CS623: Introduction to Computing with Neural Nets *(lecture-8)*

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

Hardness of Training Feedforward NN

- NP-completeness result:
  - *Avrim Blum, Ronald L. Rivest: Training a 3-node neural network is NP-complete. Neural Networks 5(1): 117-127 (1992)*Showed that the *loading problem* is hard
- As the number of training example increases, so does the training time EXPONENTIALLY

Numerous problems have been proven to be NP-complete

- The procedure is always the same:
- Take *an instance* of a *known* NP-complete problem; let this be *p.*
- Show a *polynomial time Reduction* of *p* TO an instance *q* of the problem whose status is being investigated.
- Show that the answer to *q* is *yes*, if and only if the answer to *p* is *yes.*

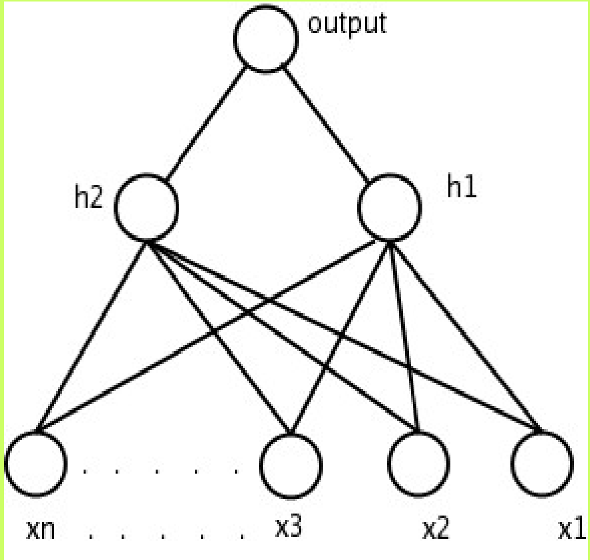
Training of NN

- Training of Neural Network is NP-hard
- This can be proved by the NP-completeness theory

- Question
  - Can a set of examples be loaded onto a Feed Forward Neural Network efficiently?

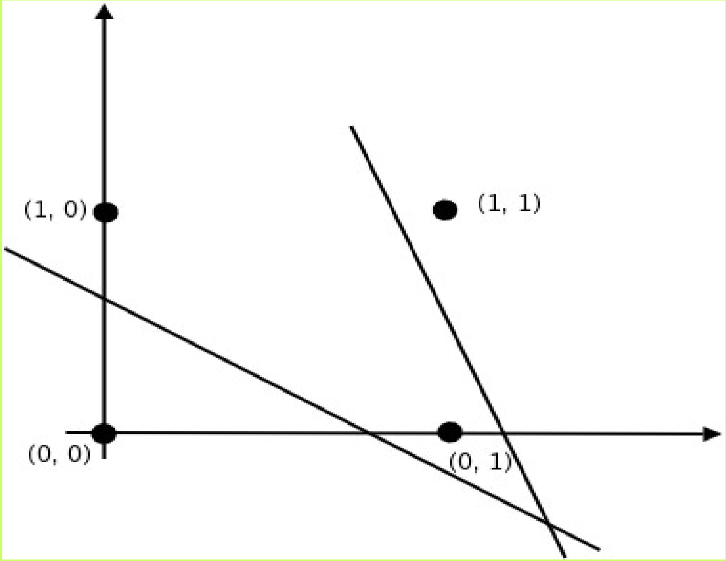
Architecture

- We study a special architecture.
- Train the neural network called 3-node neural network of feed forward type.
- ALL the neurons are 0-1 threshold neurons



Architecture

- h_1 and h_2 are hidden neurons
- They set up hyperplanes in the (n+1) dimensions space.



Confinement Problem

- Can two hyperplanes be set which confine ALL and only the positive points?
- Positive Linear Confinement problem is NP-Complete.
- Training of positive and negative points needs solving the CONFINEMENT PROBLEM.

