SMAI-M20-02: Representation and Classification

C. V. Jawahar

IIIT Hyderabad

August 12, 2020

Agenda

1. Recap/Repeat of L01
   - https://www.dropbox.com/s/ltmyx9y15hxnmm8/L1.pdf?dl=0
2. This lecture:
   - Representation as a vector
   - Nearest Neighbour Algorithm
   - Linear Classification
3. Discussions and Extensions
4. Reviews and Next Step

Recap

- Exposure:
  - Maths in the form of UG Courses and Schools.
  - Programming/Algorithms: Comfortable with adaptation/hand-on.
- Maths:
  - Typical Engg Maths. Not super advanced.
  - Topics: Linear Algebra, Probability, Differential Calculus
- Tools:
  - Most tools/libraries are in python. No plan to teach python or programming.
  - `https://jupyter.org/`
- Infrastructure:
  - Some compute (say a laptop) advisable. Though VPN/Remote/Cloud may also work in many cases.
  - Some internet connection (say 4G) expected. Course accounts buffers to take care of unfortunate network failures.

About the course: Course Coverage

1. **Part 1: Basics**
   - Mathematical Foundations; Role of Linear Algebra and Probability; Supervised Leaning Formulation and Challenges, Sample Algorithms
2. **Part II: Fundamental Algorithms**
   - Linear Methods in Machine Learning; Regression, PCA, Logistic Regression, Perceptrons, Gradient Descent, Multiclass
3. **Part III: Powerful Ideas**
   - SVMs, Kernels, Nonlinear Methods, Ensembe, Semi-Supervised, Unsupervised and Self-Supervised Learning
4. **Part IV: Neural Network Learning**
   - Artificial Neural Networks, MLP and Back Propagation; Intro to Deep Learning, Intro to CNN, RNN

Approximately 25% emphasis/time on each part.

Course Structure and Evaluation

- Lectures, Tutorials (as per time table)
  - Lec: M,W,F: 9.30-10.30AM IST
  - Tut: Wed. 5.30-7.30 (may be 1 Hr batches?)(Watch for Announcements)
- Office Hours (OH) (Additional Support; Weekly (?)):
  - Logistics, Exceptions, Semi-Technical, Administrative
  - TA (One Common; One on Reserving)
- Emails:
  - Use Office Hours, if your request can wait. Get personal attention.
  - Any course specific emails: please send to: **smai.m2020@gmail.com**
- Expected course load:
  - Regular activities (Homeworks, Assignments)
  - Some preparation (20-30 mins before the lecture)
  - Reading, Thinking, Discussing
  - Meetings: Lec (*), Tut, OH (Optional)
- **If attending, attend seriously and disciplined.** This is a large class. May be not everyone was not given opportunity.

Course Structure and Evaluation

- Regular Homeworks: 40%
  - Handwritten and Programming
  - Approximately 200 Points. Your best 80% (Approximately out of 160) will be used for grading
  - Stretch beyond the class/lecture. Learning+Evaluation is the goal.
- Quiz (3): 25%
  - Closest to the traditional exams.
  - Evaluation is the primary purpose
- Assignment (3): 25%
  - Programming and Exploration
  - Exposure to the depth is the focus.
- In Class Review/Recap: 10%
  - Objective questions and regular review. Light weight. Learning is the primary objective.
- 5-10% changes due to operational issues of uncertainties.

