

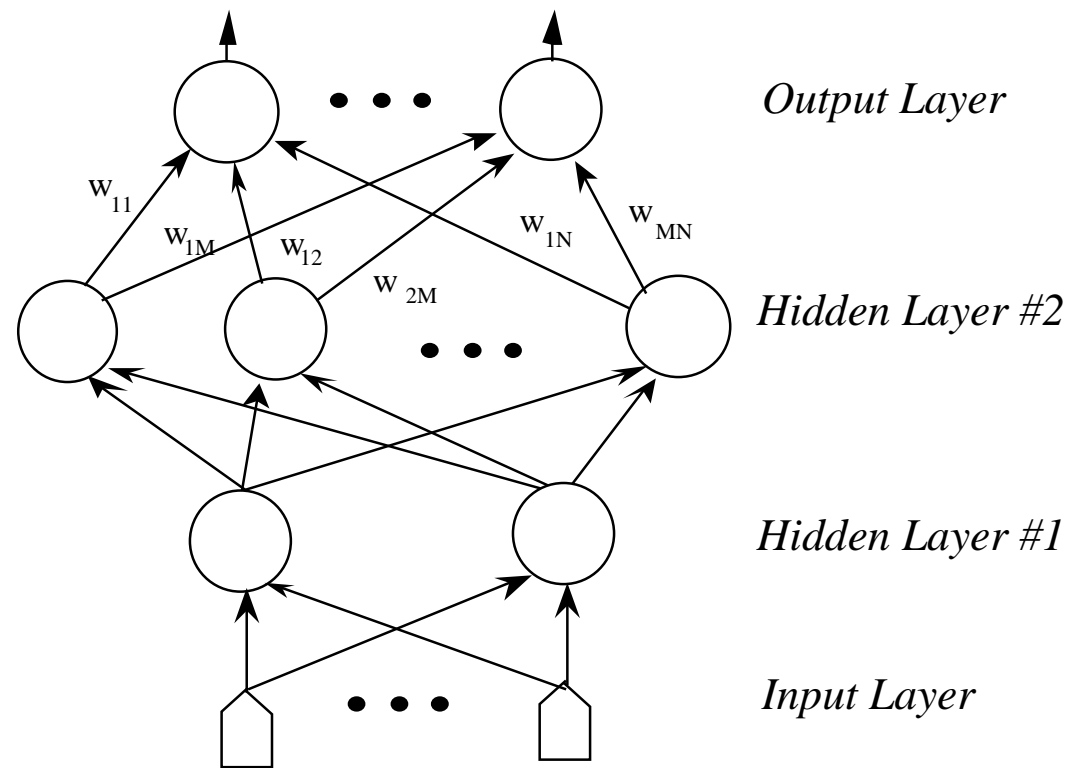
Lecture 9

MLP (I): Feed-forward Model

Outline

- Multi-Layer Perceptron Structure
- Feed Forward Model
- XOR Example
- MLP Applications

