Hypothesis Representation

Classification and Representation

Logistic Regression

Logistic Regression Model

Want $0 \leq h_\theta(x) \leq 1$

Logistic Regression Model

Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = \quad \theta^T x$$

Logistic Regression Model

Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = g(\theta^T x)$$

Logistic Regression Model

Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = g(\theta^T x)$$

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

Logistic Regression Model

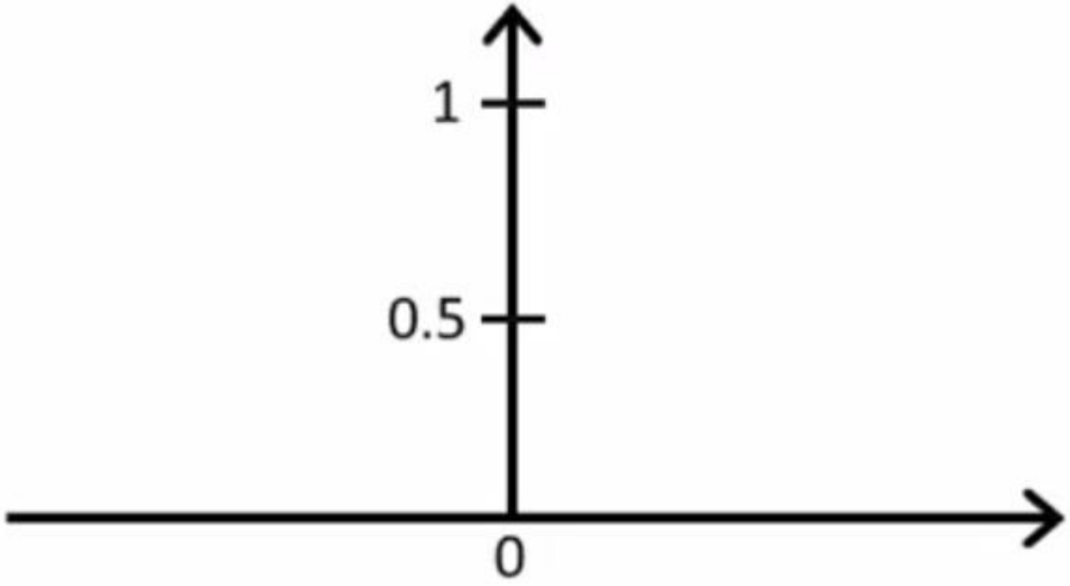
Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = g(\theta^T x)$$