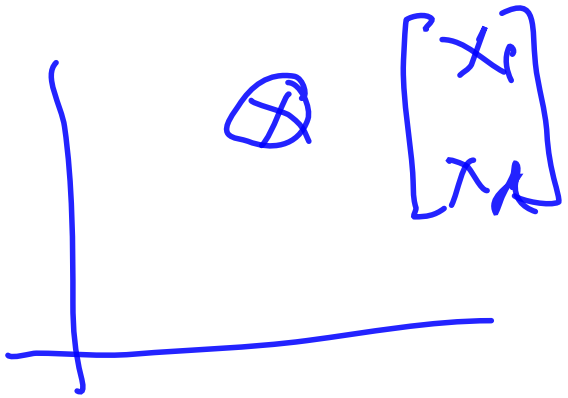
Recap: Technical Summary

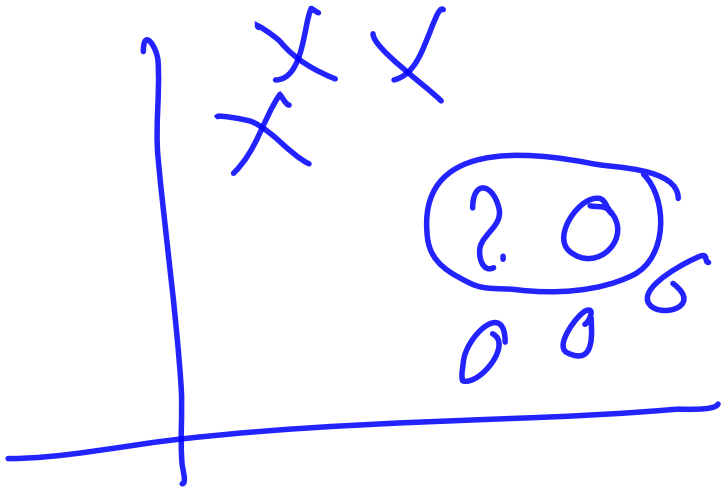
- We had seen how data can lead to learnable parameterized functions $f(\mathbf{W}, \mathbf{x})$
- The notion of "Training" and "Testing"
- We started with an overview at: `https://www.youtube.com/watch?v=8xniRSjRyCQ&feature=youtu.be` on how data can help in solving problems.

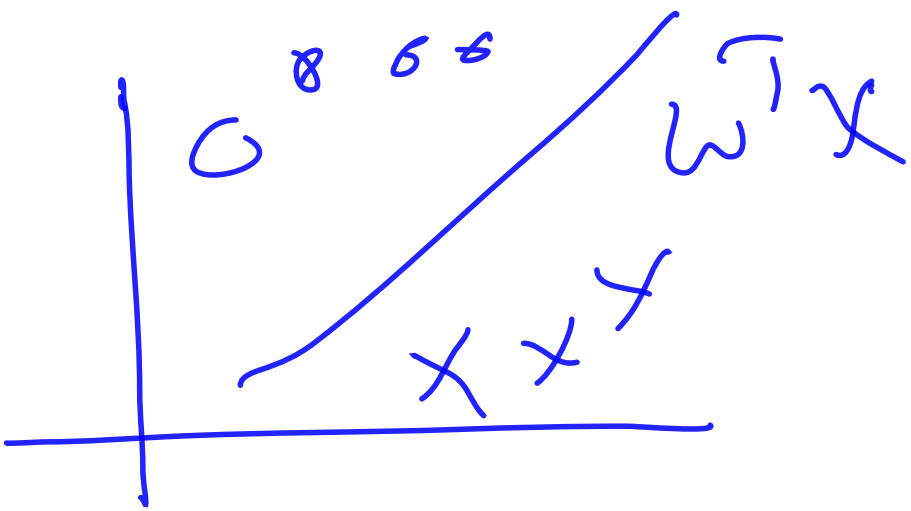
This Lecture

- How to represent in the form of $\mathbf{x} \in R^d$
- Nearest neighbour Algorithm for Classification
  - Classify based on the majority labels in the neighbourhood.
- A simple Linear Classifier
  - $sign(\mathbf{w}^T\mathbf{x})$
  - Either +ve or -ve.
- If you have not see the pre-lecture videos, please do. At least the one on "Linear Classifier"
  - Go to MS Teams
  - Go to the channel"Lectures - Information". See the posts
  - Watch the last video/link.
  - `https://www.youtube.com/watch?v=P92mkhzt6Hg&feature=youtu.be`
- (We will take 10 mins for people who might have missed; But as we move forward, do plan 30 mins of pre-lecture preparation previous day.)

