Warm-up as You Log In

Assume that exam scores are drawn independently from the same Gaussian (Normal) distribution.

Given three exam scores 75, 80, 90, which pair of parameters is a better fit?

A) Mean 80, standard deviation 3

B) Mean 85, standard deviation 7

Use a calculator/computer.

Gaussian PDF: $p(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$

Announcements

Assignments

- HW3
  - Mon, 9/28, 11:59 pm
- HW4
  - Not out until after the midterm

Schedule change this week

- Recitation slots this Friday will all be lecture (all three)

Announcements

Midterm 1

- Mon, 10/5
- See Piazza for details
- SGD not in scope for Midterm 1
- Practice exam
  - Timed (90 min) exam in Gradescope
  - Open for a 24 hour window only, Tue 7 pm to Wed 7 pm
  - Need to take the practice exam to have access to the questions
  - Also, practice exam zoom sessions
    - Tue 7 pm
    - Tue 11 pm

Plan

Last time

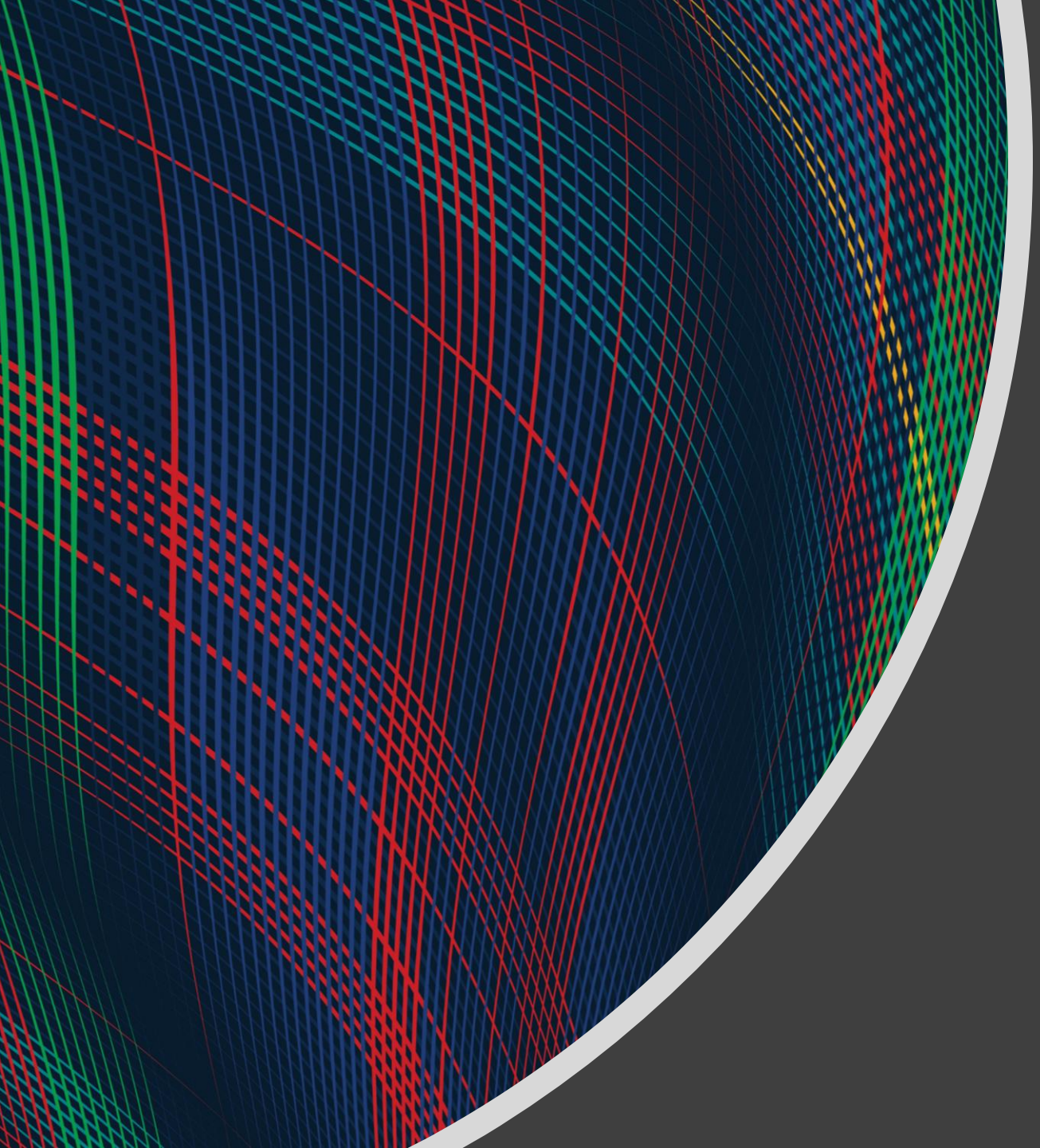
- Optimization for linear regression
  - Linear and convex functions
  - (Batch) Gradient descent
  - Closed-form solution

Today

- Stochastic gradient descent
- Logistic Regression
  - Back to classification
- Likelihood
- MLE