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

Sigmoid function

Logistic function



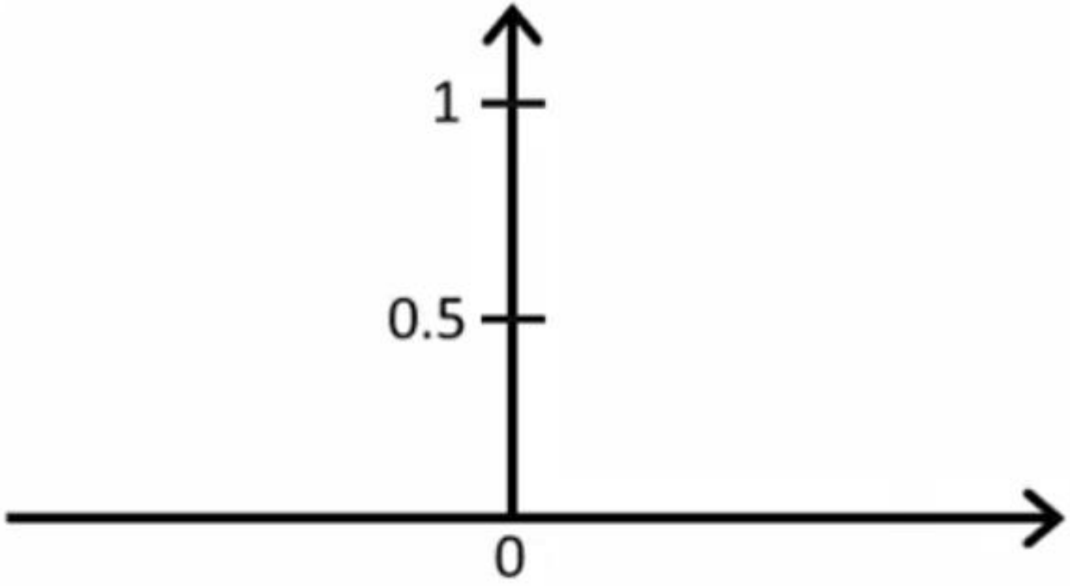
Logistic Regression Model

Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = g(\theta^T x)$$

$$\rightarrow g(z) = \frac{1}{1+e^{-z}}$$

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$



→ Sigmoid function

↳ Logistic function

Logistic Regression Model

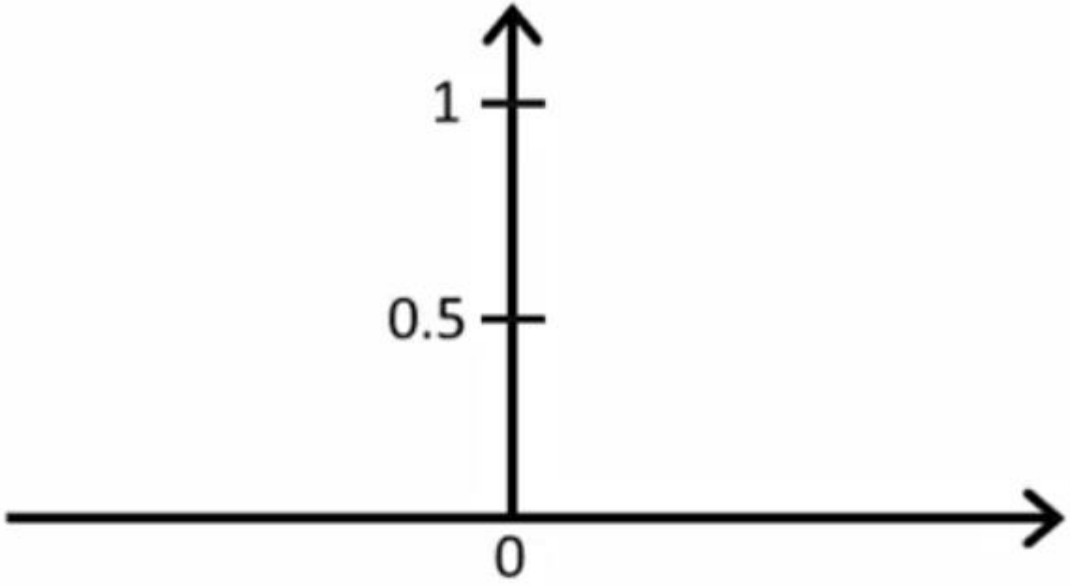
Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = g(\theta^T x)$$