Solving with Set Splitting Problem

- Set Splitting Problem
- Statement:
  - Given a set S of n elements e_1, e_2, ..., e_n and a set of subsets of S called as concepts denoted by c_1, c_2, ..., c_m, does there exist a splitting of S
  - i.e. are there two sets S_1 (subset of S) and S_2 (subset of S) and none of c_1, c_2, ..., c_m is subset of S_1 or S_2

Set Splitting Problem: example

- Example

  $S = \{s_1, s_2, s_3\}$

  $c_1 = \{s_1, s_2\}$, $c_2 = \{s_2, s_3\}$

Splitting exists

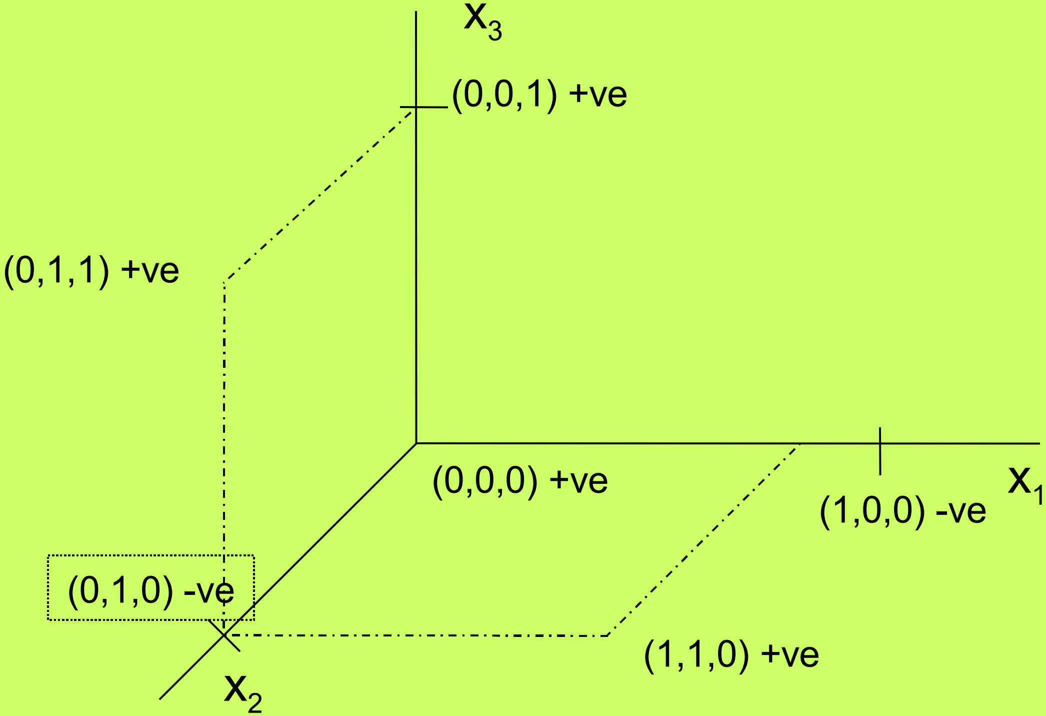
$S_1 = \{s_1, s_3\}$, $S_2 = \{s_2\}$

Transformation

- For *n* elements in S, set up an *n*-dimensional space.
- Corresponding to each element mark a negative point at unit distance in the axes.
- Mark the origin as positive
- For each concept mark a point as positive.

Transformation

- $S = \{s_1, s_2, s_3\}$
- $c_1 = \{s_1, s_2\}$, $c_2 = \{s_2, s_3\}$



Proving the transformation

- Statement
  - *Set-splitting problem has a solution if and only if positive linear confinement problem has a solution.*
- Proof in two parts: **if part** and **only if** part
- If part
  - *Given* Set-splitting problem has a solution.
  - *To show* that the constructed Positive Linear Confinement (PLC) problem has a solution
  - *i.e.* to show that since S_1 and S_2 exist, P_1 and P_2 exist which confine the positive points

Proof – If part

- P_1 and P_2 are as follows:
  - $P_1 : a_1x_1 + a_2x_2 + \ldots + a_nx_n = -1/2$ -- Eqn A
  - $P_2 : b_1x_1 + b_2x_2 + \ldots + b_nx_n = -1/2$ -- Eqn B

