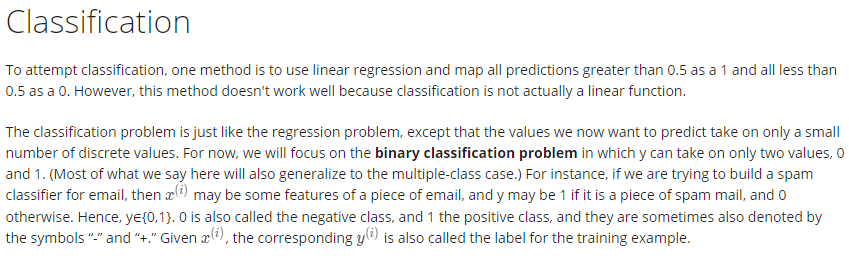
# Logistic Regression

## Classification and Representation

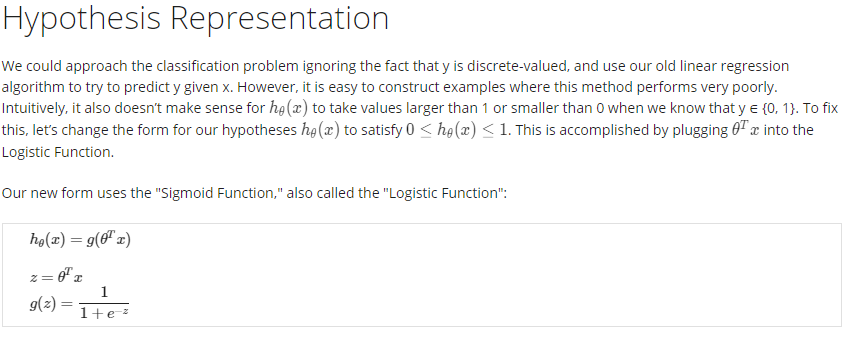
Classification

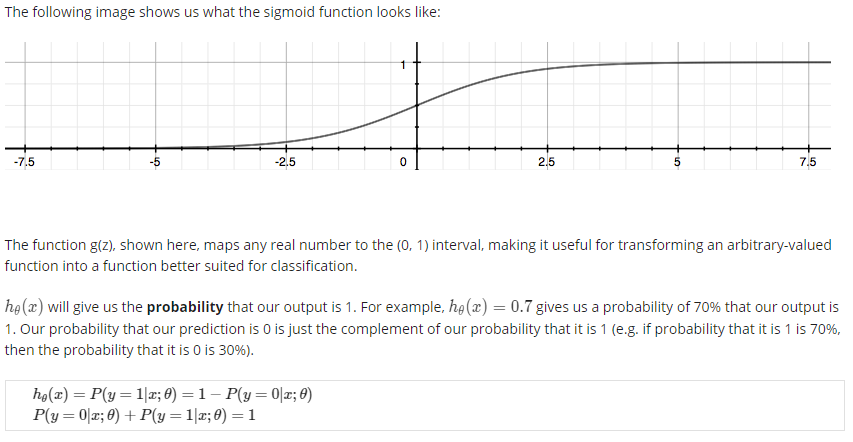
* e.g.: spam email; fraudulent transactions; malignant/benign tumors
* Focus on binary
* Threshold classifier output, evaluate h
* h can be > 1 or < 0 even for y = 0 or 1
* Logistic Regression
* 0 <= h <= 1
* Really classification algorithm; named due to historical reasons



Hypothesis Representation

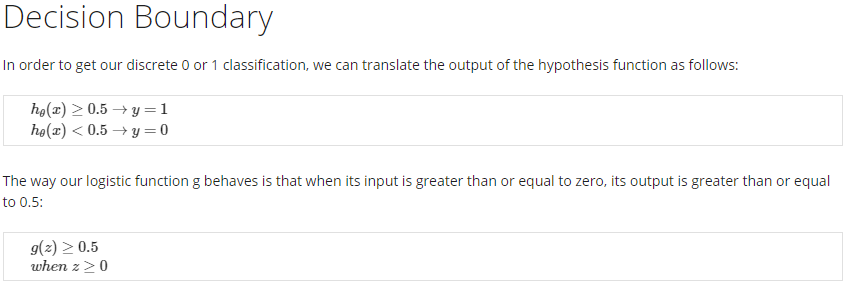
* Logistic regression model
* Sigmoid/Logistic function
* h(x) = θTx 🡪 g(θTx)
* h(x) is the estimated probability that y = 1 on input x
* h(x) = P(y = 1|x; θ)

