Logistic Regression

Given x, want →

$\hat{y} = P(y = 1 \mid x)$

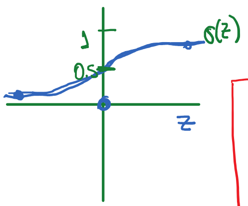
$0 \le \hat{y} \le 1x$ $\quad x \in \mathbb{R}^{n_x}$

Parameters: $\boxed{w} \in \mathbb{R}^{n_x}$, $\boxed{b} \in \mathbb{R}$

Output: $\hat{y} = \sigma(\underbrace{w^{T}x + b}_{z})$

$\hat{y} = \sigma(\underbrace{w_x^T + b}_{z})$

Sigmoid Function



$\sigma(z) = \frac{1}{2 + e^{-z}}$

$x_0 = 1$, $x \in \mathbb{R}^{n_x+1}$

$\hat{y} = \sigma(\theta^{T}x)$

$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{n_x} \end{bmatrix} \begin{array}{l} \}\, b \\ \}\, w \\ \\ \\ \end{array}$

if z is large → $\sigma(z) \approx \frac{1}{1+0} = 1$

if z is large negative number →

$\sigma(z) = \frac{1}{1 + e^{-z}} \approx \frac{2}{1 + B\text{ig Number}} \approx 0$

Sigmoid

Logistic Regression Cost Function

$\rightarrow \hat{y}^{(i)} = \sigma(w^T x^{(i)} + b)$, where $\sigma(z^{(i)}) = \frac{1}{1+e^{-z^{(i)}}}$

Given $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$, want $\rightarrow \hat{y}^{(i)} \approx y^{(i)}$

$z^{(i)} = w^T x^{(i)} + b$

$x^{(i)}$, $y^{(i)}$, $z^{(i)}$ — i-th example

Loss (error) function:

$\times$ ~~$\mathcal{L}(\hat{y}, y) = \frac{1}{2}(\hat{y} - y)^2$~~

$$\mathcal{L}(\hat{y}, y) = -\left(y \log(\hat{y}) + (1-y)\log(1-\hat{y})\right)$$

if $y = 1$: $\mathcal{L}(\hat{y}, y) = -\log \hat{y} \leftarrow$ want $\log \hat{y}$ large, want $\hat{y}$ large

if $y = 0$: $\mathcal{L}(\hat{y}, y) = -\log(1-\hat{y}) \leftarrow$ want $1-\hat{y}$ large ... want $\hat{y}$ small

Cost function: $J(w,b) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(\hat{y}^{(i)}, y^{(i)})$

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