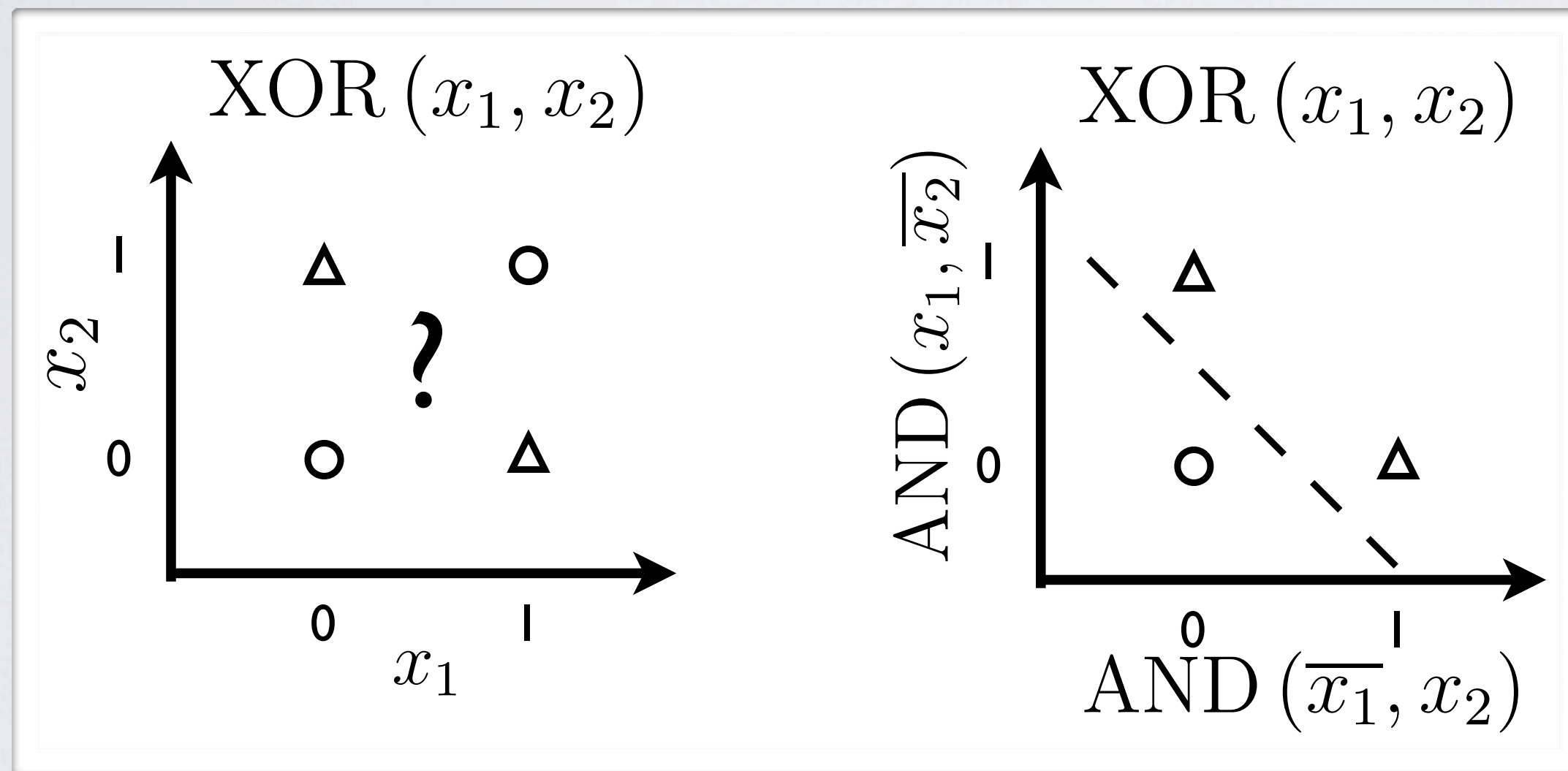# Neural networks

Feedforward neural network - capacity of neural network

ARTIFICIAL NEURON

Topics: capacity of single neuron

• Can't solve non linearly separable problems...