Multi-Layer Perceptron Structure



A Three Layer Feed-forward Multi-Layer Perceptron

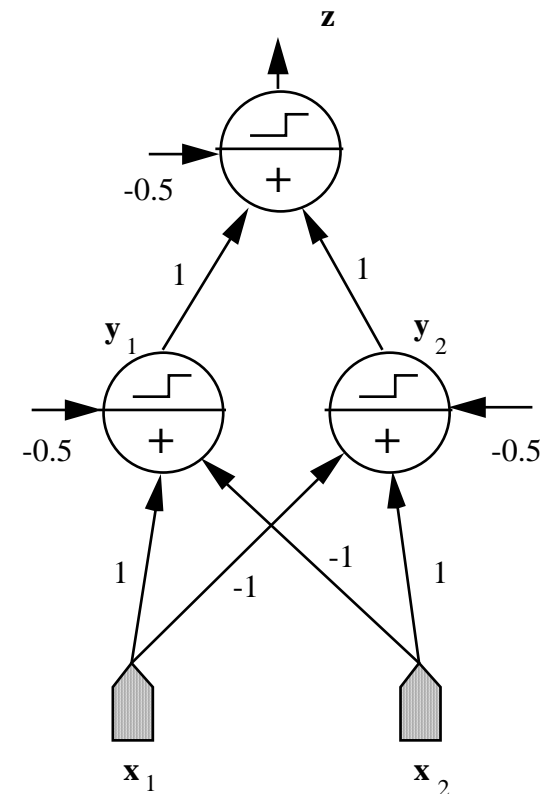
Two Layer Perceptron XOR Gate

Let $x_1, x_2 \in \{0, 1\}$, then

$$y_1 = \text{sgn}(x_1 - x_2 - 0.5) = x_1 \text{ AND } \bar{x}_2$$

$$y_2 = \text{sgn}(x_2 - x_1 - 0.5) = x_2 \text{ AND } \bar{x}_1$$

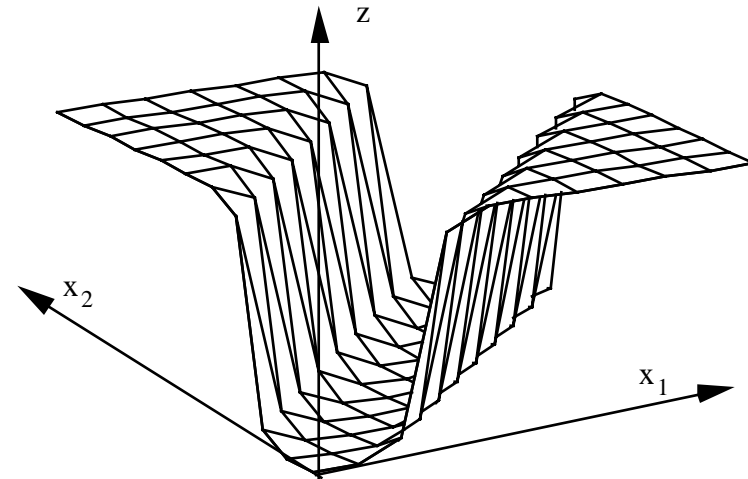
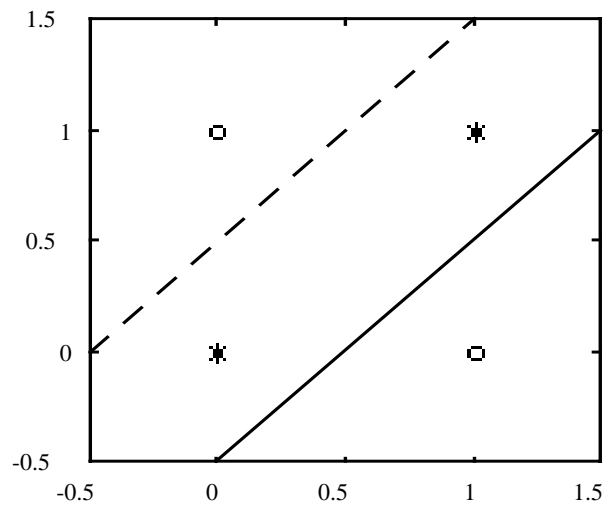
$$z = \text{sgn}(y_1 + y_2 - 0.5) = y_1 \text{ OR } y_2$$



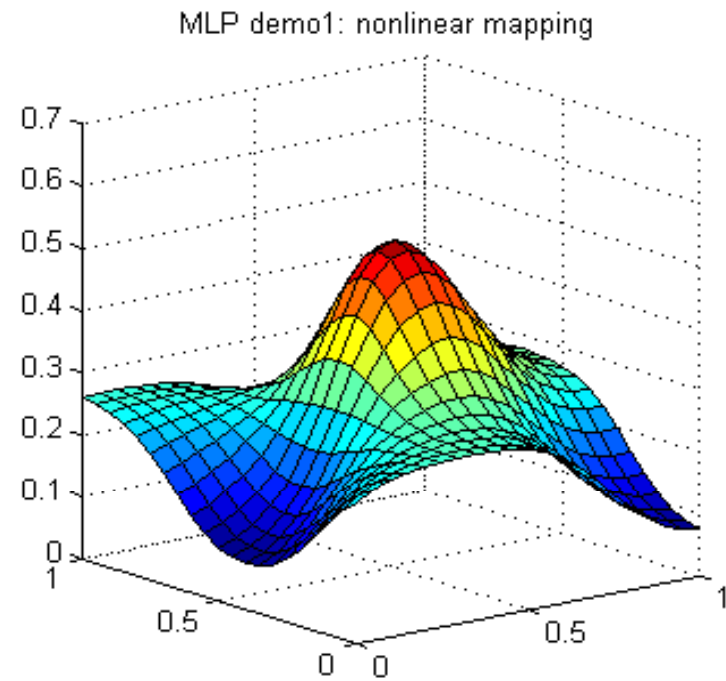
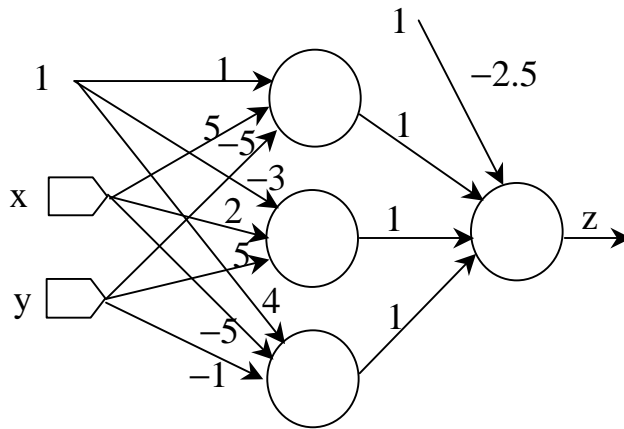
Decision Boundaries of XOR