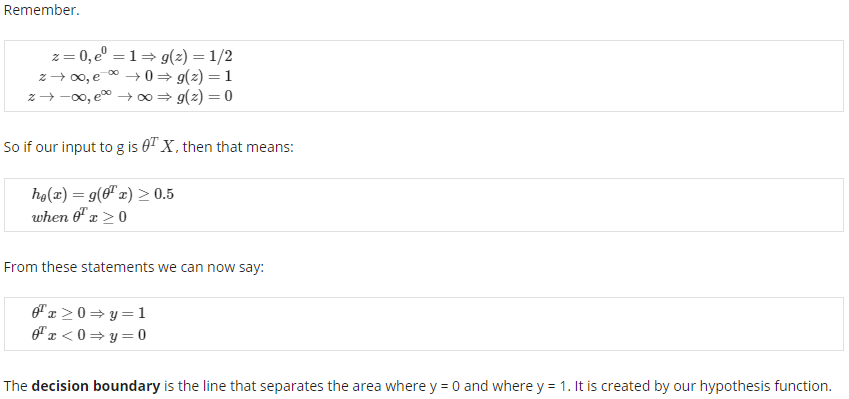


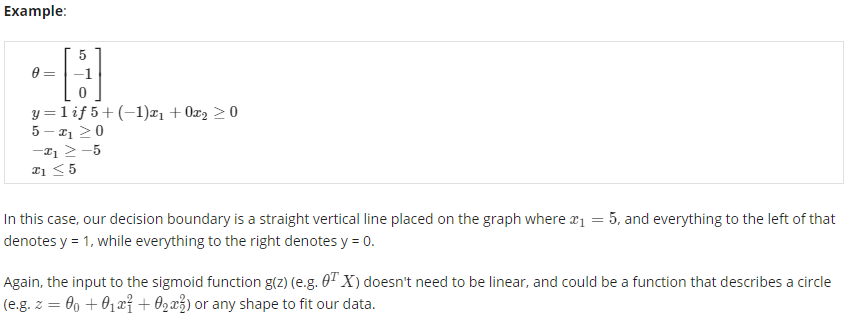


Decision Boundary

* independent of dataset for now; later on: how to fit parameters using data



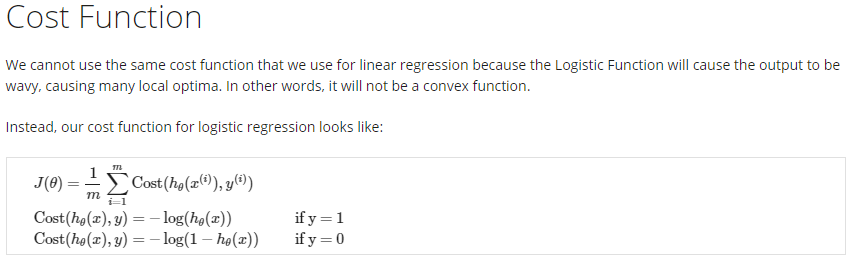


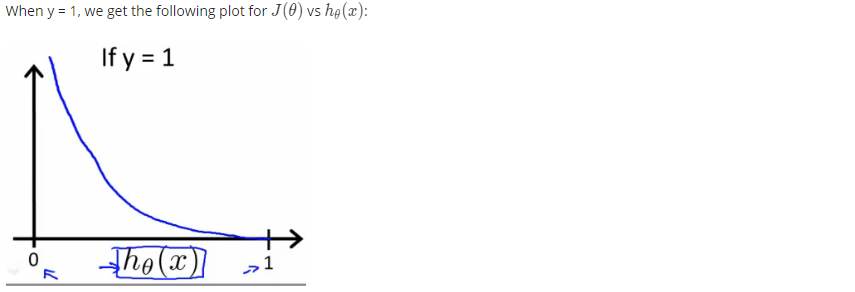


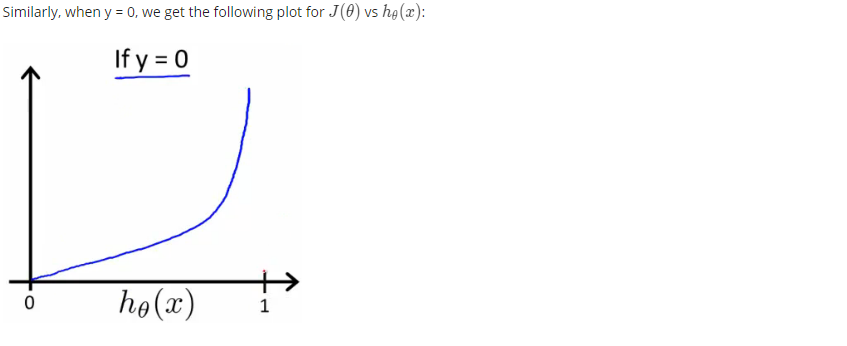
## Logistic Regression Model

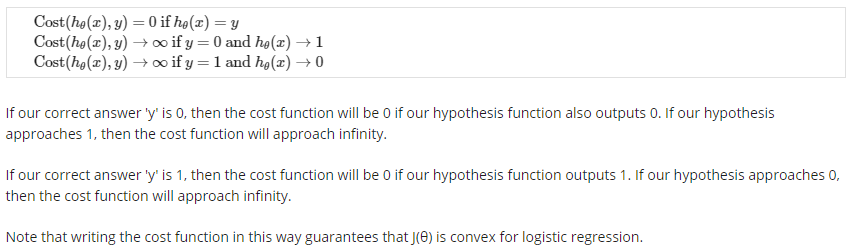
Cost Function