• ... unless the input is transformed in a better representation

NEURAL NETWORK

Topics: single hidden layer neural network

- Hidden layer pre-activation:

$$\mathbf{a}(\mathbf{x}) = \mathbf{b}^{(1)} + \mathbf{W}^{(1)}\mathbf{x}$$

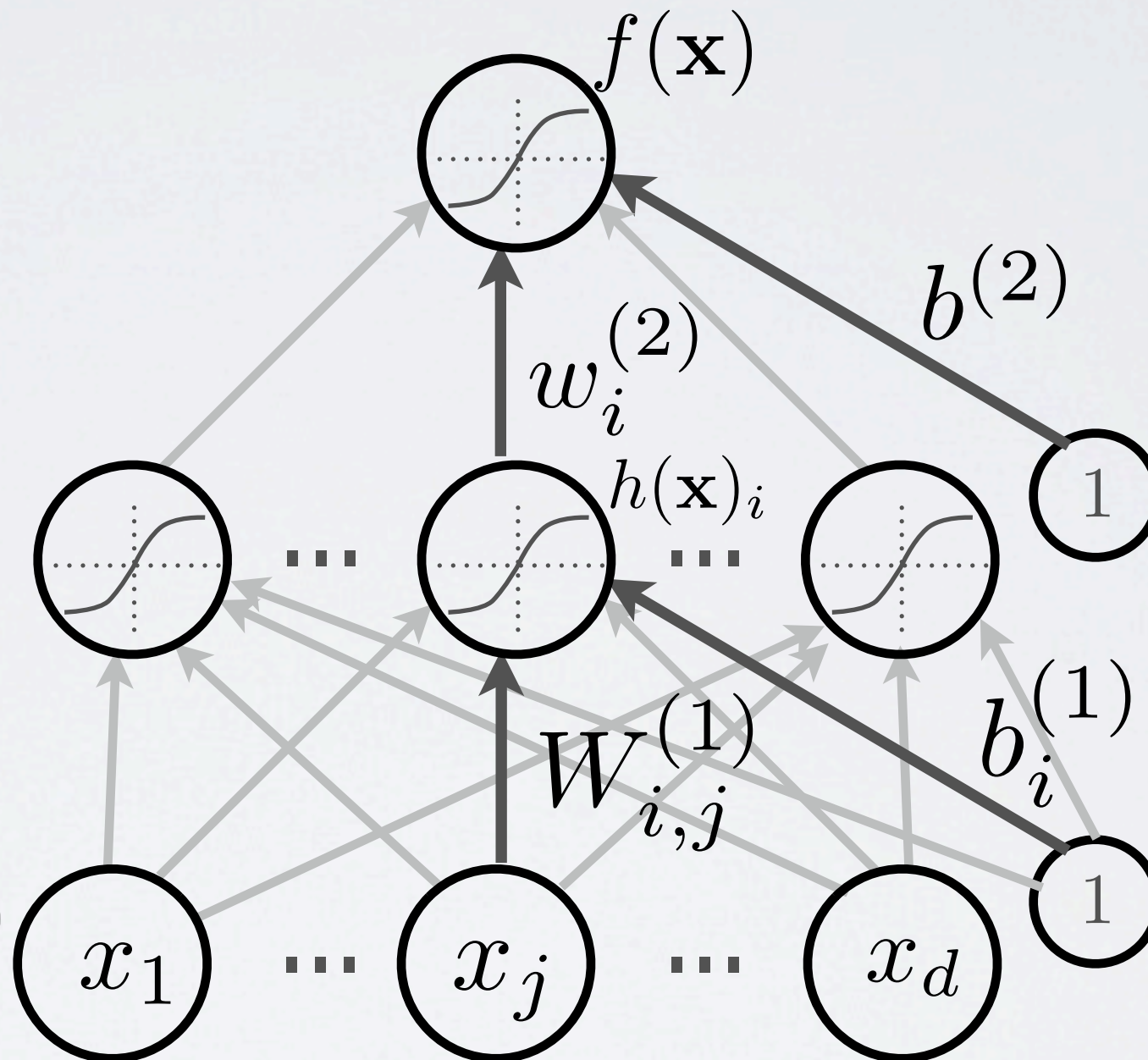$$\left(a(\mathbf{x})_i = b_i^{(1)} + \sum_j W_{i,j}^{(1)} x_j \right)$$

