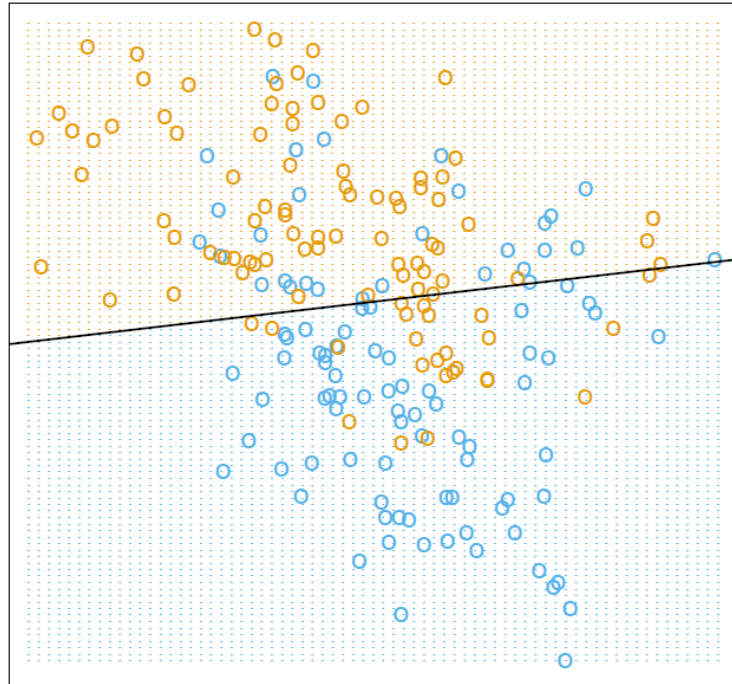


Linear Classifiers

Linear Regression of 0/1 Response



Some slides from:
Andrew Ng

CSC411: Machine Learning and Data Mining, Winter 2017

Michael Guerzhoy
1

Classification vs. Regression

- Classification: for the example (x_1, x_2, \dots, x_n) predict the label y (e.g., face recognition)
- Regression: for the example (x_1, x_2, \dots, x_n) predict a real number y (e.g., house price prediction)

Classification with two classes

- If there are only two classes, transform, e.g.,
orange => 1
blue => 0
to turn the classification problem into a regression problem

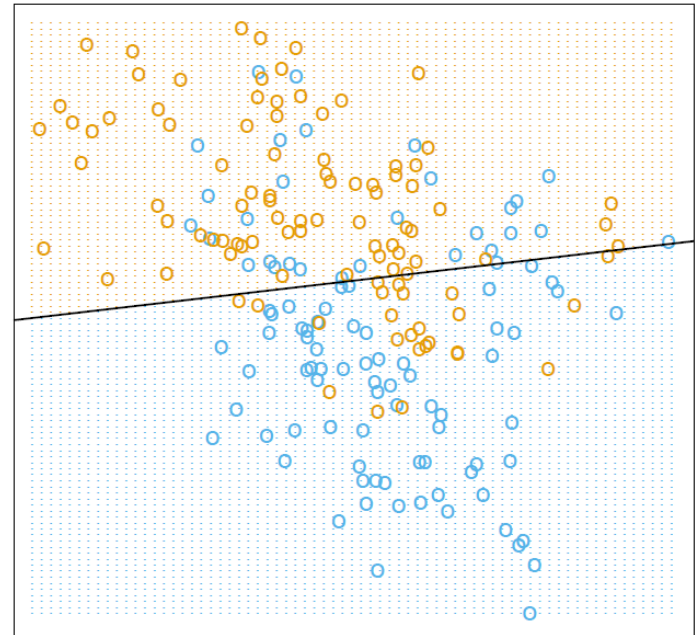
- Find the best

$$h_{\theta}(x) = \theta^T x$$

- Predict:

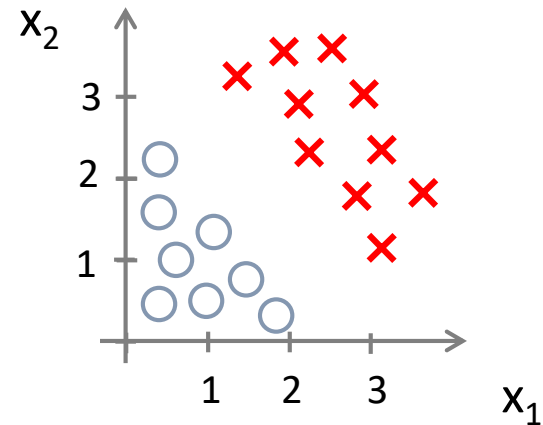
$$\begin{cases} 1, & h_{\theta}(x) > .5 \\ 0, & \text{otherwise} \end{cases}$$

Linear Regression of 0/1 Response



What is the equation of the decision boundary?