* Make J convex





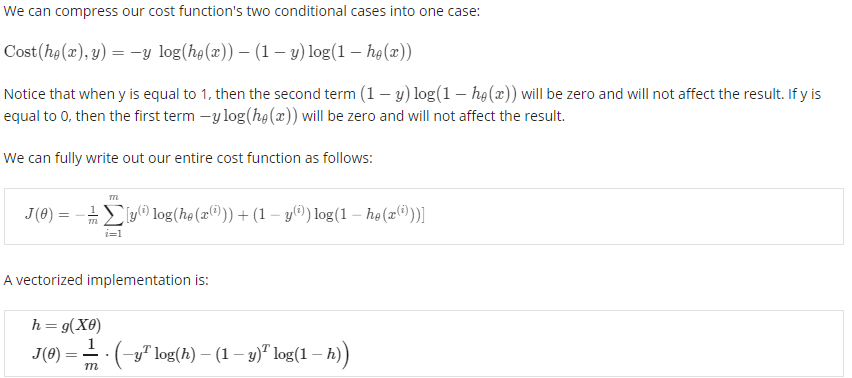


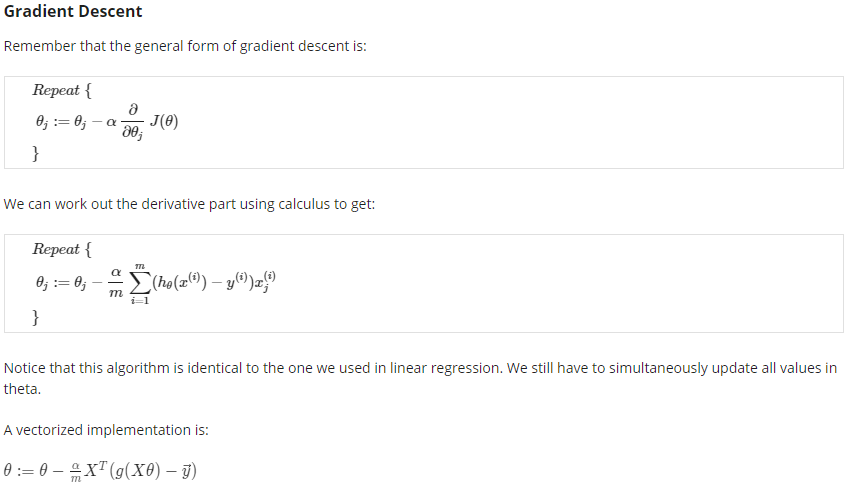


Simplified Cost Function and Gradient Descent

* Logistic regression cost function: derived from statistics using the principal of maximum likelihood estimation; efficiently find parameters for different models; convex
* Same algorithm as of linear regression except h is different



