6_Optimization_in_Logistic_Regression

08 August 2024 10:50

Logistic Regression: -

(x, y) $\quad x \in \mathbb{R}^{n \times m}$ → $n \rightarrow$ # of features in each sample.

$y \in \{0, 1, \ldots, m\}$ $\quad m \rightarrow$ # samples.

$k =$ # classes.

$x.\text{shape} = (n, m)$

$y.\text{shape} = (1, m)$

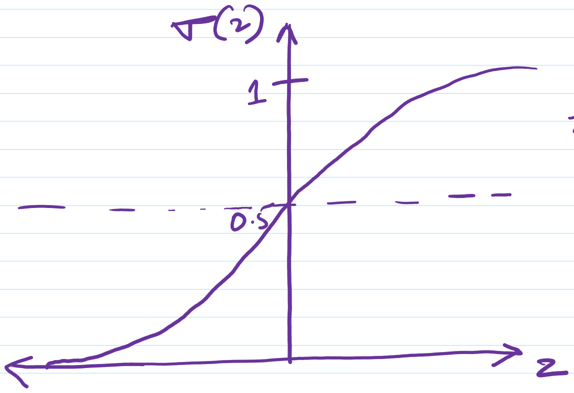
Let $z = w\,b + x$

$h = \sigma(z)$

$\sigma(z) = \dfrac{1}{1+e^{-z}}$

$\hat{y}$

0.5



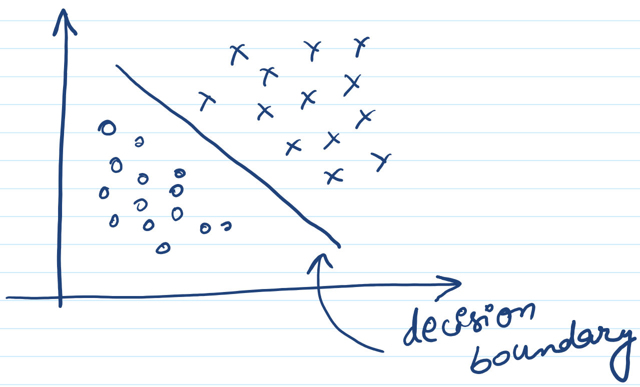
$z\uparrow \quad \sigma(z) = \dfrac{1}{1+0} = 1$

$z\downarrow \quad \sigma(z) = \dfrac{1}{1+\infty} = 0$

sigmoid function

$\hat{y} =$ "1" if $h \geq 0.5$

"$\hat{y} = 0$" if $h < 0.5$



$$h_\theta(x) = g(\theta_0 + \theta_1 x_2 + \theta_2 x_3)$$

$$= \theta^T x$$

where $\theta = [\theta_0 \quad \theta_1 \quad \theta_2]$ $\quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$\hat{y} = h_\theta(x) = \dfrac{1}{1+e^{-\theta^T x}}$, with thresholding

Cost Function :—

Binary cross-entropy loss.

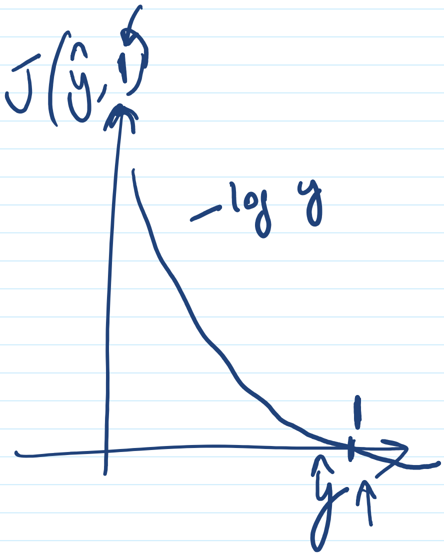
$$\boxed{J(\hat{y}, y) = -\left(\underline{y \log \hat{y}} + \underline{(1-y) \log(1-\hat{y})}\right)}$$

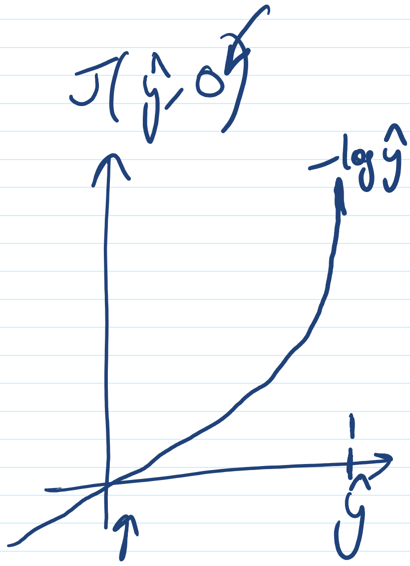
$\Rightarrow$ if $y = 1$

$$J(\hat{y}, y) = -\log \hat{y}$$

if $y = 0$

$$J(\hat{y}, y) = -\log(1-\hat{y})$$





Total cost,

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} J(\hat{y}_i, y_i)$$

$$J(\theta) = \frac{1}{m} \sum_{i=1} L(\hat{y}_i, y_i)$$

$$= -\frac{1}{m} \left[\sum_{i=1}^{m} \left(y_i \log \hat{y}_i + (1-y_i) \log (1-\hat{y}_i) \right) \right]$$

$$= -\frac{1}{m} \left[\sum_{i=1}^{m} y_i \log h_\theta(x_i) + (1-y_i) \log (1-h_\theta(y_i)) \right]$$

$$\min_{\theta} J(\theta)$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

Gradient Descent :-

Repeat {

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

}

Calculate $\frac{\partial}{\partial \theta_j} J(\theta)$