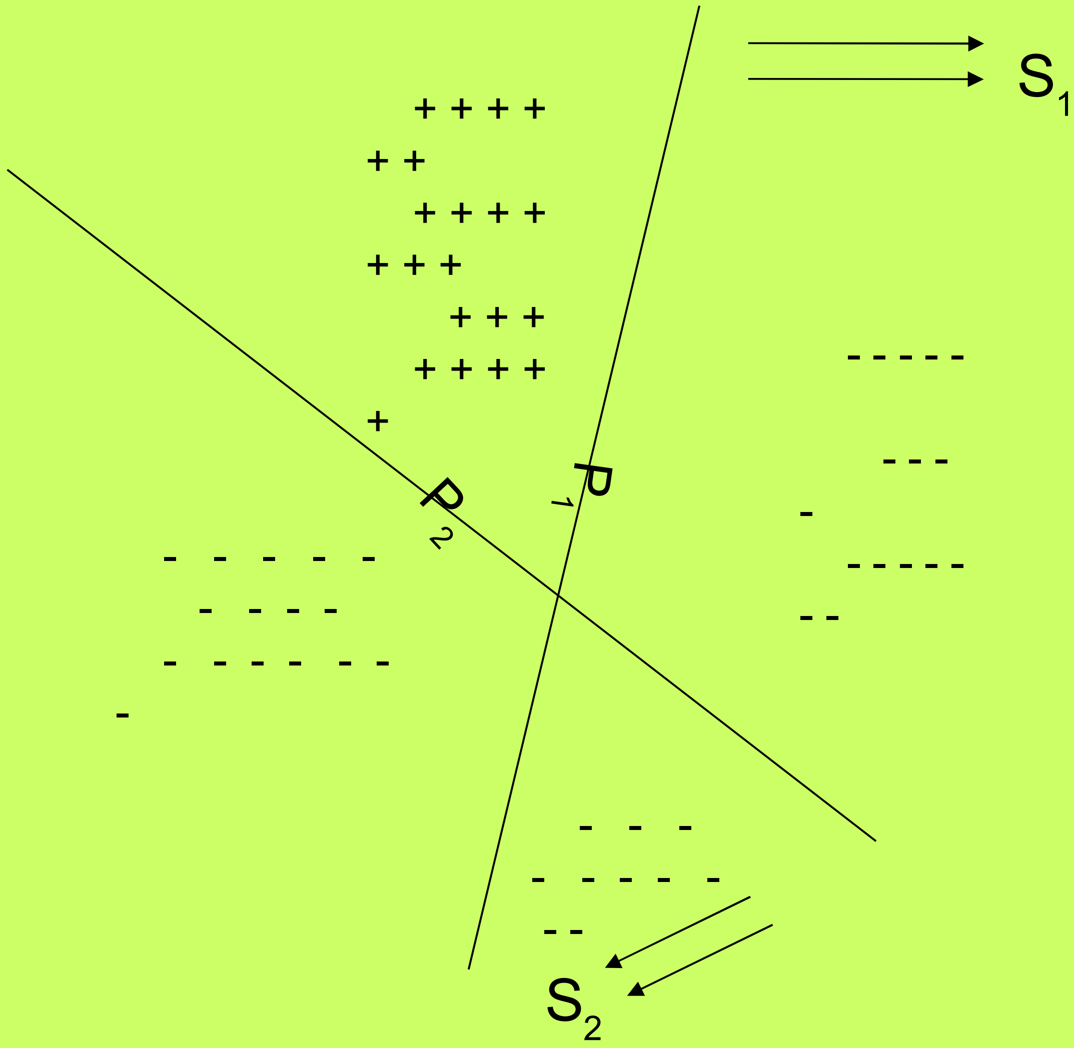
$$a_i = -1, \quad \text{if } s_i \in S_1$$

$$= n, \quad \text{otherwise}$$

$$b_i = -1, \quad \text{if } s_i \in S_2$$

$$= n, \quad \text{otherwise}$$

Representative Diagram



Proof (If part) – Positive points

- For origin (a +ve point), plugging in $x_1 = 0 = x_2 = ... = x_n$ into P_1 we get, $0 > -1/2$
- For other points
  - +ve points correspond to c_i's
  - Suppose c_i contains elements $\{s_1^i, s_2^i, ..., s_{ni}^i\}$, then at least one of the s_j^i cannot be in S_1

    $\therefore$ co-efficient of $x_j^i = n$,

    $\therefore$ LHS > -1/2
- Thus +ve points for each c_i belong to the same side of P_1 as the origin.
- Similarly for P_2.