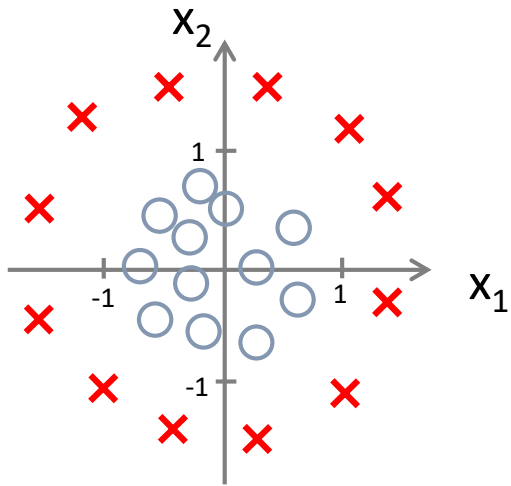
Decision boundary shapes



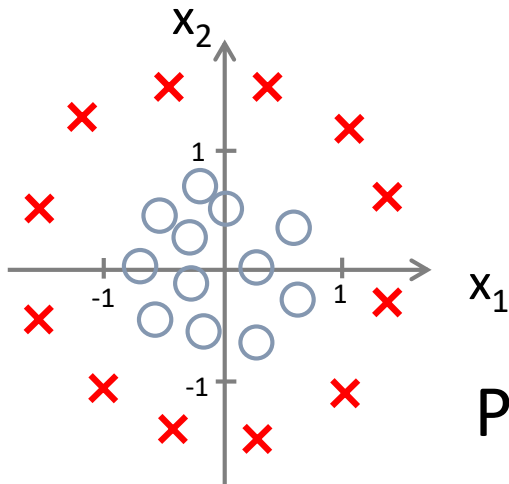
$$h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

Predict $y = 1$ if $-3 + x_1 + x_2 \geq 0$

Decision boundary shapes



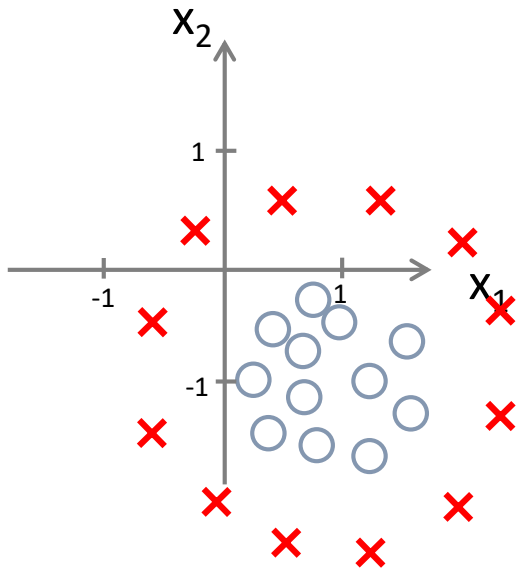
Decision boundary shapes



$$h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$$

Predict $y = 1$ if $-1 + x_1^2 + x_2^2 \geq 0$

What is the equation for a good decision boundary?



Multiclass Classification

Email foldering/tagging : Work, Friends, Family, Hobby

$y = 1$ $y = 2$ $y = 3$ $y = 4$

Features: x_1 : 1 if “extension” is in the email, 0 otherwise

x_2 : 1 if “dog” is in the email, 0 otherwise

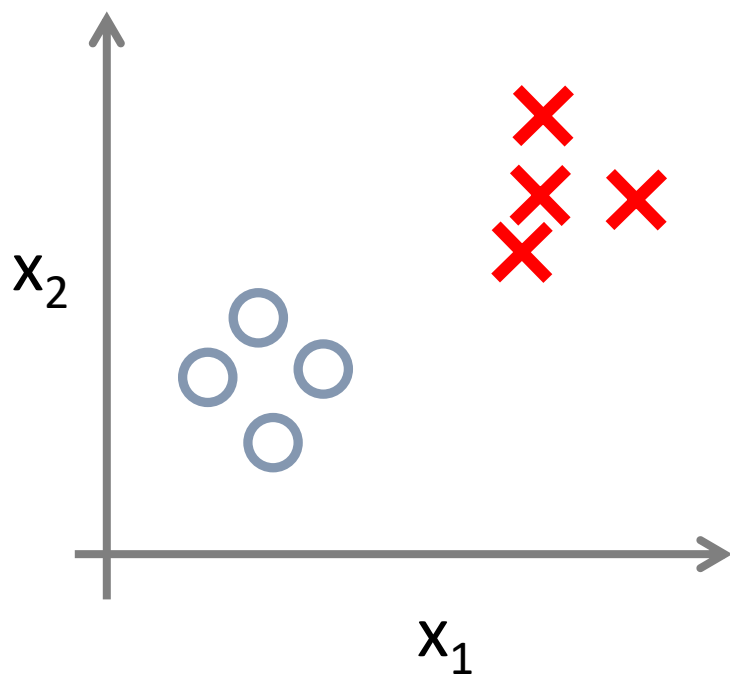
...

