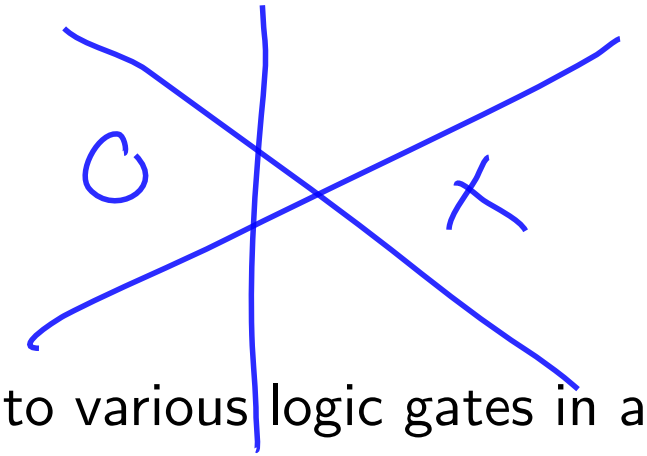
# SMAI-M20-L19: Logistic Regression

C. V. Jawahar

IIIT Hyderabad

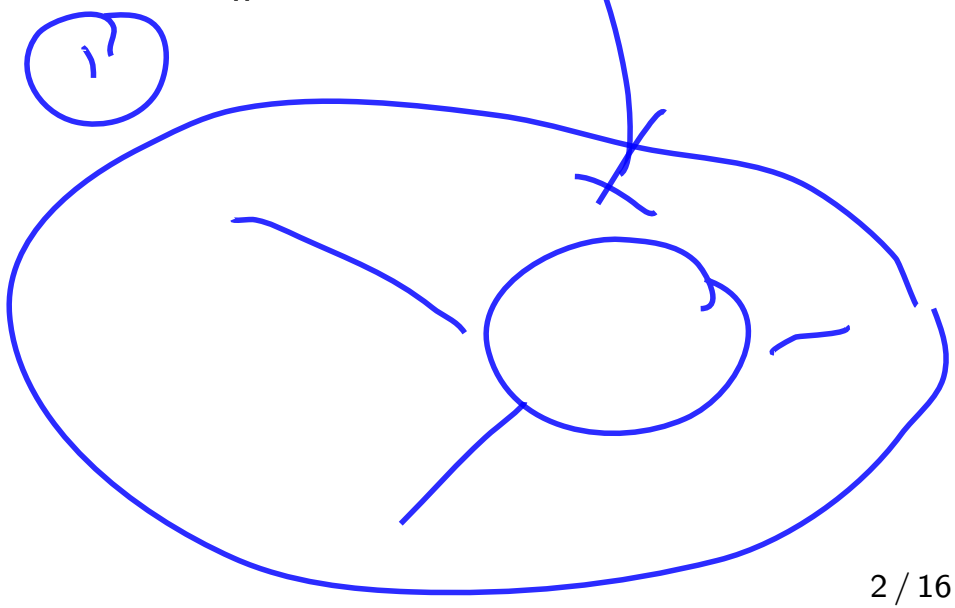
September 23, 2020

# Class Review



What weights can lead to various logic gates in a Single Layer Perceptron?

- with $\{-1, +1\}$logic or $\{0, 1\}$ logic
- AND, OR, NAND, NOR, NOT etc.
- with various definitions of "activations" $\phi()$.

# Blank

# Announcements

**Quiz 1**

- Mostly based on topics upto (including PCA). No focus on perceptron, GD, Logistic Regression.
- Assumes. you followed lecture videos, class reviews, discussion points, home works.
- Four parts (each like a class review). Each of 5 questions. Total: 20 questions. Modular.
  - No discussions with friends, experts, classmates during the quiz.
  - Not a quiz of 20 question. Think of it as four class reviews.
  - One part is numerical. Keep a calculator, if required.
  - One part may ask you to refer to the class material. Keep course material, your notes, a computer accessible. Keep a pen/paper handy.
  - Each part: 10 mins within a span of 15 mins (say Part 1 during 6.30-6.45pm). Take care of brief network/power outages. Brief brakes required by individuals.
  - Any issues, email course email address. (not to individuals). Any clarifications/announcements will be in the "Quiz Channel".
  - People who miss quiz due to any genuine issue, do not attempt. Write email. We will see how to take care, including a possible extra one.

# Recap:

- Supervised Learning:
  - Notions of Training, Validation and Testing; Loss Function and Optimization, Generalization, Overfitting, Occam's razor, Model Complexity, Bias and Variance, Regularization.
  - Performance Metrics, Estimating error using validation set.
- Approaches:
  - Optimal Decision as $\omega_1$ if $P(\omega_1|\mathbf{x}) \geq P(\omega_2|\mathbf{x})$ else $\omega_2$, MLE
  - Dimesnionality Reduction and Representation ( Feature Selection, PCA, Neural Embeddings)
  - Application of PCA: Eigen Face
  - Matrix Factorization for Data Matrices (SVD, Eigen Decomposition)
  - Application of Matrix Factorization: LSI, Matrix Completion, Recommendation Systems)
  - Nearest Neighbour, Linear Discriminants
  - Gradient Descent
  - Linear Regression: Closed form, GD, Regularization, Optimization
  - Perceptron Algorithm and Neuron Model
  - Logistic Regression