Proof (If part) – Negative points

- -ve points are the unit distance points on the axes
  - They have only one bit as 1.
  - Elements in S_1 give rise to m_1 -ve points.
  - Elements in S_2 give rise to m_2 -ve points.
- -ve points corresponding to S_1
  - If $q_i \varepsilon S_1$ *then* x_i *in* P_1 *must have co-efficient -1*

    $\therefore LHS = -1 < -1/2$

What has been proved

- Origin (+ve point) is on one side of P_1
- +ve points corresponding to c_i's are on the same side as the origin.
- -ve points corresponding to S_1 are on the opposite side of P_1

Illustrative Example

- Example
  - S = $\{s_1, s_2, s_3\}$
  - $c_1 = \{s_1, s_2\}$, $c_2 = \{s_2, s_3\}$
  - Splitting : $S_1 = \{s_1, s_3\}$, $S_2 = \{s_2\}$
- +ve points:
  - (<0, 0, 0>,+), (<1, 1, 0>,+), (<0, 1, 1>,+)
- -ve points:
  - (<1, 0, 0>,-), (<0, 1, 0>,-), (<0, 0, 1>,-)

Example (contd.)

- The constructed planes are:
- P_1 :
  - $a_1x_1 + a_2x_2 + a_3x_3 = -1/2$
  - $-x_1 + 3x_2 - x_3 = -1/2$
- P_2:
  - $b_1x_1 + b_2x_2 + b_3x_3 = -1/2$
  - $3x_1 - x_2 + 3x_3 = -1/2$

Example (contd.)

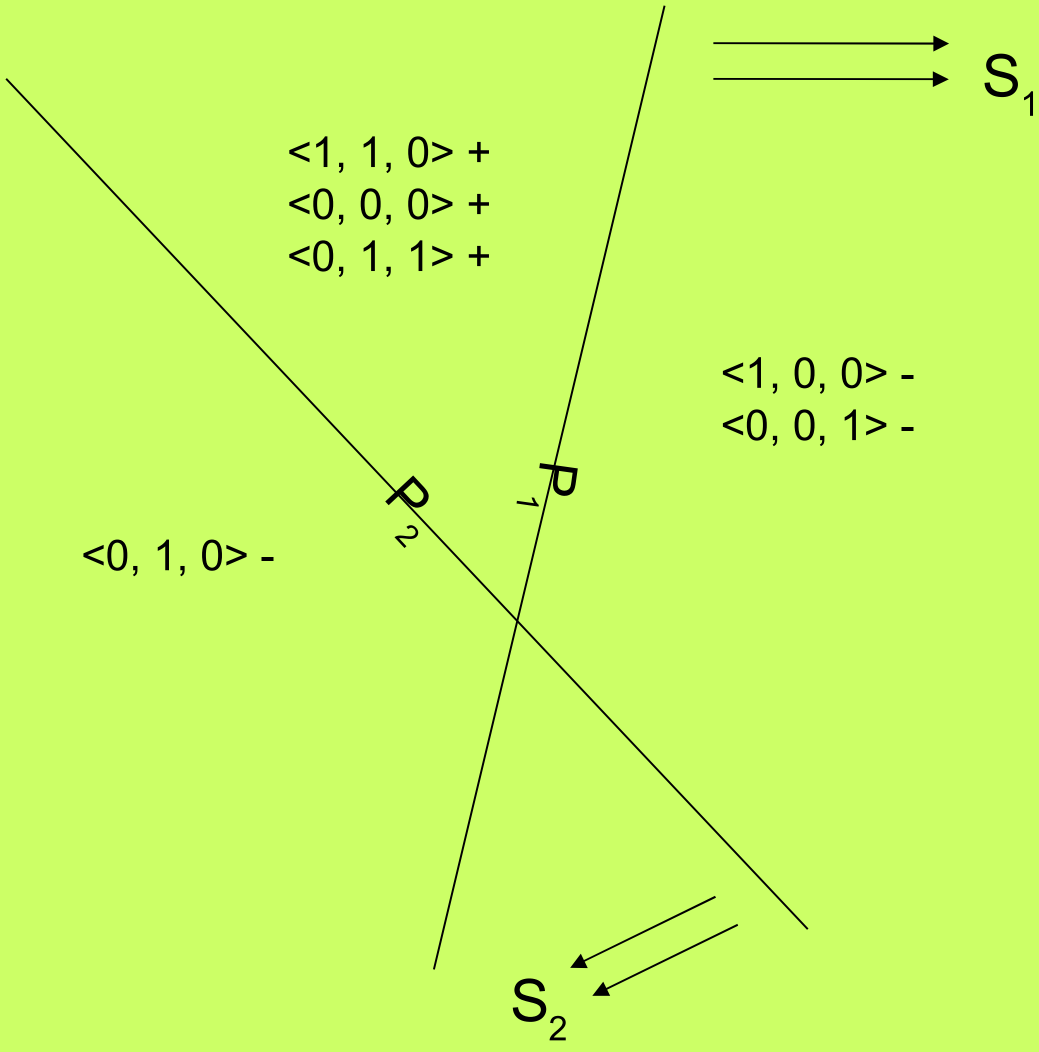
- P_1: $-x_1 + 3x_2 - x_3 = -1/2$
- <0, 0, 0>: LHS = 0 > -1/2,
  - ∴ <0, 0, 0> is +ve pt (similarly, <1,1,0> and <0,1,1> are classified as +ve)
- <1, 0, 0>: LHS = -1 < -1/2,
  - ∴ <1, 0, 0> is -ve pt
- <0, 0, 1>: LHS = -1 < -1/2,
  - ∴ <0, 0, 1> is -ve pt