Medical diagrams: Not ill, Cold, Flu

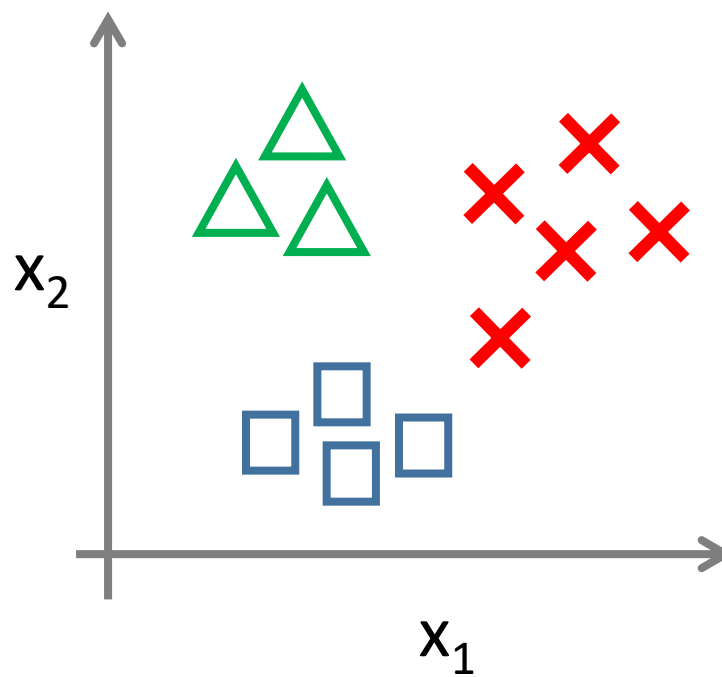
$y = 1$ $y = 2$ $y = 3$

Features: temperature, cough presence, ...

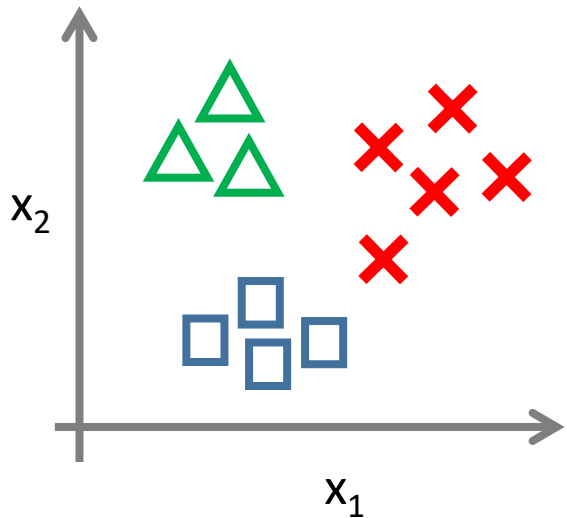
Binary classification:





Multi-class classification:




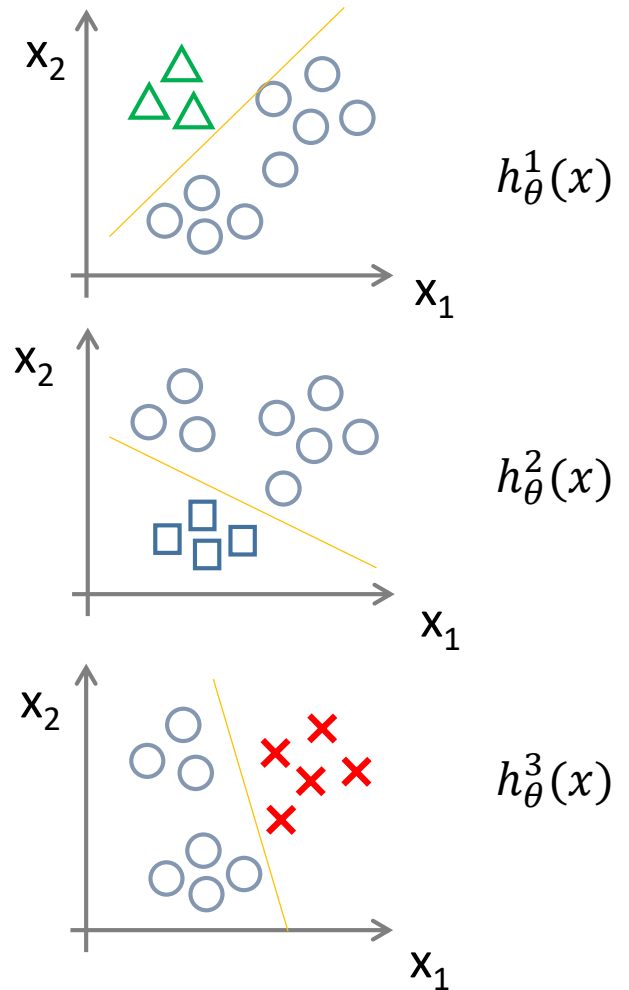
One-vs-all (one-vs-rest):



Class 1: 

Class 2: 

Class 3: 



Output the i such that $h_{\theta}^i(x)$ is the largest
(Idea: a large $h_{\theta}^i(x)$ means that the classifier is “sure”)