- Hidden layer activation:

$$\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{a}(\mathbf{x}))$$
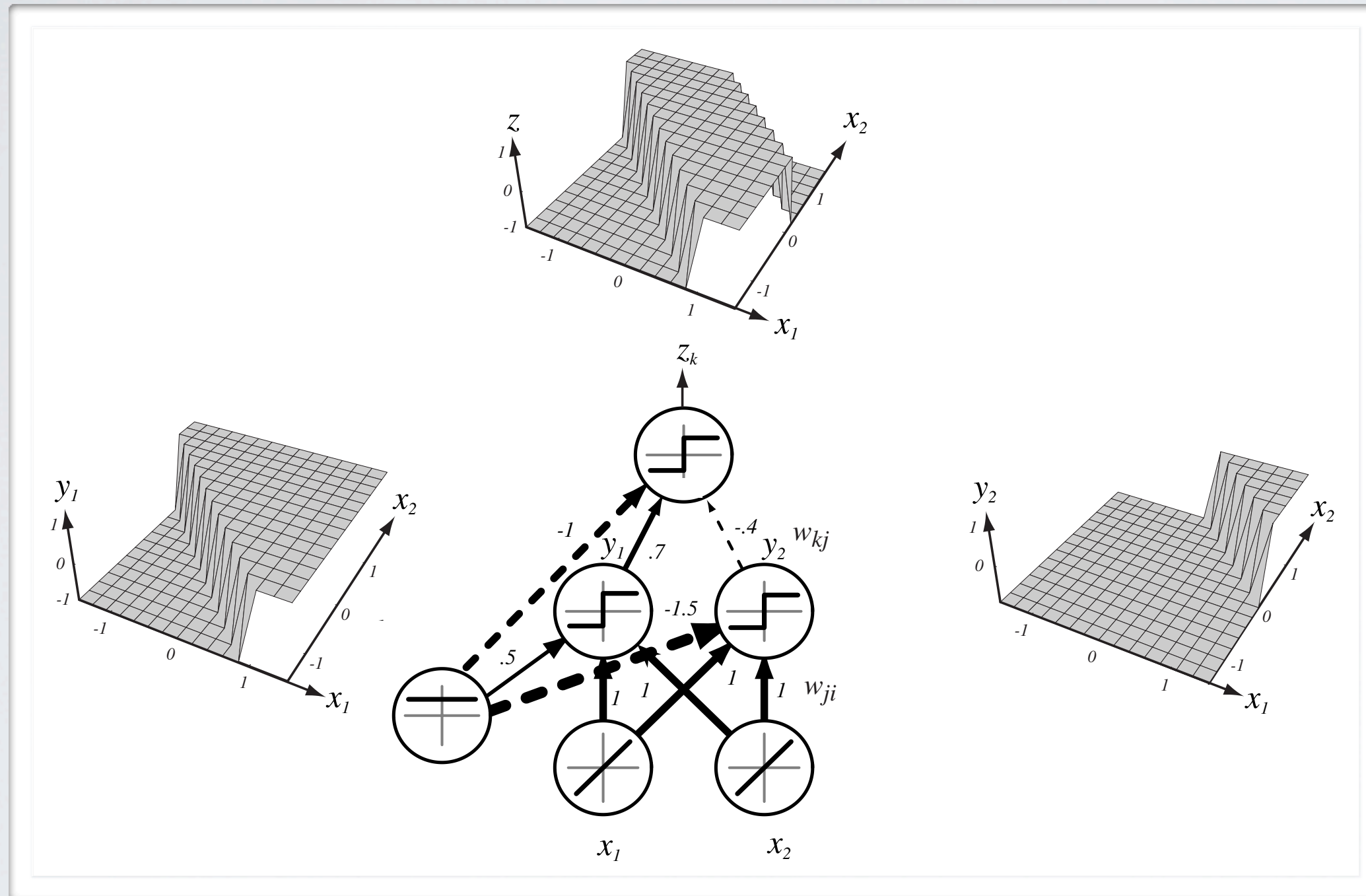
- Output layer activation:

$$f(\mathbf{x}) = o\left(b^{(2)} + \mathbf{w}^{(2)^\top}\mathbf{h}^{(1)}\mathbf{x} \right)$$