But <0,1,0> is classified as +ve, i.e., cannot classify the point of S_2.

Example (contd.)

- P_2 : $3x_1 - x_2 + 3x_3 = -1/2$
- <0, 0, 0> : LHS = 0 > -1/2
  - ∴ <0, 0, 0> is +ve pt
- <1, 1, 0> : LHS = 2 > -1/2
  - ∴ <1, 1, 0> is +ve pt
- <0, 1, 1> : LHS = 2 > -1/2
  - ∴ <0, 1, 1> is +ve pt
- <0, 1, 0> : -1 < -1/2
  - ∴ <0, 1, 0> is -ve pt

Graphic for Example



Proof – Only if part

- Given +ve and -ve points constructed from the set-splitting problem, two hyperplanes P_1 and P_2 have been found which do positive linear confinement
- To show that S can be split into S_1 and S_2

Proof - Only if part (contd.)

- Let the two planes be:
  - P_1: $a_1x_1 + a_2x_2 + \ldots + a_nx_n = \theta_1$
  - P_2 : $b_1x_1 + b_2x_2 + \ldots + b_nx_n = \theta_2$
- Then,
  - S_1 = {elements corresponding to -ve points separated by P_1}
  - S_2 = {elements corresponding to -ve points separated by P_2}

Proof - Only if part (contd.)

- Since P_1 and P_2 *take care of* **all** -ve points, their union is equal to S ... (proof obvious)
- **To show:** No c_i is a subset of S_1 and S_2
- *i.e.*, there is in c_i at least one element $\notin S_1$

  -- Statement (A)

Proof - Only if part (contd.)

- Suppose $c_i \subset S_1$, then every element in c_i is contained in S_1
- *Let* $e_1^i, e_2^i, \ldots, e_{mi}^i$ *be the elements of* c_i *corresponding to each element*
- *Evaluating for each co-efficient, we get,*
  - $a_1 < \theta_1, \quad a_2 < \theta_1, \ \ldots, \quad a_{mi} < \theta_1$ *-- (1)*
  - *But* $a_1 + a_2 + \ldots + a_m > \theta_1$ *-- (2)*
  - *and* $0 > \theta_1$ *-- (3)*
- *CONTRADICTION*

What has been shown

- Positive Linear Confinement is NP-complete.
- Confinement on any set of points of one kind is NP-complete (easy to show)
- The architecture is special- only one hidden layer with two nodes
- The neurons are special, 0-1 threshold neurons, NOT sigmoid
- Hence, can we generalize and say that FF NN training is NP-complete?
- Not rigorously, perhaps; but strongly indicated