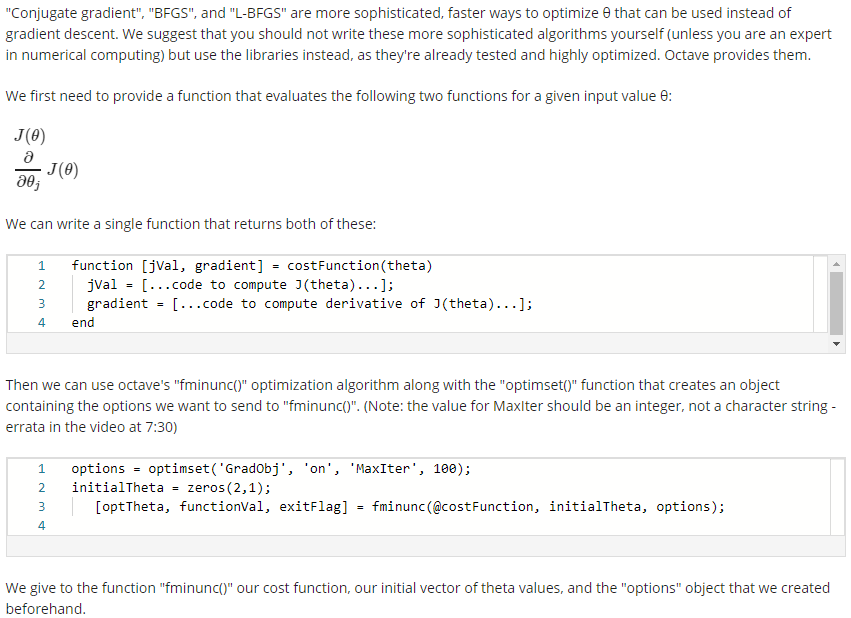




Advanced Optimization

* Optimization algorithms
* Gradient Descent
* Conjugate Gradient
* BFGS
* L-BFGS

