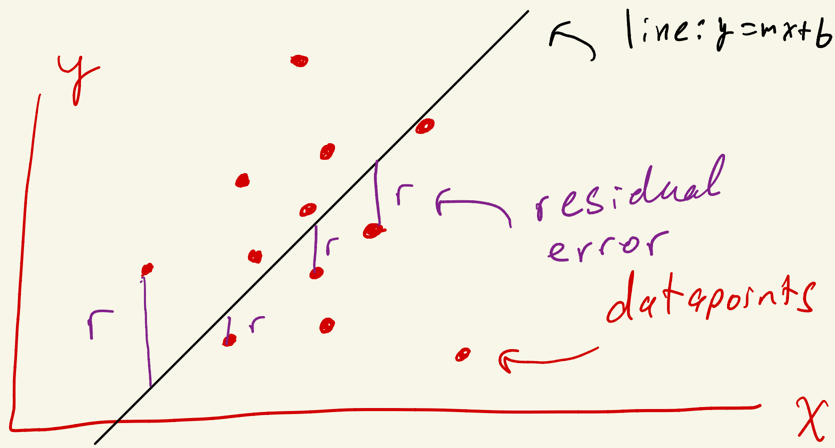
EECS 551 Discussion 4

- Classification via linear regression

* Linear Regression: Finding the "line of best fit" to minimize residual error between points & line.

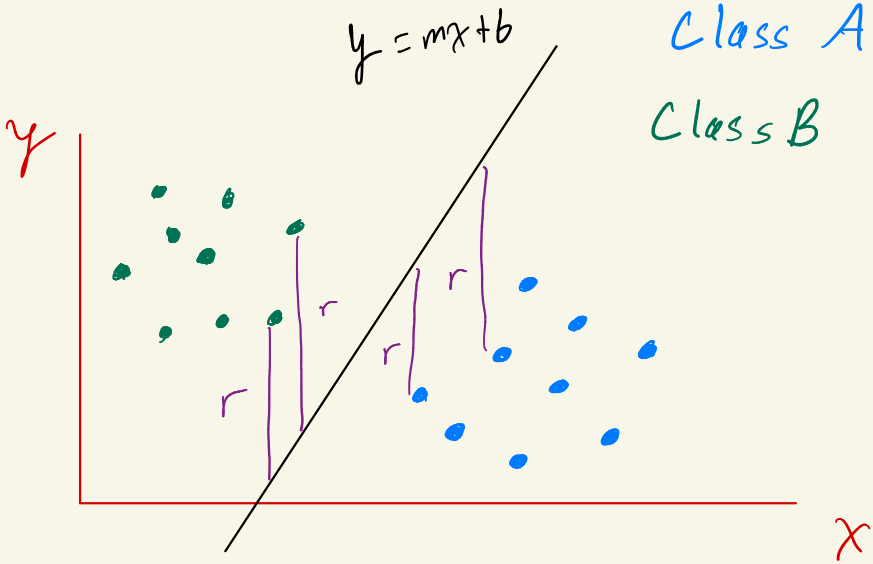


When finding this line, we want to "learn" the coefficients m & b from our data.

Now, imagine finding a line to fit data from two distinct classes:

Class A

Class B



This line will separate the two classes (if they are linearly separable) & we can use the line for classification of a novel sample drawn from either Class A or Class B.

* Least Squares Linear Regresssion

Goal: Solve $\hat{x} = \underset{x}{\operatorname{argmin}} \|y - Ax\|_2^2$

given a set of training data.

labels

"feature matrix"

Predictor

We know $\hat{x} = A^+ y$.

* What is a feature matrix?

- Each row of A corresponds to a training sample, each column corresponds to a different feature.

- Therefore, when we compute Ax, we are taking inner products between feature vectors & our predictor... →

So to make a prediction of a new/test sample, compute

$$\hat{y} = a^T \hat{x}$$

prediction

feature vector of test sample.

learned classifier, has coefficients learned from training to match labels

In this task, training samples are images. The features of each image we are concerned with are

We include a "1" feature to allow for an offset b in our line

[mean of Image, middle column mean of image, 1]

We are using labels $y_i \in \{-1, 1\}$.

* Also, $y \simeq A\hat{x}$, so our predictor will not perfectly find labels $\hat{y} \in \{-1, 1\}$.

* It then makes sense to use 0 as the decision boundary. $\hat{y} = \text{sign}(a^T\hat{x})$.

- To visualize the decision boundary, solve $a^T\hat{x} = 0$ & plot a_2 in terms of a_1 & coefficients of $\hat{x}$.