$$\rightarrow g(z) = \frac{1}{1+e^{-z}}$$

$\theta^T x$

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$



→ Sigmoid function

→ Logistic function

Logistic Regression Model

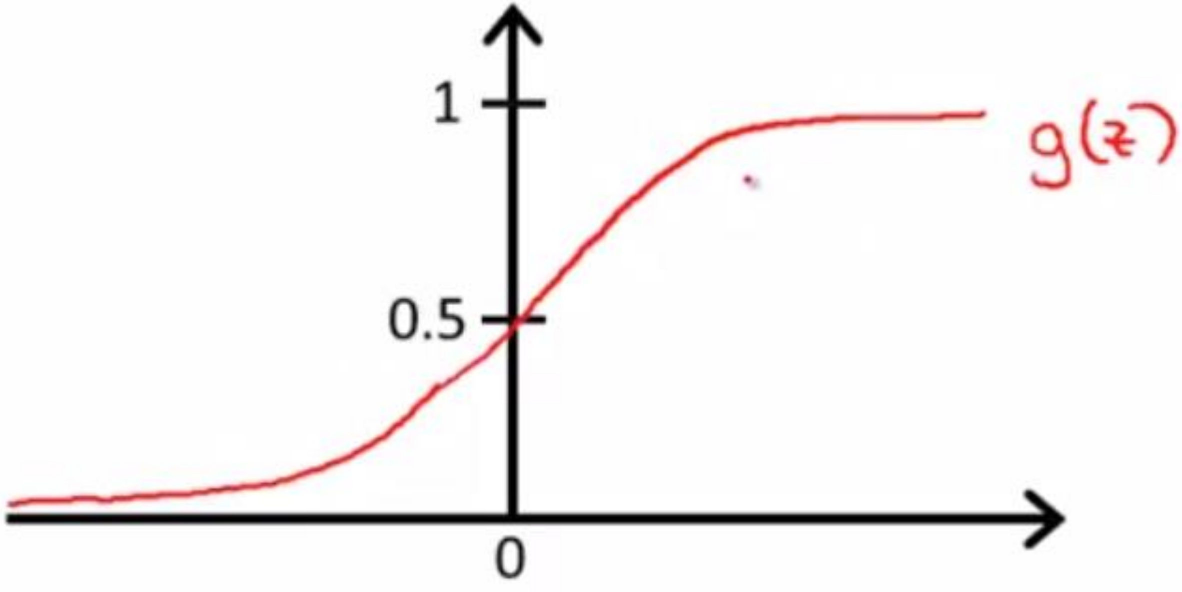
Want $0 \le h_\theta(x) \le 1$

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

$$h_\theta(x) = g(\theta^T x)$$

$$\rightarrow g(z) = \frac{1}{1+e^{-z}}$$

$\theta^T x$



→ Sigmoid function

→ Logistic function

Logistic Regression Model

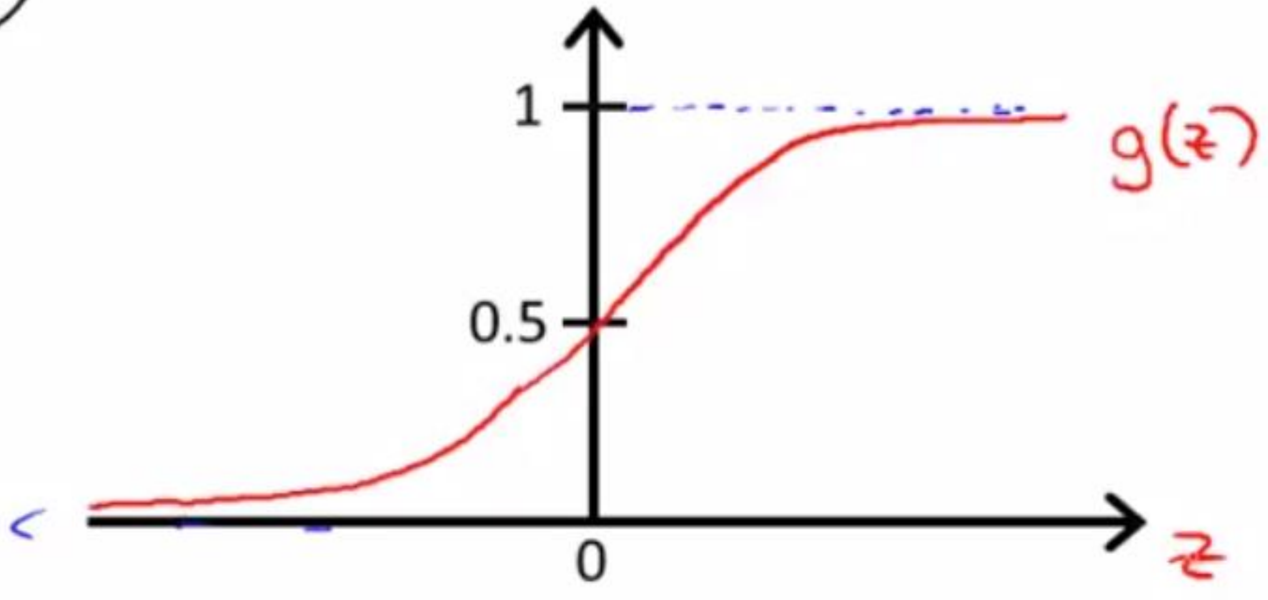
Want $0 \leq h_\theta(x) \leq 1$

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

$$h_\theta(x) = g(\theta^T x)$$

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

$\theta^T x$



Sigmoid function

Logistic function

Interpretation of Hypothesis Output

$h_\theta(x)$ = estimated probability that y = 1 on input x

$h_\theta(x)$

Interpretation of Hypothesis Output

$h_\theta(x)$ = estimated probability that $\boxed{y = 1}$ on input x

Example: If $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$

$$h_\theta(x) = 0.7$$

$h_\theta(x)$

Interpretation of Hypothesis Output

$h_\theta(x)$ = estimated probability that $\boxed{y = 1}$ on input x

Example: If $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$

$$h_\theta(x) = 0.7$$

Tell patient that 70% chance of tumor being malignant

Interpretation of Hypothesis Output

$h_\theta(x)$

$h_\theta(x)$ = estimated probability that $\boxed{y = 1}$ on input x

Example: If $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$

$$h_\theta(x) = 0.7$$

$y=1$

Tell patient that 70% chance of tumor being malignant

$h_\theta(x) = P(y=1 \mid x ; \theta)$

"probability that y = 1, given x, parameterized by θ"

Exercise

- Suppose we want to predict, from data *x* about a tumor, whether it is malignant (*y*=1) or benign (*y*=0). Our logistic regression classifier outputs, for a specific tumor, $h_\theta(x) = P(y = 1|x; \theta) = 0.7$, so we estimate that there is a 70% chance of this tumor being malignant. What should be our estimate for the probability the tumor is benign?
- 0.7^2
- 0.7-0.3
- 0.7-0.5
- 0.3

Interpretation of Hypothesis Output

$h_\theta(x)$ = estimated probability that $\boxed{y = 1}$ on input x

$h_\theta(x)$

Example: If $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$

$h_\theta(x) = 0.7$

$y = 1$

Tell patient that 70% chance of tumor being malignant

$h_\theta(x) = P(y=1|x;\theta)$

"probability that y = 1, given x, parameterized by θ"

$y = 0$ or 1

$\rightarrow P(y = 0|x;\theta) + P(y = 1|x;\theta) = 1$

$P(y = 0|x;\theta) = 1 - P(y = 1|x;\theta)$

Interpretation of Hypothesis Output

$h_\theta(x)$

$h_\theta(x)$ = estimated probability that $\boxed{\text{y = 1}}$ on input x

Example: If $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$

$h_\theta(x) = 0.7$ y=1

Tell patient that 70% chance of tumor being malignant

$h_\theta(x) = P(y=1|x;\theta)$

"probability that y = 1, given x, parameterized by θ"

y = 0 or 1

$$\rightarrow P(y=0|\cancel{x;}\theta) + P(y=1|\cancel{x;}\theta) = 1$$

$$P(y=0|x;\theta) = 1 - P(y=1|x;\theta)$$