Wrap-up (Stochastic) Gradient Descent

Previous lecture slides



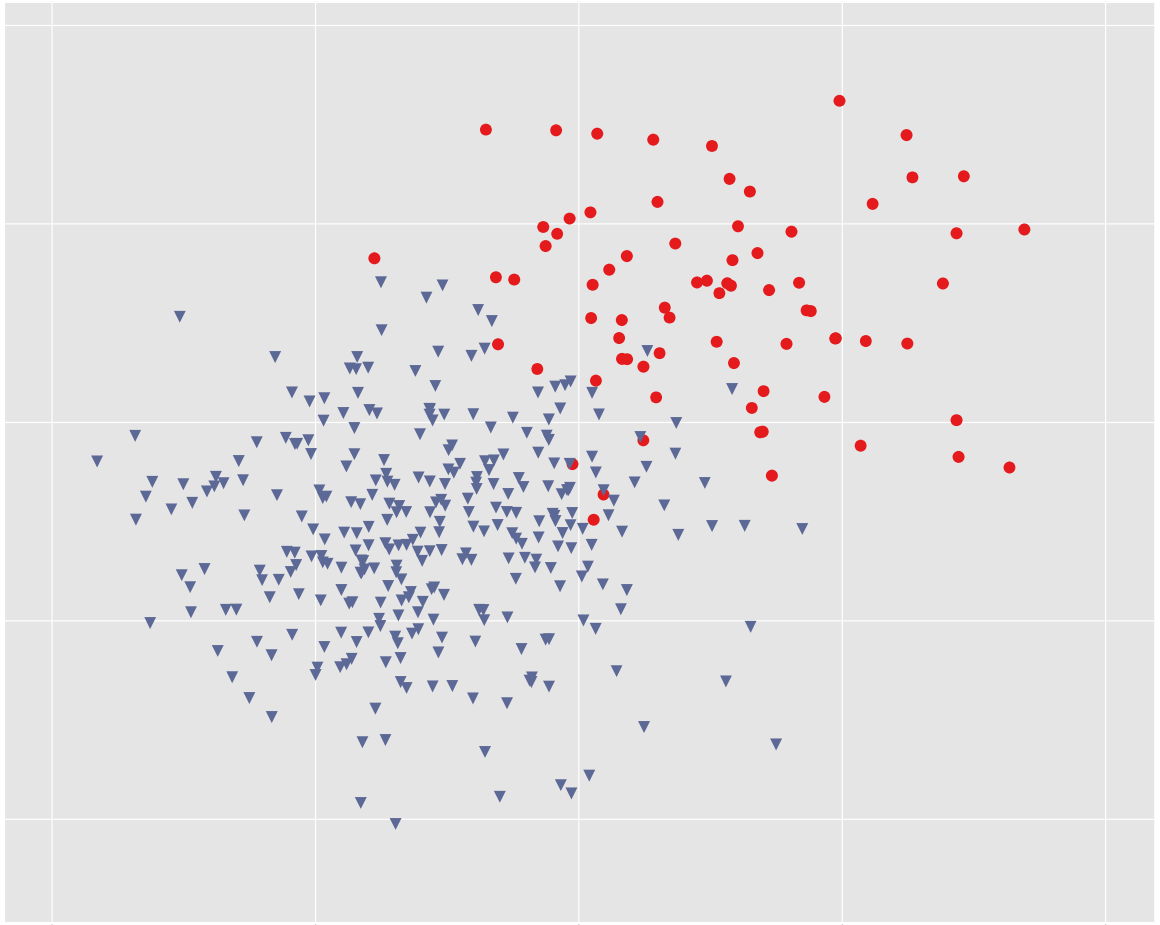
Introduction to Machine Learning

Logistic Regression

Instructor: Pat Virtue

Prediction for Cancer Diagnosis

Learn to predict if a patient has cancer ($Y = 1$) or not ($Y = 0$) given the input of two test results, X_A and X_B.