- Linear Hyper-planes as decision boundaries

$$x_1 - x_2 - 0.5 = 0; \quad \text{and} \quad x_2 - x_1 - 0.5 = 0$$



MLP Nonlinear Mapping



MLP Feed-forward model

Notation

- (k) – Index of individual feature vectors, $1 \leq k \leq K$.
- (ℓ) -- Layer index, superscript, $0 \leq \ell \leq L$.
 $\ell = 0 \Rightarrow$ input layer, $\ell = L \Rightarrow$ output layer

- i, j – i^{th} and j^{th} neuron in each layer, subscript

Example:

- $z_i^{(\ell)}(k)$: the output of i^{th} neuron in the ℓ^{th} layer corresponding to the k^{th} feature vector.
- $w_{ij}^{(\ell)}$: the value of the synaptic weight that connect the output of the j^{th} neuron at $\ell - 1^{\text{th}}$ layer to the i^{th} neuron at the ℓ^{th} layer. The value of the weight is updated once every *epoch*.

MLP Feed-forward model

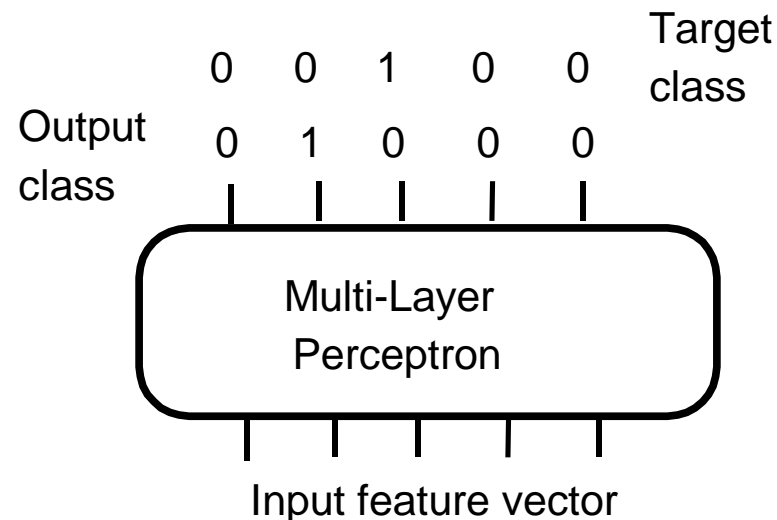
$$z_i^{(\ell)}(k) = f(u_i^{(\ell)}(k)) = 1/(1 + \exp[-u_i^{(\ell)}(k)])$$

$$u_i^{(\ell)}(k) = \sum_{j=1}^N w_{ij}^{(\ell)} a_j^{(\ell-1)}(k) + \theta_i^{(\ell-1)} = \sum_{j=0}^N w_{ij}^{(\ell)} z_j^{(\ell-1)}(k)$$

- Note that $z_0^{(\ell-1)}(k) \equiv 1$, and $w_{i0}^{(\ell)} \equiv \theta_i^{(\ell-1)}$
- The input layer usually consists of linear elements. Thus, a 2-layer MLP will have two layers of non-linear neurons: the hidden layer, and the output layer.

Applications to Classification

- Classification: Match output class to target class. MLP assigns each input feature vector to a membership of a particular class i .



Applications to Approximation

- Approximation (regression, modeling) :
Targets are real numbers instead of binary class membership.

