

- Constraining the model complexity
  1. L1/L2 Regularisation → Add weights to cost calculation
  2. Dropout
  3. Early stopping
- Data Augmentation

L1 Regularisation

Rmbr

L1 loss — absolute errors

L2 loss — squared errors.

$$w = \text{minimise loss} + \text{Regularisation}$$

$$\text{minimise loss} + \lambda \times \text{Regulariser (Model)}$$

λ ↓ strength

Linear Regression:

$$y' = \sum_i w_i x_i$$

$$\text{Loss} = \sum (y_i - w_i x_i)^2$$

$$w = \arg\min \sum (y_i - w_i x_i)^2$$

LASSO

$$w^{L1} = \arg\min \sum (y - y')^2 + \lambda |w|$$

↳ absolute values of weights

pushes the weights closer to 0

L2 Regularisation

$$w^{L2} = \arg\min \sum (y - y')^2 + \frac{\lambda}{2} \|w\|^2$$

↓

Ridge Regression

Tuning λ (0 to 1)

too strong → underfit

too small → still overfits

Dropout → will discuss later

Early stopping → discussed earlier

Data augmentation :
create variations of data