# Classification

Subcategory of supervised learning:

* Goal to predict category or class of new data instances based on an algorithm learned from training examples

Binary classification:

* Classifying instances into one or two classes (eg spam or not spam)

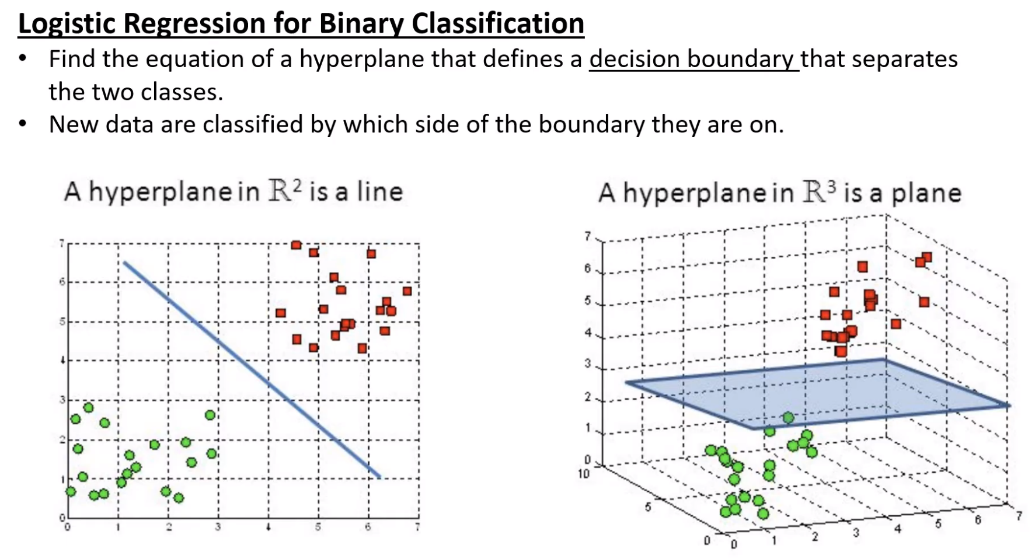
Multiclass classification:

* Three or more classes (facial recognition)

Many types:

* Logistic regression, k-nearest neighbor, naïve bayes classification, support vector machines, decision trees, deep learning

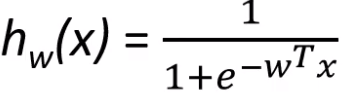
## Logistic regression for binary classification



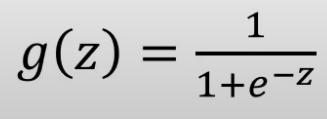
Find equation of hyperplane that defines a decision boundary that separates two classes

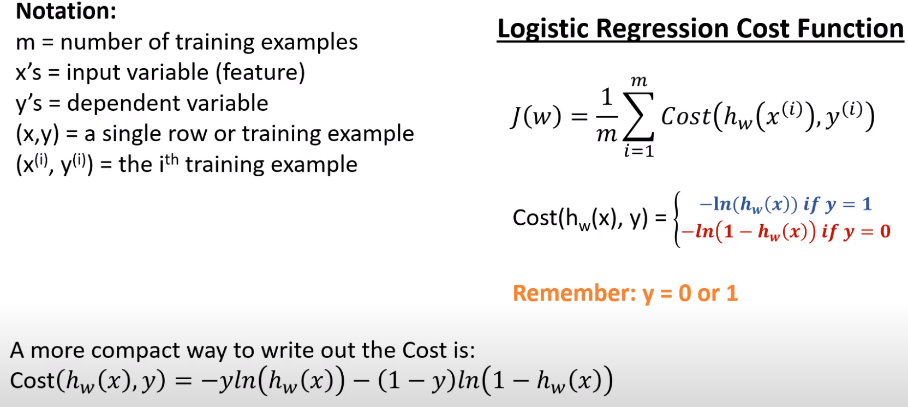
* New data classified by which side of the boundary they are on

Hypothesis function:

* 
* Below 0.5, class 0, above, class 1

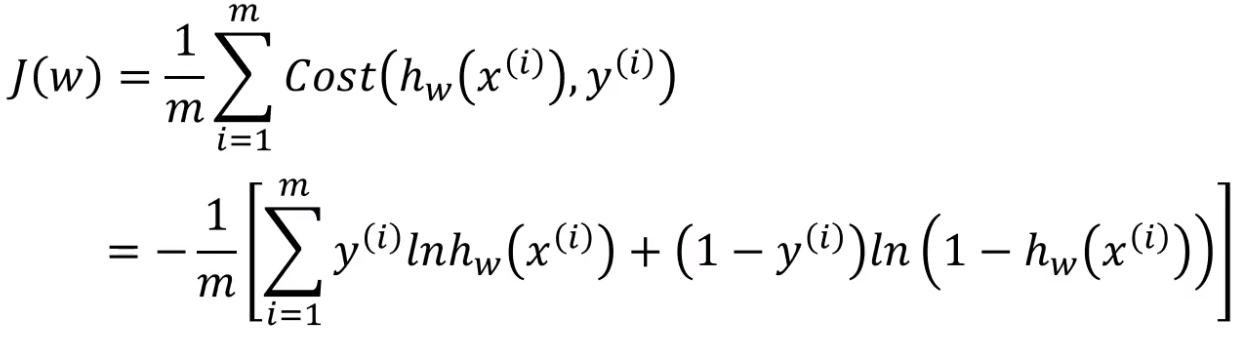
Logistic function (or sigmoid):

