WEEK 6
Logistic Regression

\[ h_\theta(x) = \theta^T x \]

Threshold classifier output \( h_\theta(x) \) at 0.5:

If \( h_\theta(x) \geq 0.5 \), predict “y = 1”

If \( h_\theta(x) < 0.5 \), predict “y = 0”
Logistic Regression

Threshold classifier output $h_\theta(x)$ at 0.5:

If $h_\theta(x) \geq 0.5$, predict “y = 1”

If $h_\theta(x) < 0.5$, predict “y = 0”
Logistic Regression

Classification: \( y = 0 \) or \( 1 \)

\[ h_\theta(x) \] can be \( > 1 \) or \(< 0 \)

Logistic Regression: \( 0 \leq h_\theta(x) \leq 1 \)
(a classification algorithm)
Logistic Regression

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

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

Mean the same thing

Andrew Ng
Logistic Regression

Interpretation of Hypothesis Output

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

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

\[ h_\theta(x) = 0.7 \]

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 \( \theta \)"

\[ P(y = 0|x; \theta) + P(y = 1|x; \theta) = 1 \]

\[ P(y = 0|x; \theta) = 1 - P(y = 1|x; \theta) \]
Logistic Regression

Decision Boundary

\[ h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2) \]

Predict "y = 1" if \[-3 + x_1 + x_2 \geq 0\]
Logistic Regression

Non-linear decision boundaries

\[ h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2) \]

Predict "y = 1" if \(-1 + x_1^2 + x_2^2 \geq 0\)
Logistic Regression

Training set: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})\}$

m examples $\quad x \in \begin{bmatrix} x_0 \\ x_1 \\ \ldots \\ x_n \end{bmatrix} \quad x_0 = 1, y \in \{0, 1\}$

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

How to choose parameters $\theta$?
Logistic Regression

Cost function

Linear regression:
\[ J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \frac{1}{2} (h_\theta(x^{(i)}) - y^{(i)})^2 \]

Cost:
\[ \text{Cost}(h_\theta(x^{(i)}), y^{(i)}) = \frac{1}{2} (h_\theta(x^{(i)}) - y^{(i)})^2 \]
Logistic Regression

Logistic regression cost function

\[
J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \text{Cost}(h_\theta(x^{(i)}), y^{(i)})
\]

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

To fit parameters \( \theta \):

\[
\min_{\theta} J(\theta)
\]

To make a prediction given new \( x \):

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

Gradient Descent

\[
J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_\theta(x^{(i)})) \right]
\]

Want \( \min_\theta J(\theta) \):

Repeat \{ \\
\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \\
\} \quad \text{(simultaneously update all} \ \theta_j)
Logistic Regression

Multiclass classification

Email foldering/tagging: Work, Friends, Family, Hobby
y=1  y=2  y=3  y=4

Medical diagrams: Not ill, Cold, Flu
y=1  y=2  y=3

Weather: Sunny, Cloudy, Rain, Snow
y=1  y=2  y=3  y=4
Logistic Regression

Binary classification:

Multi-class classification:
Logistic Regression

One-vs-all (one-vs-rest):

\[ h^{(i)}_{\theta}(x) = P(y = i | x; \theta) \quad (i = 1, 2, 3) \]
Consider a logistic regression model with weights $\beta = (-\ln(4), \ln(2), -\ln(3))$. A given document has feature vector $x = (1, 1, 1)$. Now, please provide your answer in the form of a fraction $\frac{a}{b}$.

What is the probability that the document is about sports?

Consider a logistic regression model with weights $\beta = (0.5, 0.25, 1)$. A given document has feature vector $x = (1, 0, 1)$. NOTE: for this problem you will be exponentiating certain quantities. You do not need to write out your answer as a number, but instead in terms of $\exp()$ values, e.g., $P = 1 + 2\exp(-1)$.

What is the probability that the document is about sports?

What is the probability that it is not about sports?