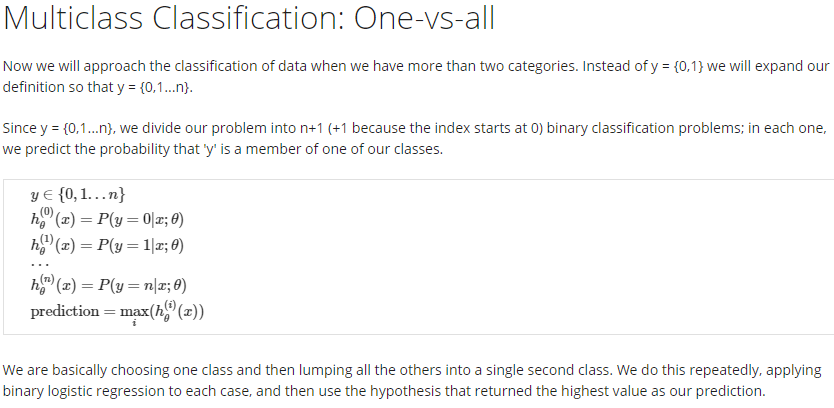


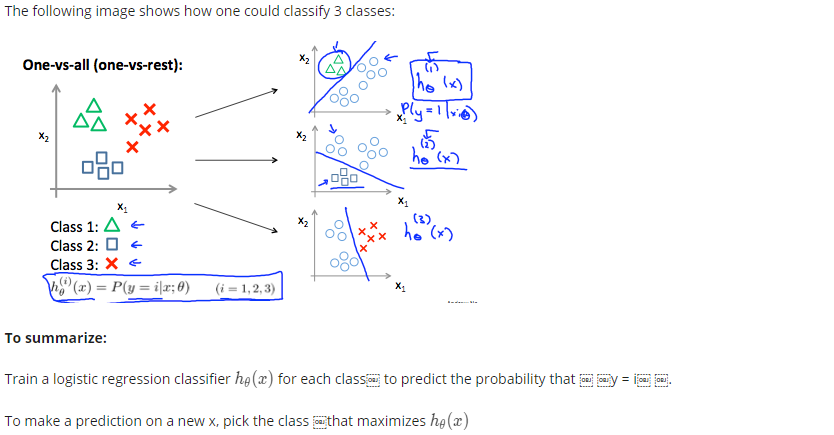


## Multiclass Classification

Multiclass Classification: One-vs-all

* One-vs-rest



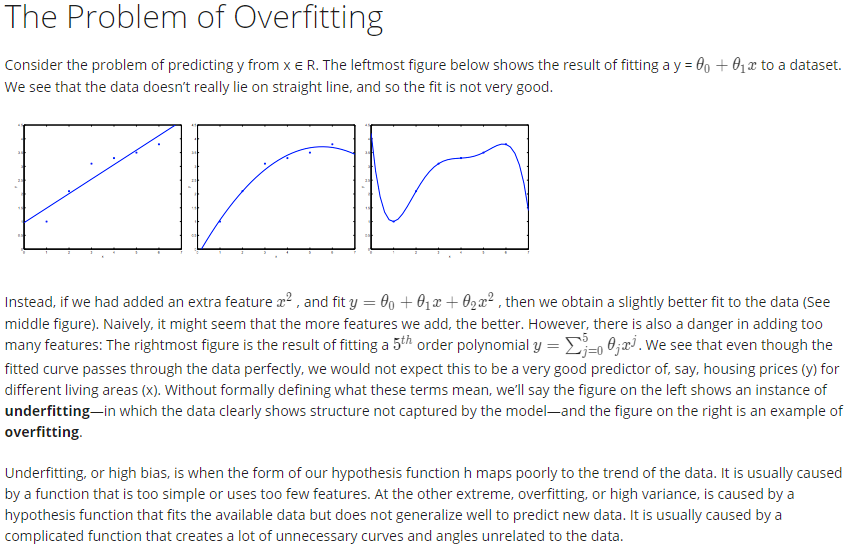


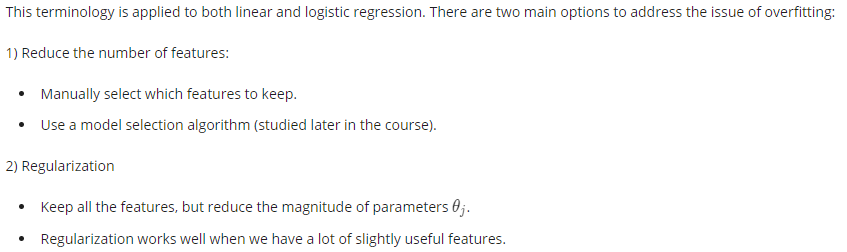
# Regularization

## Solving the Problem of Overfitting

The Problem of Overfitting

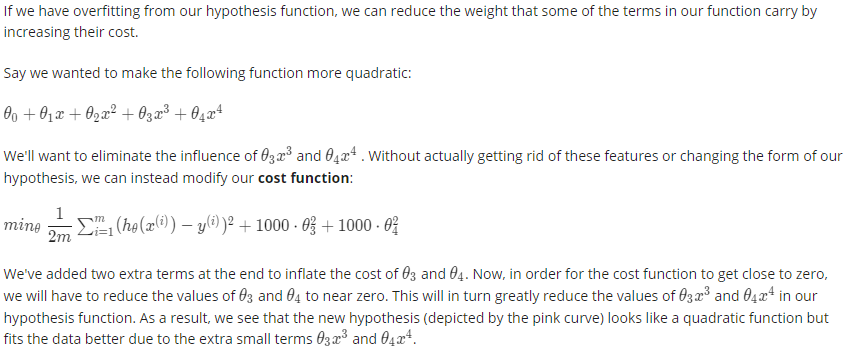
* Underfitting = high bias
* Overfitting = high variance
* Address overfitting
* Reduce number of features
* Manually select which to keep
* Model selection algorithm
* Regularization
* Keep all features but change weights
* Avoid losing info about problem