output activation function

$f(\mathbf{x})$

$b^{(2)}$

$w_i^{(2)}$

$h(\mathbf{x})_i$

1

$W_{i,j}^{(1)}$

$b_i^{(1)}$

1

$x_1 \quad \cdots \quad x_j \quad \cdots \quad x_d \quad 1$
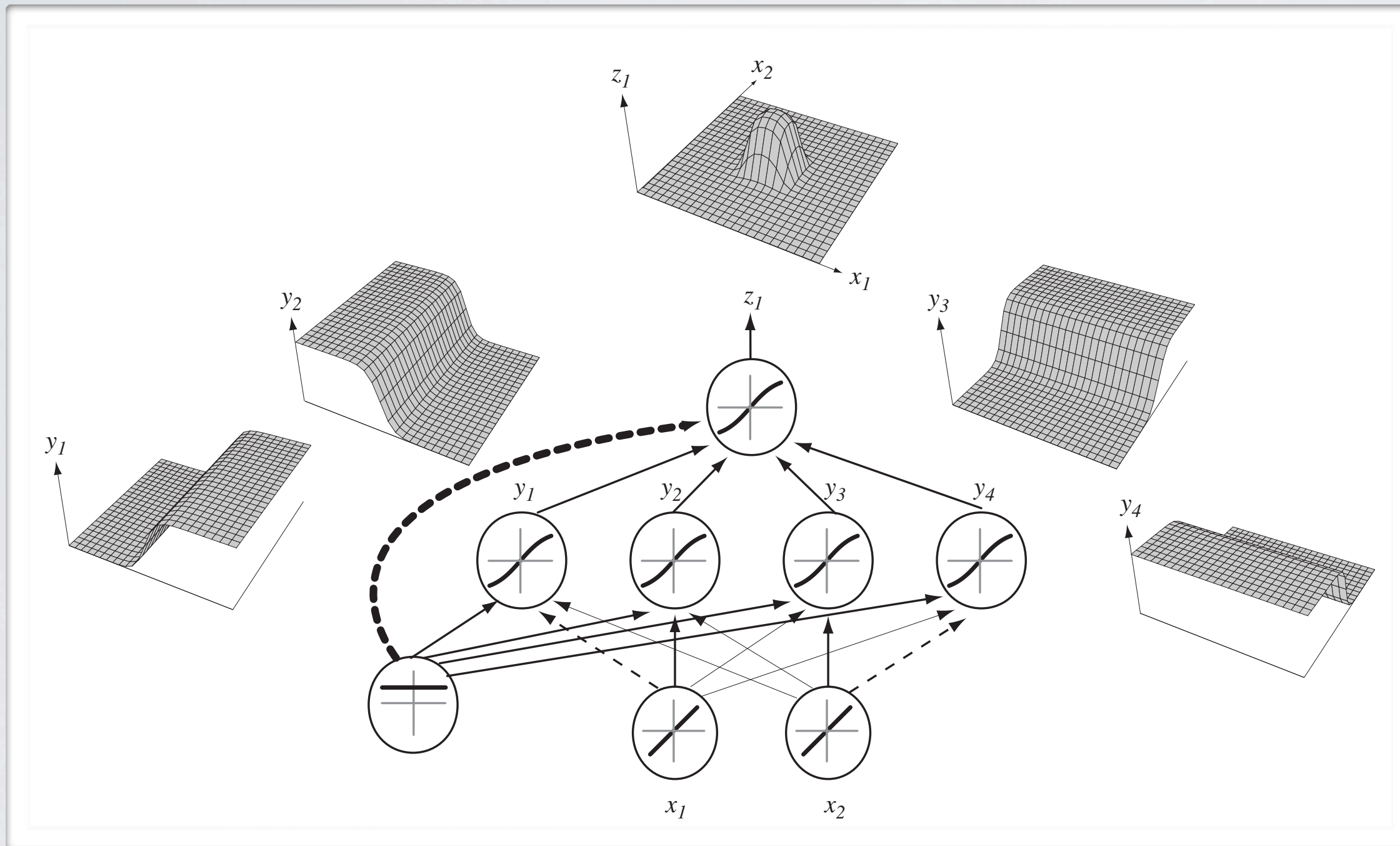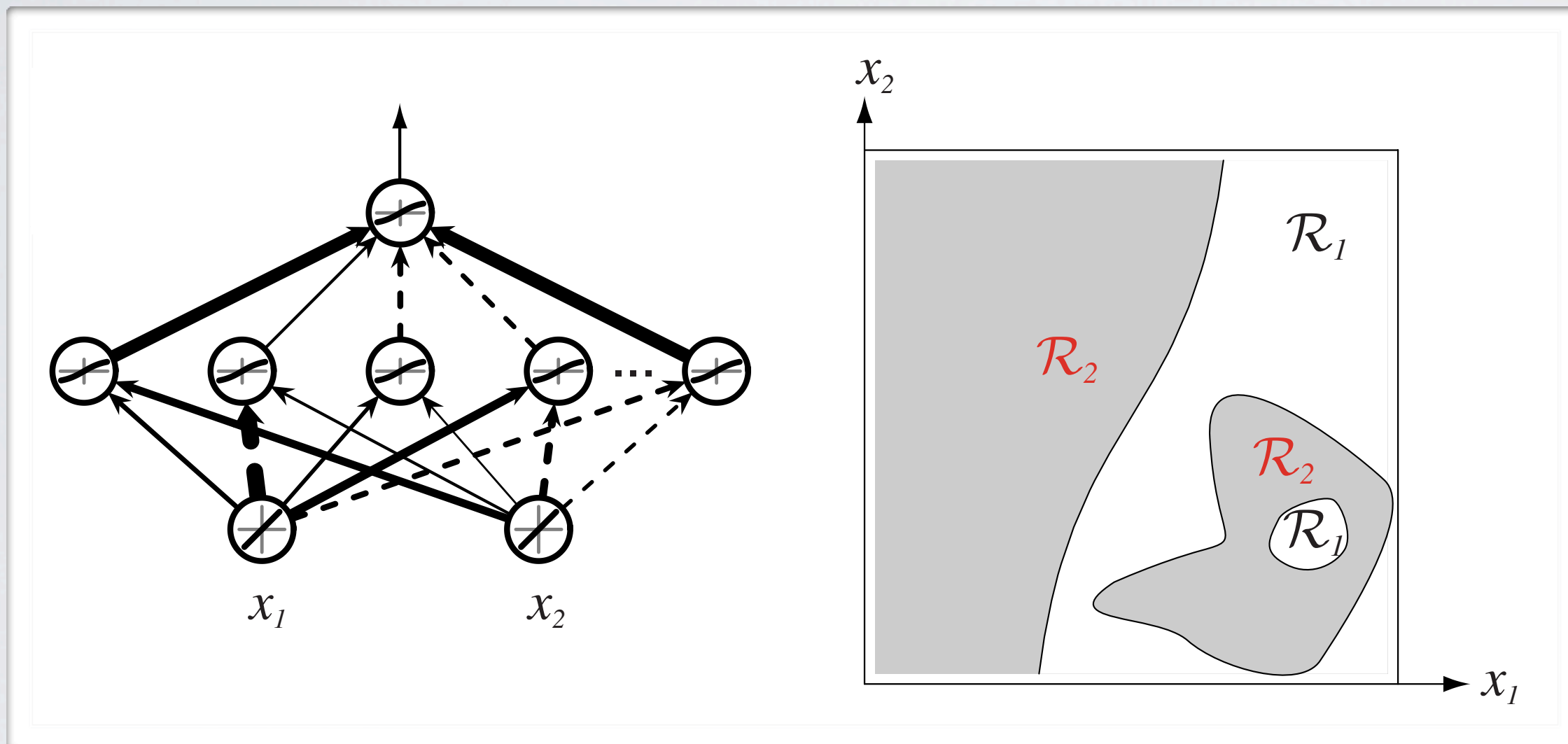
CAPACITY OF NEURAL NETWORK

Topics: single hidden layer neural network

(from Pascal Vincent's slides)

CAPACITY OF NEURAL NETWORK

Topics: single hidden layer neural network

(from Pascal Vincent's slides)

CAPACITY OF NEURAL NETWORK

Topics: single hidden layer neural network

(from Pascal Vincent's slides)

CAPACITY OF NEURAL NETWORK

Topics: universal approximation

• Universal approximation theorem (Hornik, 1991):

‣ "a single hidden layer neural network with a linear output unit can approximate any continuous function arbitrarily well, given enough hidden units"

• The result applies for sigmoid, tanh and many other hidden layer activation functions

• This is a good result, but it doesn't mean there is a learning algorithm that can find the necessary parameter values!