Blank







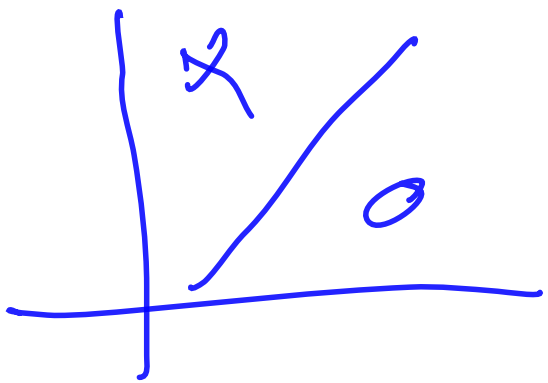
$$\text{Sign}(\vec{\omega}^T x)$$

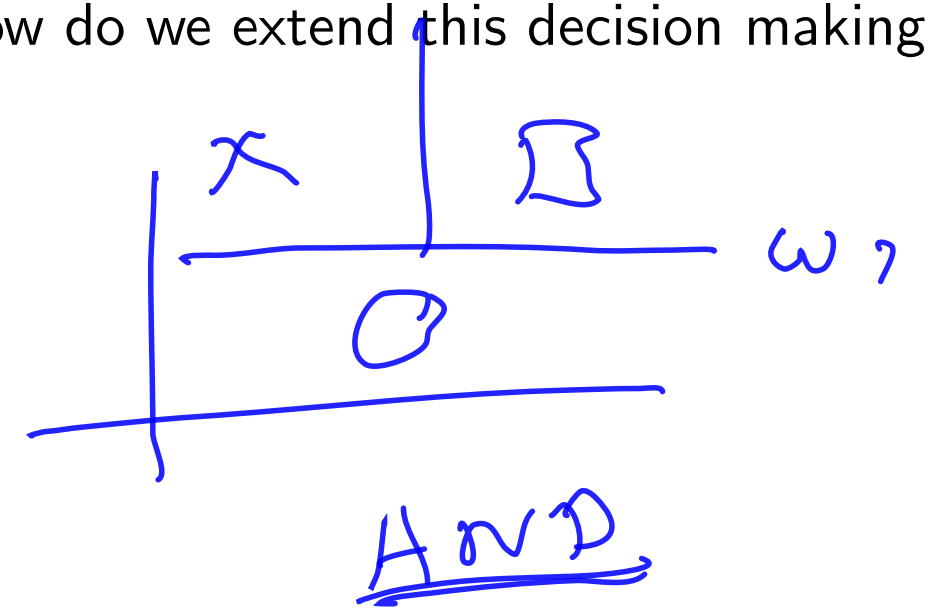
Discussions

Q: We know, if we have two classes, we can classify as:

$sign(\mathbf{w}^T\mathbf{x})$

Assume we have three classes, how do we extend this decision making logic?





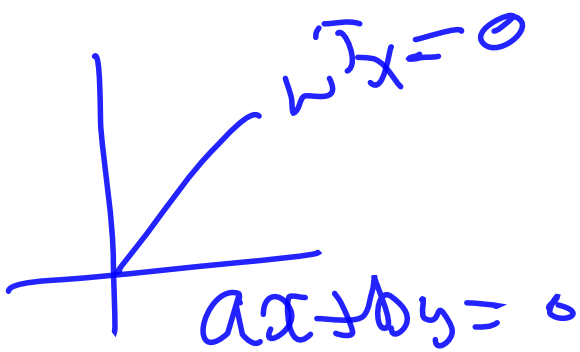
Blank

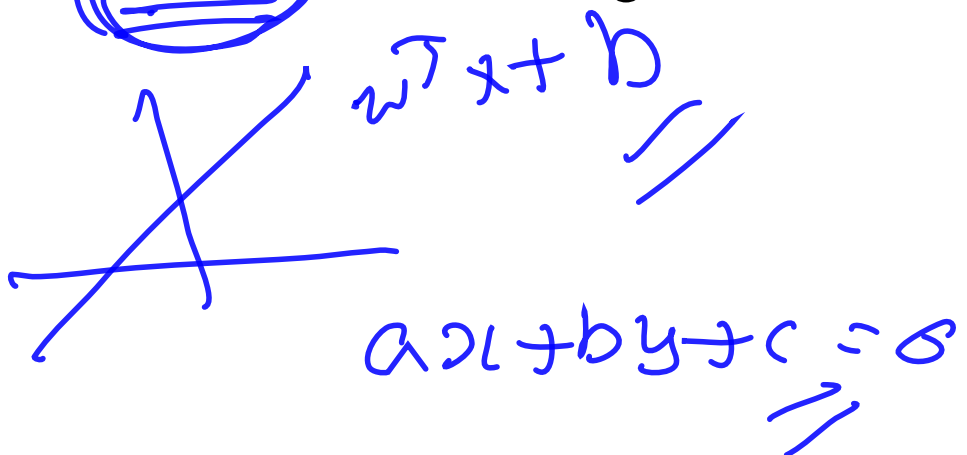
Discussions

$w = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ $x = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