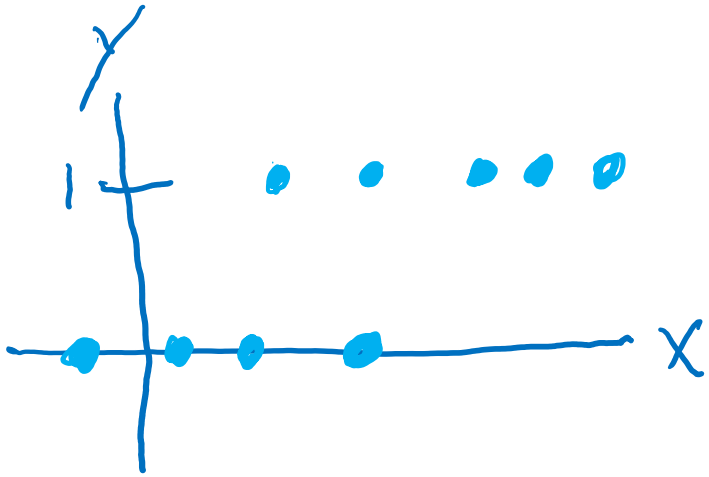


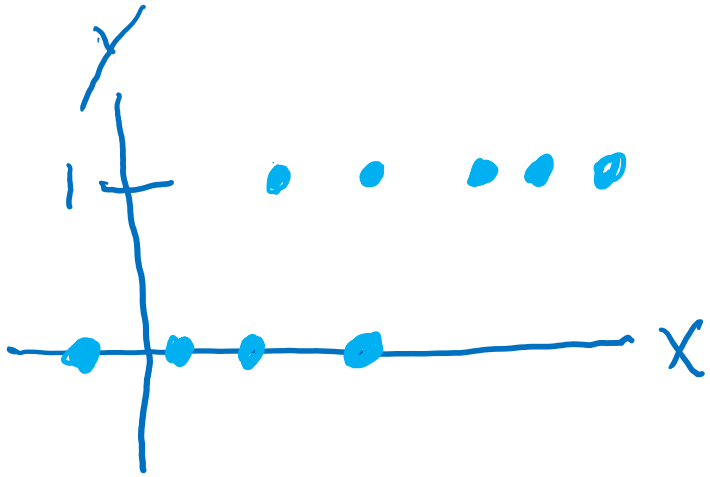
Prediction for Cancer Diagnosis

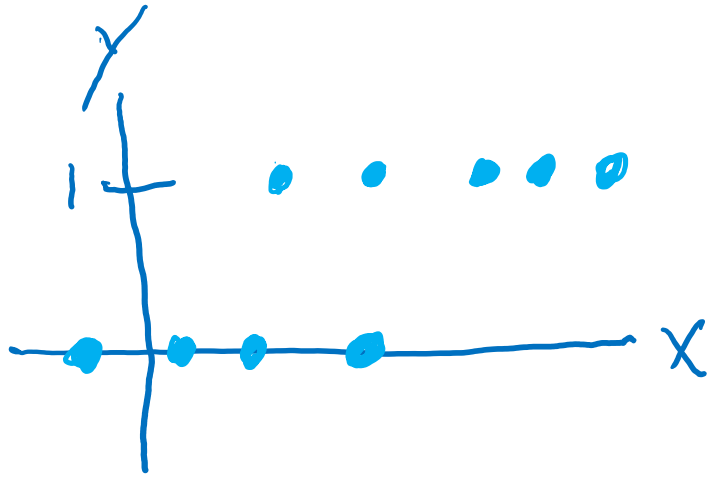
Learn to predict if a patient has cancer ($Y = 1$) or not ($Y = 0$) given the input of just one test result, X_A.

Building on a Linear Model

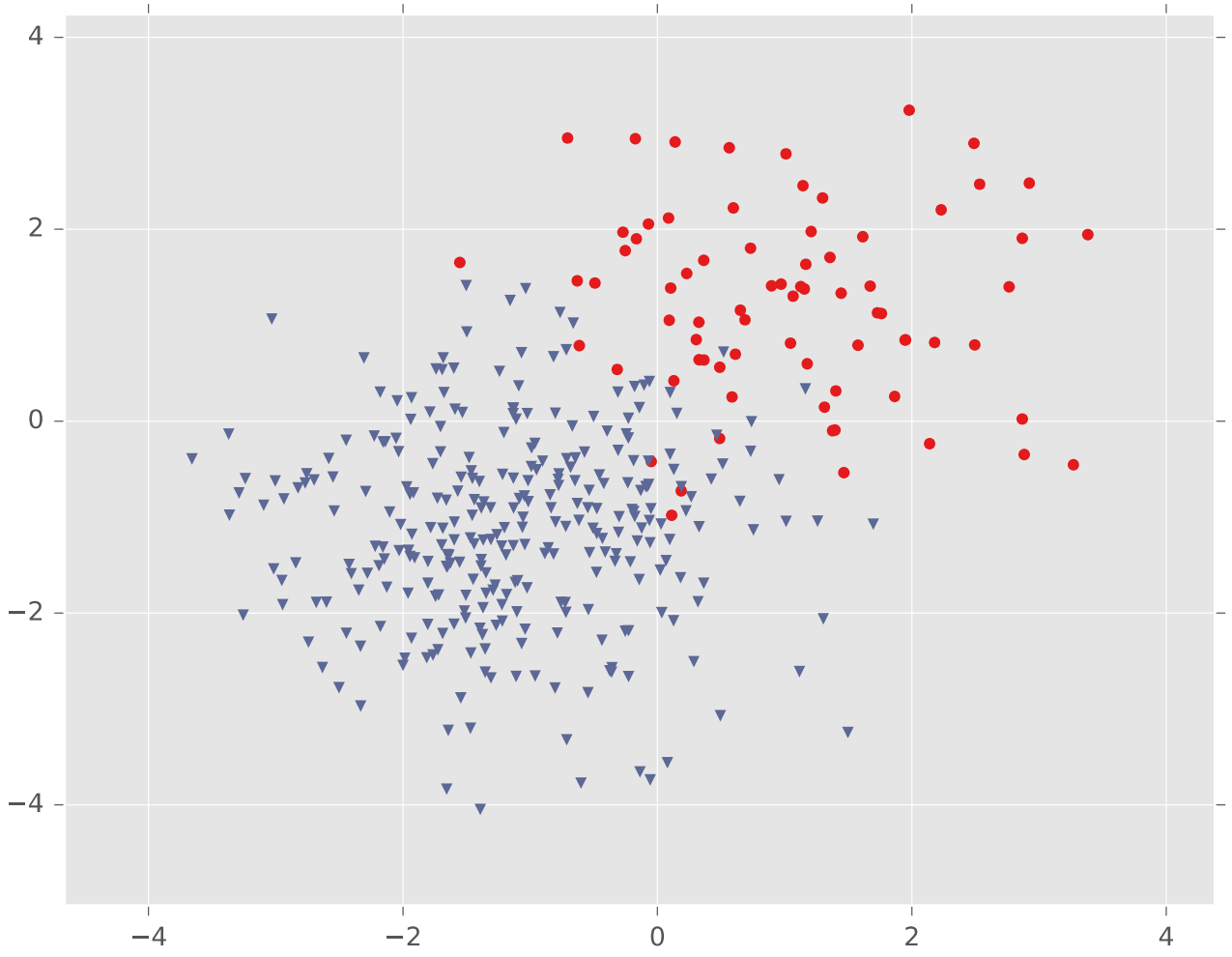
Linear vs Thresholded Linear vs Logistic Linear