# This Lecture:

1. **Logistic Regression - II**
   - How do we formulate the Logistic Regression Objective as MLE?
2. **LDA/ Fisher**
   - Linear Discriminant Analysis or Fisher Discriminant
   - Supervised Dimensionality Reduction
3. **Multi Class Classification**
   - Why it is non-trivial for multi-class?
   - Practical issues (i) computational complexity (ii) Fusion/Decision making scheme

**Questions? Comments?**

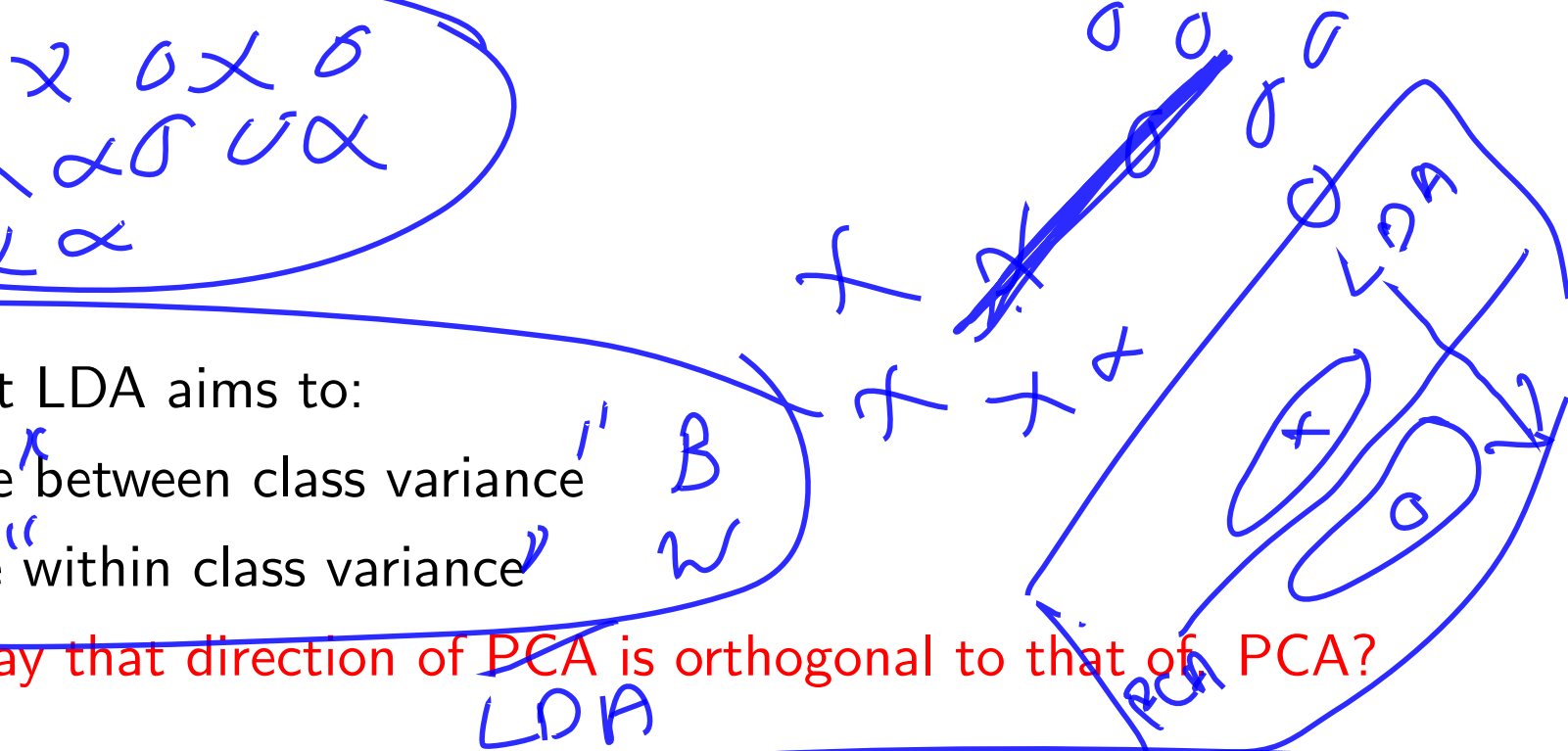
# Blank

# Discussions Point - I

We know that LDA aims to:

- maximize between class variance
- minimize within class variance

Is it true to say that direction of PCA is orthogonal to that of PCA?



# Blank

If comp → PCA

class → LDA

Linear

"Best Dim Red for class"

⇒ ?

# Blank

# Blank

# Discussions Point -II

Comment on the following three ways of designing multi-class classifiers (number of classes = K):

1. One vs Rest Classifiers and some fusion scheme.
2. Pairwise Classification with Majority Voting
3. Binary Hierarchical Classification

based on the following dimensions:

A Number of classifiers to be trained

B Number of classifiers to be evaluated for testing a single sample.

C Difficulty of the classification problem that each classifier solves (if required, assume classes are Multivariate Gaussian)

Comments A1, A2, A3, B1, B2, B3, C1, C2, C3 and finally what do you prefer? why? when?

# Blank

# Blank

# Blank

# What Next:? (next three)

1. Logistic Regression
2. Multi Class Classification (beyond binary)
3. More Dimensionality Reduction Schemes (eg. LDA/Fisher)