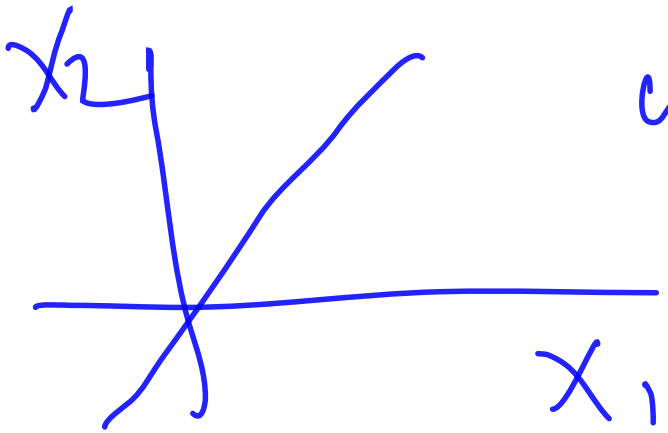
Q: $\mathbf{w}^T\mathbf{x}$ is a line passing through origin. How do we characterize "general" line?

- Ans1: $\mathbf{w}^T\mathbf{x} + b$
- Ans2: $\mathbf{w}'^T\mathbf{x}'$ where $\mathbf{x}'$ is an augmented vector. i.e., $\mathbf{x}' = [\mathbf{x}^T, 1]^T$. (There is notational convenience for this.)
- (Refer the usage of "bias" in neural networks. You will appreciate this better)
- In general, when we write as $\mathbf{w}^T\mathbf{x}$, often $\mathbf{x}$ is augmented and bias is absorbed.

