27 May 2022 21:10

$f(x) = \underbrace{2}x^2 + \underbrace{3}x + \underbrace{5}$

Training Data → $\boxed{\text{Model}}$ → Model has learnt features

$\boxed{\text{Loss}}$

↳ How far the model predicts to some data to the truth.

↙ Optimization (Gradient Descent)

↓ Loss ↓

Model → linear (straight line)

$\underset{\text{output result}}{y} = m\underset{\text{input}}{x} + c$ | $y = m_1x_1 + m_2x_2 + m_3x_3 + \ldots + c$

House Price — Square footage of a house

↖ Linear Regression

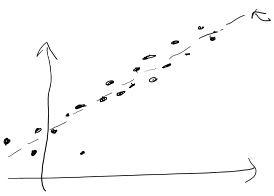
$\boxed{y = w^T X + b}$

$\begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Training Data:

House Price	Area
-	-
-	-
-	-

n — n is large



← line of best-fit $y = mx+c$

y_i ← predicted

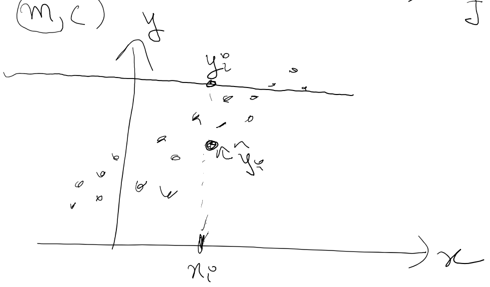
$\hat{y}_i$ ← truth

Loss fn

↳ $J = \sum |y_i - \hat{y}_i|$ ← absolute loss

↳ $J = \sum_i (y_i - \hat{y}_i)^2$ ← squared loss

(m, c)



Gradient Descent:

$$\theta_i = \theta_i - \alpha \frac{\partial J}{\partial \theta_i}$$

$\theta_i \to m, c$

hyperparameters

$\boxed{\alpha > 0}$

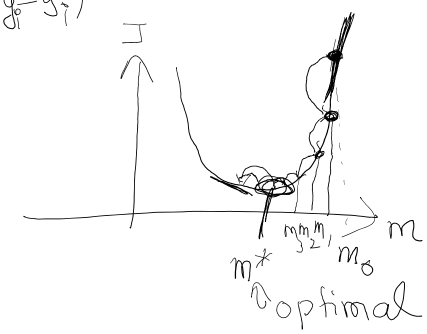
num_epochs (100 / 1000)

$$\boxed{m = m - \alpha \frac{\partial J}{\partial m}}$$

$$\boxed{c = c - \alpha \frac{\partial J}{\partial c}}$$

Line

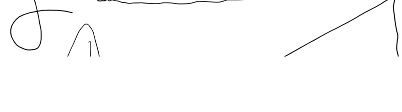
$$J = \sum_i (y_i - \hat{y}_i)^2$$



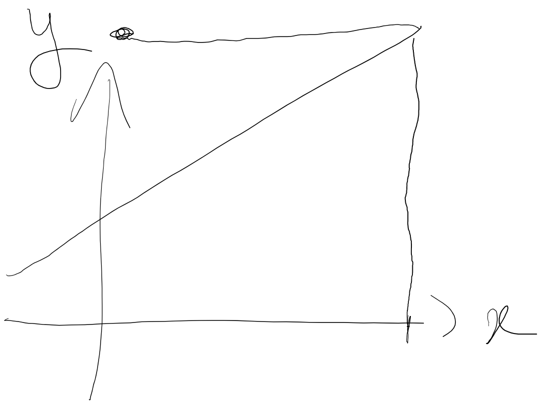
↑ optimal m

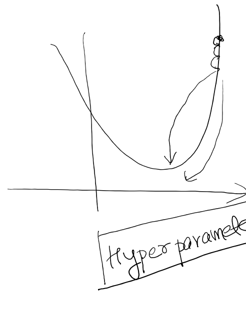
$\frac{\partial J}{\partial m}$ | c

$c_0 \to c_1 \to c_2 \longrightarrow c^*$



ar
Regression





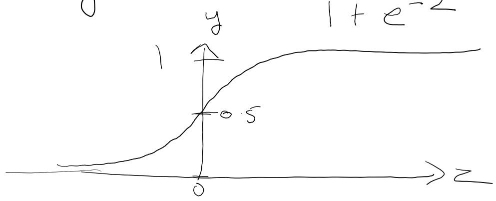
Classification → output in [0,1]

A B C

Logistic Regression

$z = mx + c$

$y = \text{sigmoid}(z) = \dfrac{1}{1+e^{-z}}$



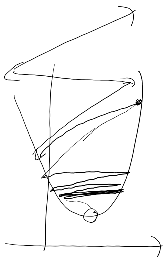
Categorical Cross-Entropy Loss

Loss → Binary Cross-Entropy Loss → 0 Yes Cat, → 1 No Dog

$J = \sum \hat{y}_i \log y_i + (1-\hat{y}_i)\log(1-y_i)$

$y_i \in [0,1]$ $\hat{y}_i \to 0, \to 1$

Gradient Descent $\dfrac{\partial J}{\partial m}$ $\dfrac{\partial J}{\partial c}$



$\boxed{\text{or}}$

$\alpha \rightarrow$ chosen by us