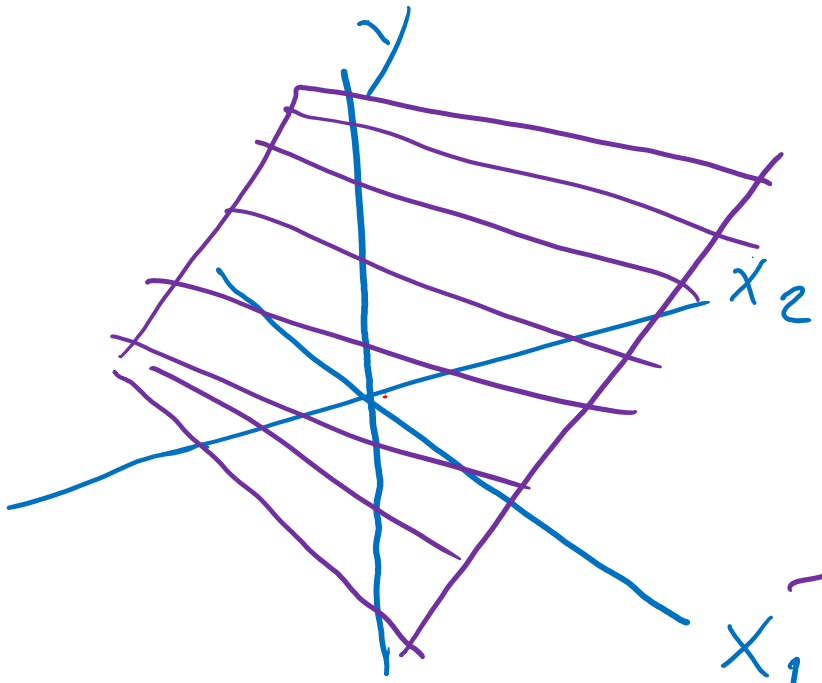




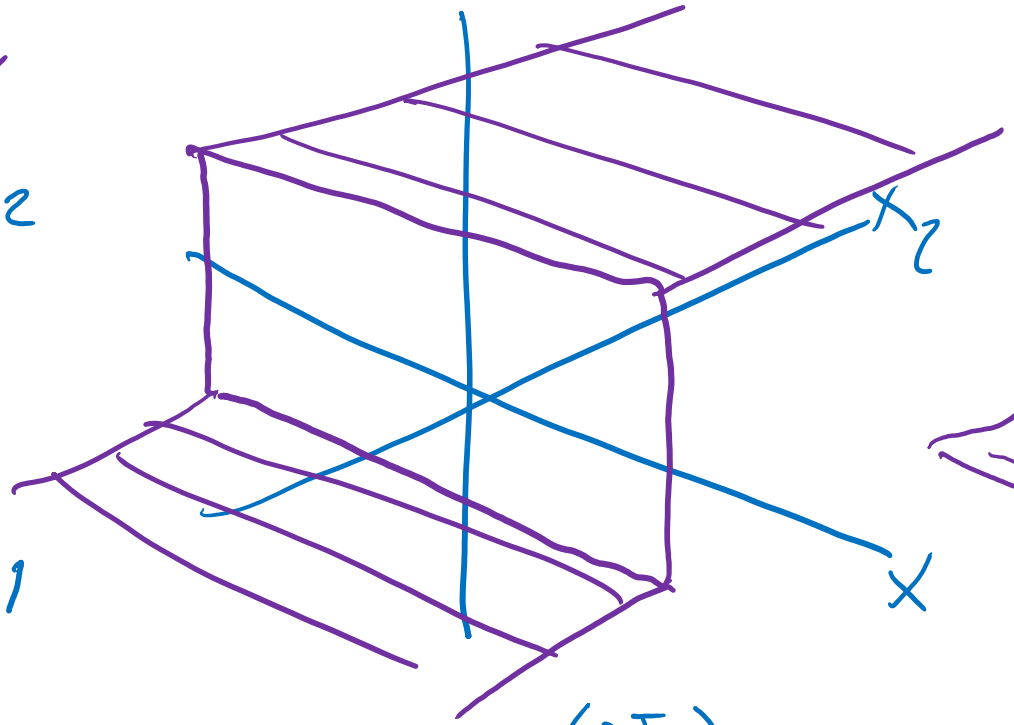
Building on a Linear Model



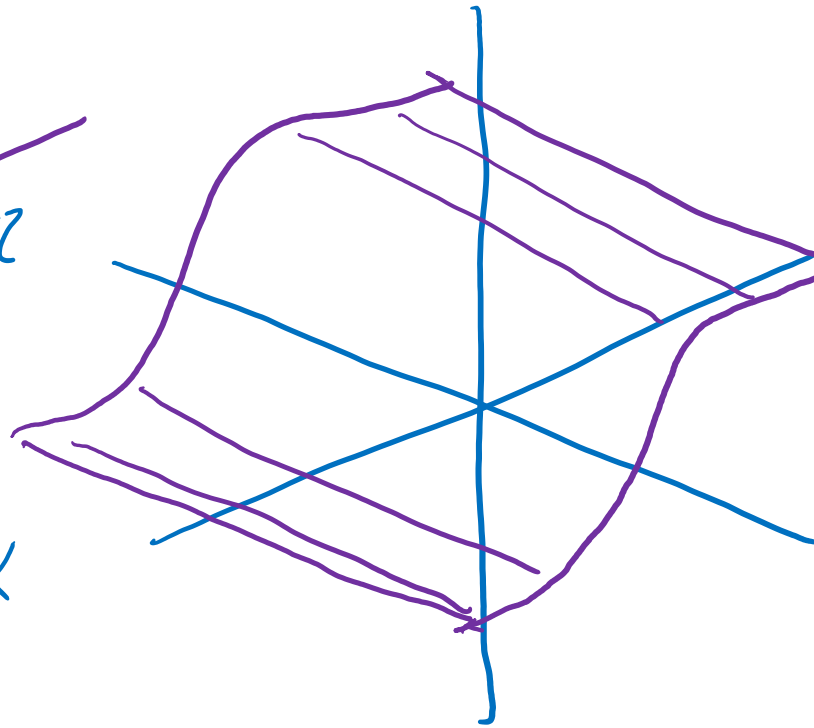
Building on a Linear Model



$y = \vec{\theta}^T \vec{x}$



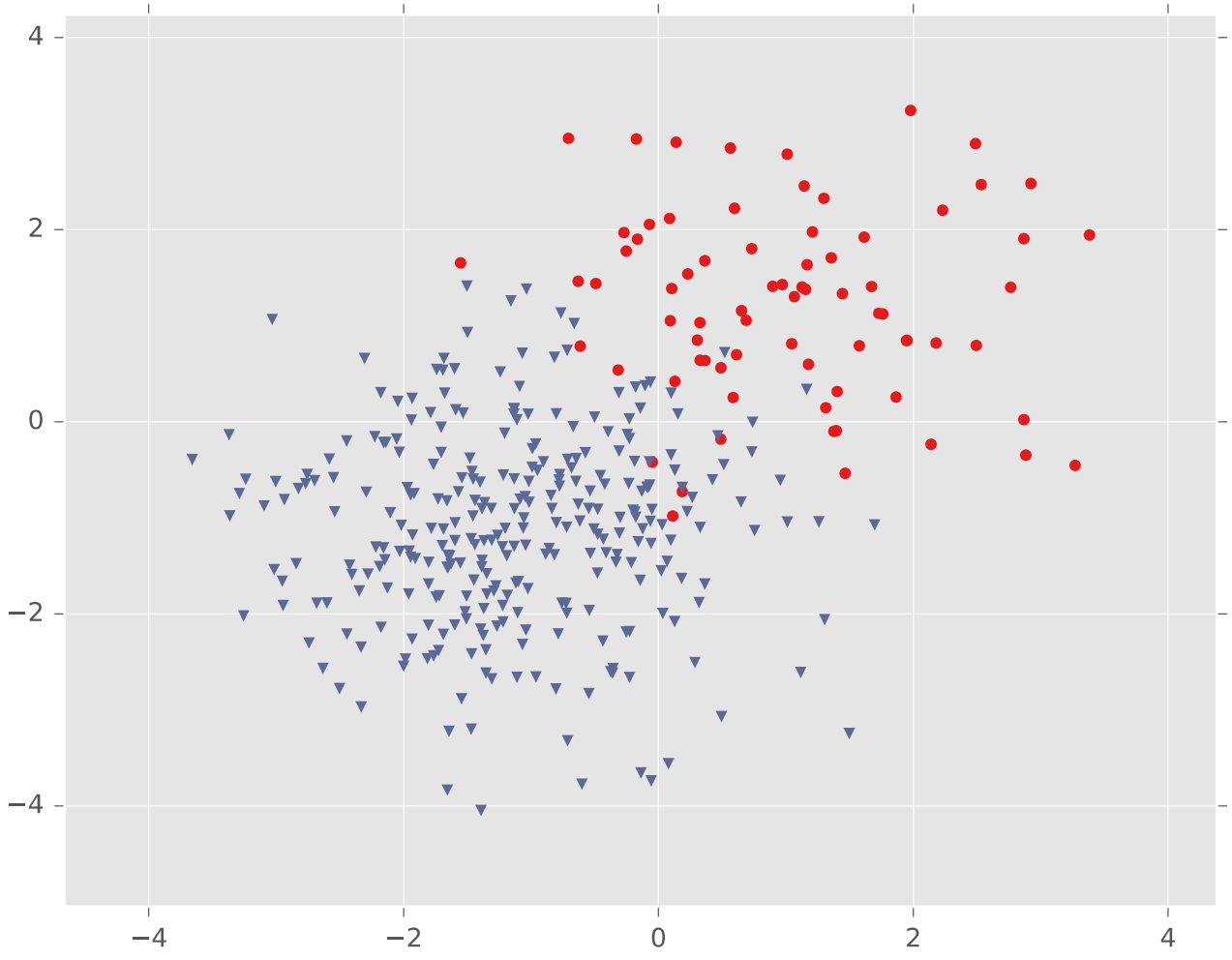
$y = \text{sign}(\theta^T x)$



$y = \sigma(\theta^T x)$

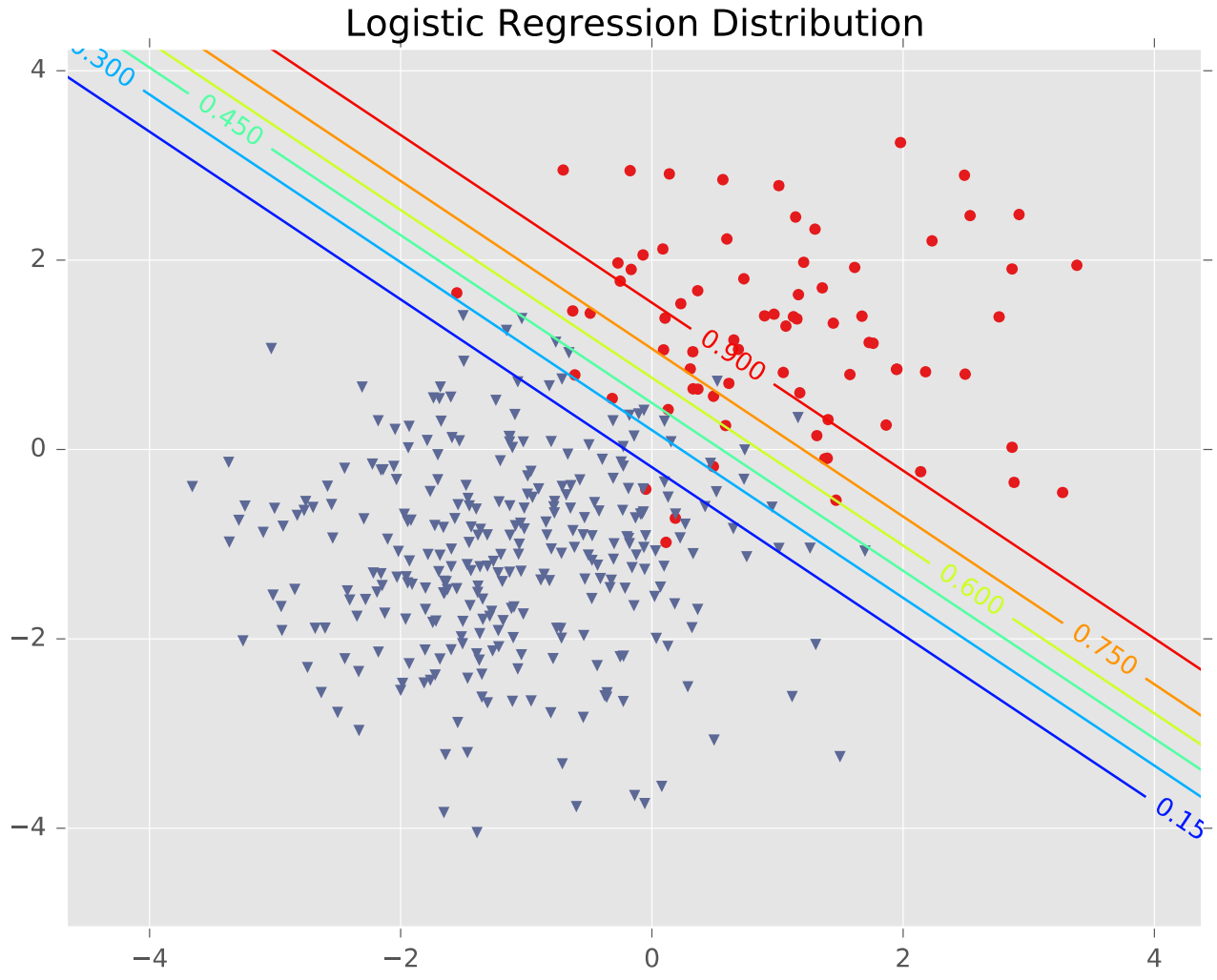
$p(y|x)$