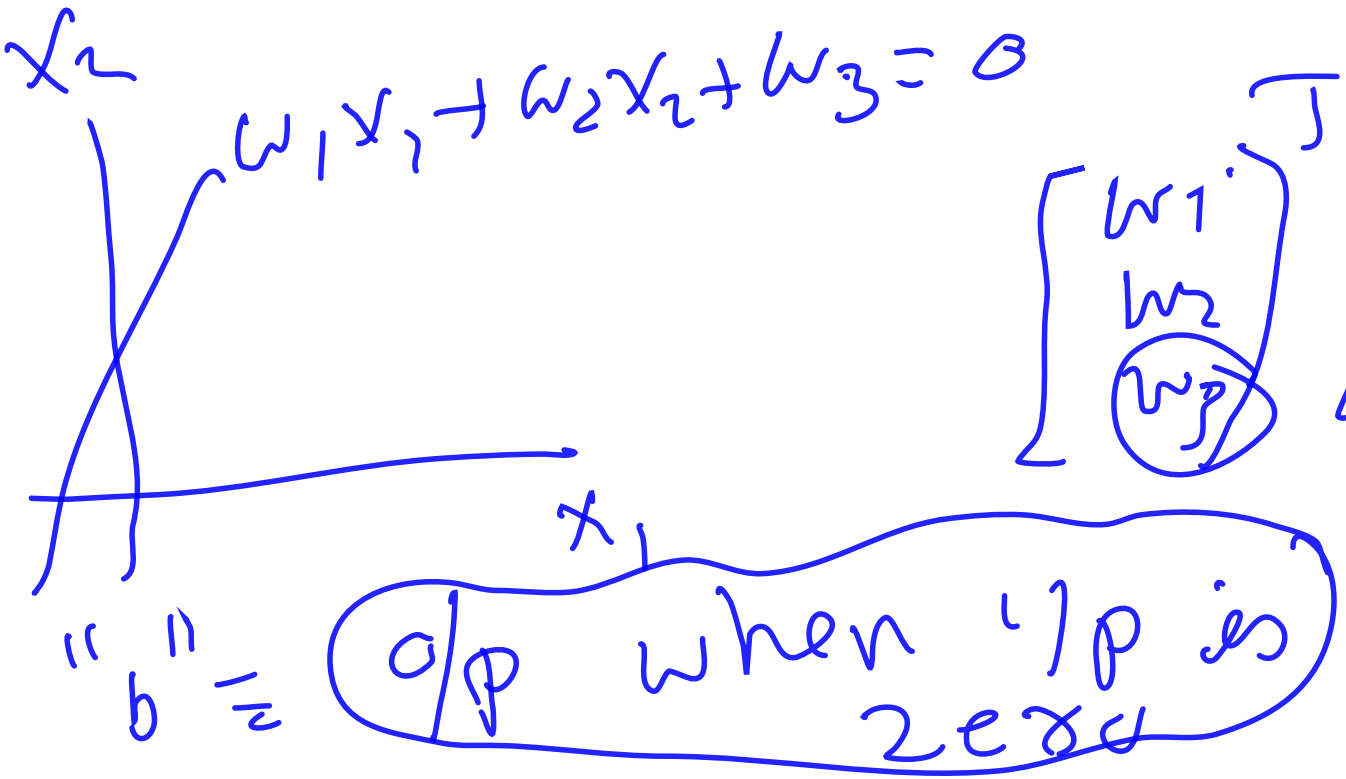




Blank



$w_1 x_1 + w_2 x_2 = 0$ $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$



$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = 0$

"b" = o/p when i/p is zero

$\begin{bmatrix} x \\ 1 \end{bmatrix}$

Feature Transformation

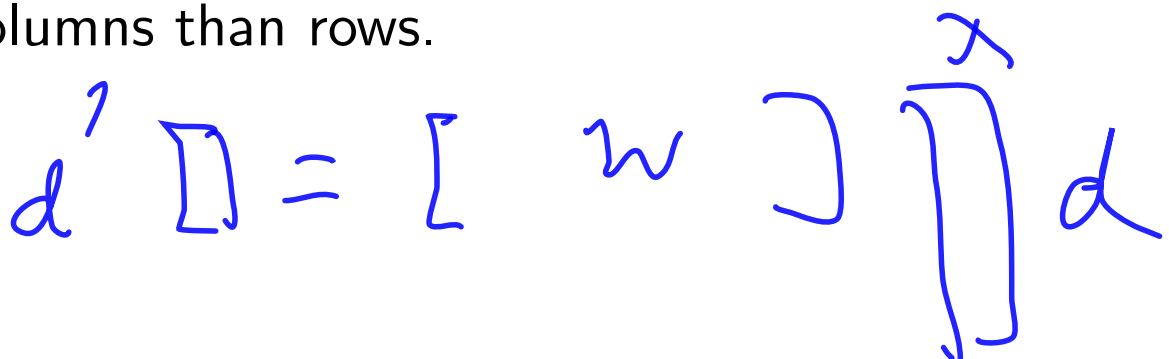
- In general, this feature transformation

$$\phi : \mathbf{x} \rightarrow \mathbf{x}'$$

  is a useful trick.
  - **Goal:** Get better features starting from "raw" measurements/features.
- The feature transformations like:

$$\mathbf{x}' = \mathbf{W}\mathbf{x}$$

  can lead to dimensionality reduction, when the matrix $\mathbf{W}$ has more columns than rows.



Blank

Review Question

What is the angle between the two lines characterized by

$$\mathbf{w}_1 = [1, 1]^T$$

$$\mathbf{w}_2 = [1, -1]^T$$

(a) 0 (b) 45 (c) 90 (d) 120

Review Question

We stop two people at random. What is the probability that they were born on the same day of the week?
(a) $\frac{1}{7}$ (b) $\frac{1}{7^2}$ (c) $\frac{1}{7+7}$ (d) $\frac{1}{2}$ (e) None of the above

What Next:?

- We will use this week (until 15 Aug) for streamlining the Course:
- Topics: Performance Metrics of ML solutions.
- Logistics:
  - Office Hours (Watch MS Team Announcements and Channels)
  - Start using Skikha (look for instructions in the Channels)
  - Formal Details on Moodle; TAs active (by Weekend or Early next week)
- Watch Channels in "Teams" regularly.