Lecture 19 Logistical Regression

* Trying to find a threshold that distinguishes points in one class from points in another
* parameters of line used as tool to assign likeness to data points
  + Changing parameters of line lets us control the likeliness we assign to regions of our space
    - We can amplify or de-amplify this which allows us to control how quickly or slowly we transition from predicting a particular class
    - This allows us to grow or shrink the uncertainty region around the threshold
  + Represent likeliness as a probability
    - Points closest to the threshold line should have a probability of ½ representing that they can likely be in either class
    - Points further from the threshold should have a larger probability result
      * Further from the line -> more certain point belongs to a particular class
    - Sigmoid function represents the above properties
      * Special case of logistic function
      * Comparable to the normal distribution PDF
  + To get a prob of a particular point
    - Plug point into equation of threshold
    - Then plug result into the sigmoid function to get prob
    - For a binary distribution P(B) = 1-P(A)
* How do we know if the model describes the data well
  + Try to find the parameters of line that make the data we saw as likely as possible
* Finding a threshold that generations a sigmoid function that best describes the data we observed
* Describing the data well means finding what makes the data we saw the likeliest
  + The more data we gather the easier this is the predict
* Nudging the parameters of the line in a direction that should improve the likeliness of our data we can mold the sigmoid function to our data and capture what distinguishes points in one class from points in another

Multiclass linear regression

* Finds a less pronounces sigmoid function
* Adapt sigmoid function to return multiple thresholds
* One V rest classification