Logistic Regression



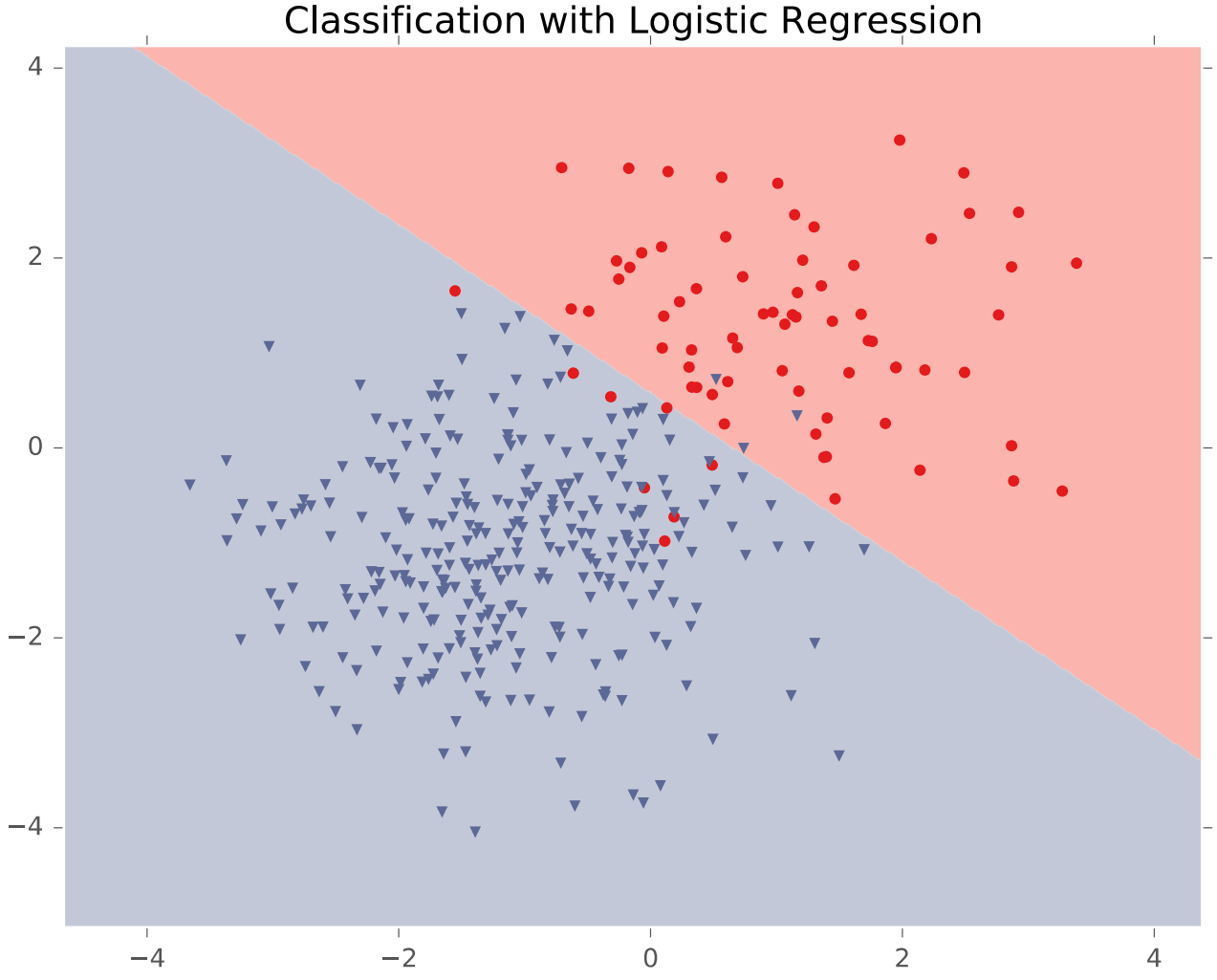
Slide credit: CMU MLD Matt Gormley

Logistic Regression



Slide credit: CMU MLD Matt Gormley

Logistic Regression

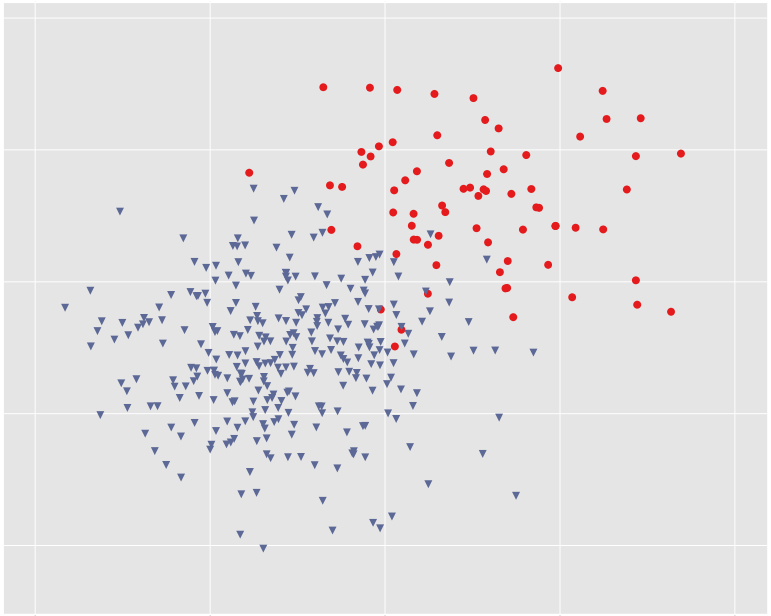


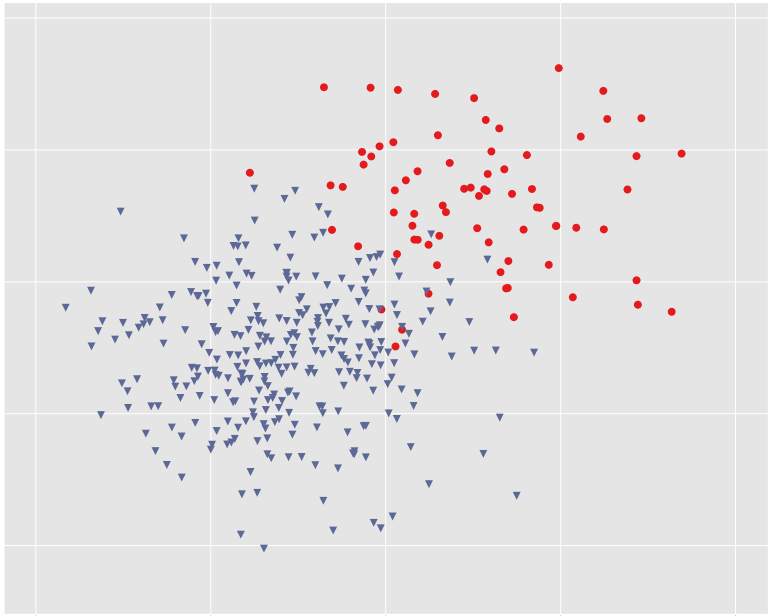
Slide credit: CMU MLD Matt Gormley

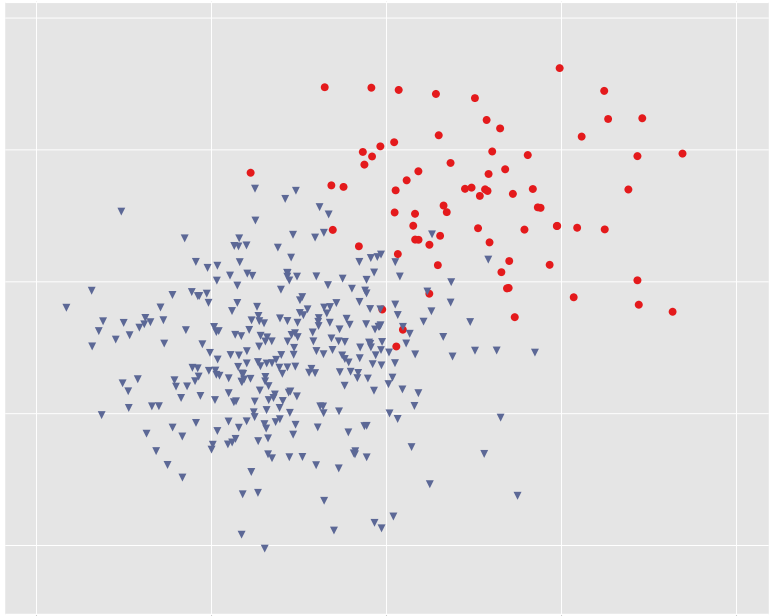
Prediction for Cancer Diagnosis

Learn to predict if a patient has cancer ($Y = 1$) or not ($Y = 0$) given the input of just one test result, X_A.

$$p(Y = 1 \mid \boldsymbol{x}, \boldsymbol{\theta}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^T \boldsymbol{x}}}$$







LIKELIHOOD AND MLE

Likelihood

Likelihood: The probability (or density) of random variable Y taking on value y given the distribution parameters, $\boldsymbol{\theta}$.

Likelihood

Likelihood: The probability (or density) of random variable Y taking on value y given the distribution parameters, $\boldsymbol{\theta}$.

Grades

Warm-up as You Log In

Assume that exam scores are drawn independently from the same Gaussian (Normal) distribution.

Given three exam scores 75, 80, 90, which pair of parameters is a better fit?

A) Mean 80, standard deviation 3

B) Mean 85, standard deviation 7

Use a calculator/computer.

Gaussian PDF: $p(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$

Likelihood

Likelihood: The probability (or density) of random variable Y taking on value y given the distribution parameters, $\boldsymbol{\theta}$.

i.i.d.: Independent and identically distributed