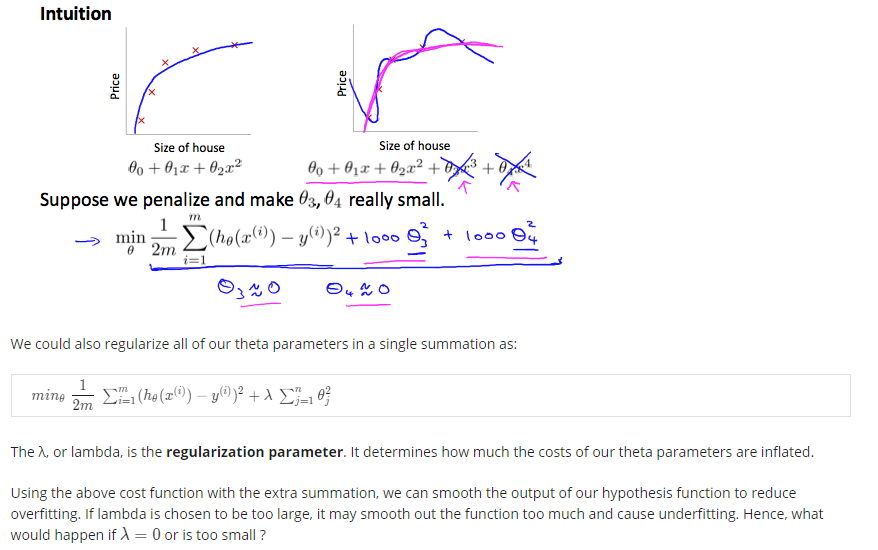




Cost Function



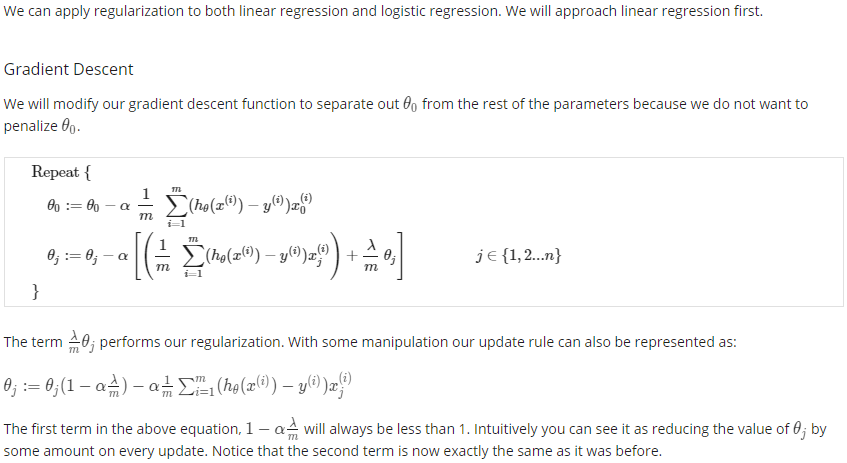


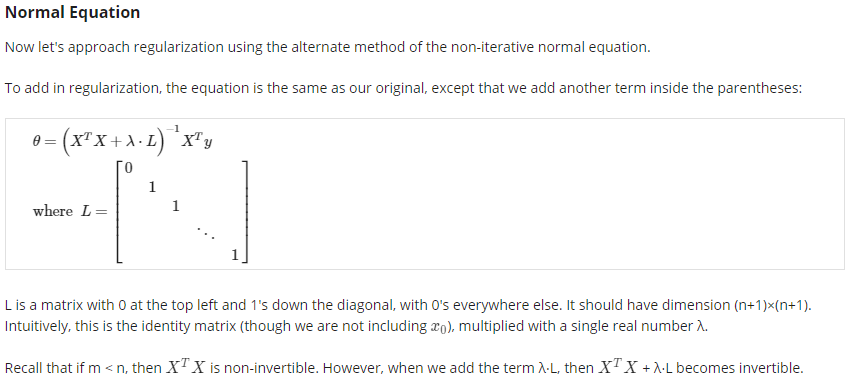


Regularized Linear Regression

* Reduce θ values







Regularized Logistic Regression