* 

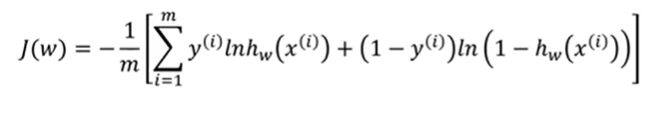
Interpretation:

* Given problem with classes, y, equal to either 0 or 1, hw(x) is probability of class x given input of x
* wTx defines decision boundary

Cost function

* J(w)
* Will use gradient descent to find Ws that minize J(w)
* Squared error not good choice, combined with g(z) it is non convex and hard to minimize – getting trapped in local minimum
* 

## Gradient Descent



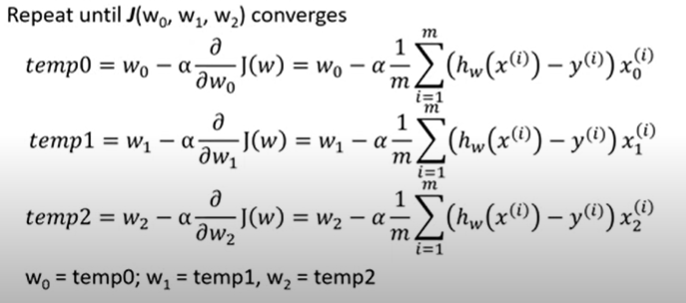
We want to minimize J(w)

wj = wj - α wj’ J(w)

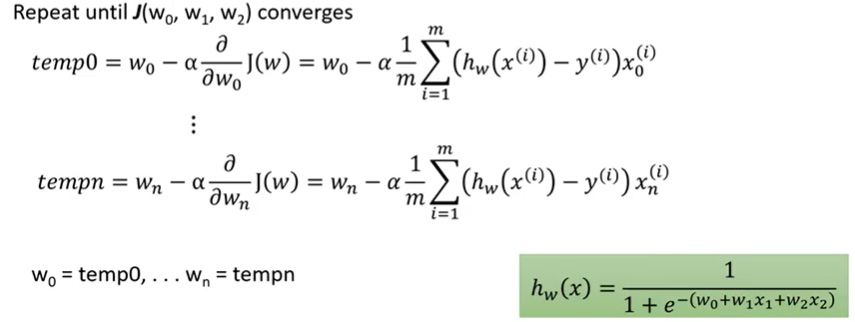
Full equation using chain rule:

* 

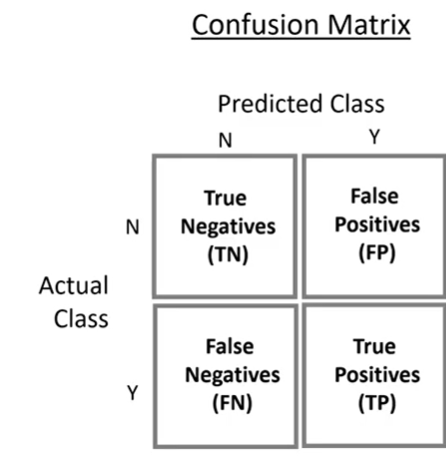
With two features:

* 
* Might look identical to regression, but hypothesis function still way different

With n features

* 

Everything we learned about α in linear regression applies to logistic regression



# Evaluating a Binary Classifier

Example:

* Suppose we have a test that predicts if person will get leukemia, apply test to 1,000,000
* Test gives accurate answer 98% of the time
* Is accuracy a good measure?
* Four possible outcomes:
  + True Positives (TP)
  + True Negative (TN)
  + False Positives (FP) // Type 1 Error
  + False Negatives (FN) // Type 2 Error
* Questions we want to answer:
  + Accuracy
    - How often is classifier correct?
  + Precision
    - When it predicts yes, how often is it correct?
    - TP / TP + FP
  + Recall
    - What percentage of positives instances are correctly predicted?
    - TP / TP + FN
* Merely using “Is named Luke” as the test results in:
  + Accuracy of of 98%!
  + Precision of 1.4%
  + Recall of 0.5%

F1 Score:

* Tradeoff between precision and recall
* Measure of how often classifier is correct and if classifier is correctly predicting positive cases
* Range is [0, 1] where larger is better
* 2 \* 1 / (1/P) + (1/Recall)