

Introduction to Neural Networks (CPE 520)

HW#13: Dec. 17, 2019

Q.1: Find the expression for the Inverse for the linear matrix equations (systems):

(a): $\mathbf{y}_{n \times 1} = \mathbf{A}_{n \times m} \mathbf{x}_{m \times 1}$ where $n < m$ and $\text{rank} = n$

(b): $\mathbf{y}_{n \times 1} = \mathbf{A}_{n \times m} \mathbf{x}_{m \times 1}$ where $n > m$ and $\text{rank} = m$

Q.2: (a) What is the difference between Ridge and LASSO regressions?
(b) Write the expressions for them.

Q.3: Convolve the following expressions:

$$\begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix} * \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

Q.4: A number of DNN architectures use 1x1 convolution, what does it mean and what is its purpose?

Q.5: Give examples of different pooling methods used in DNN.

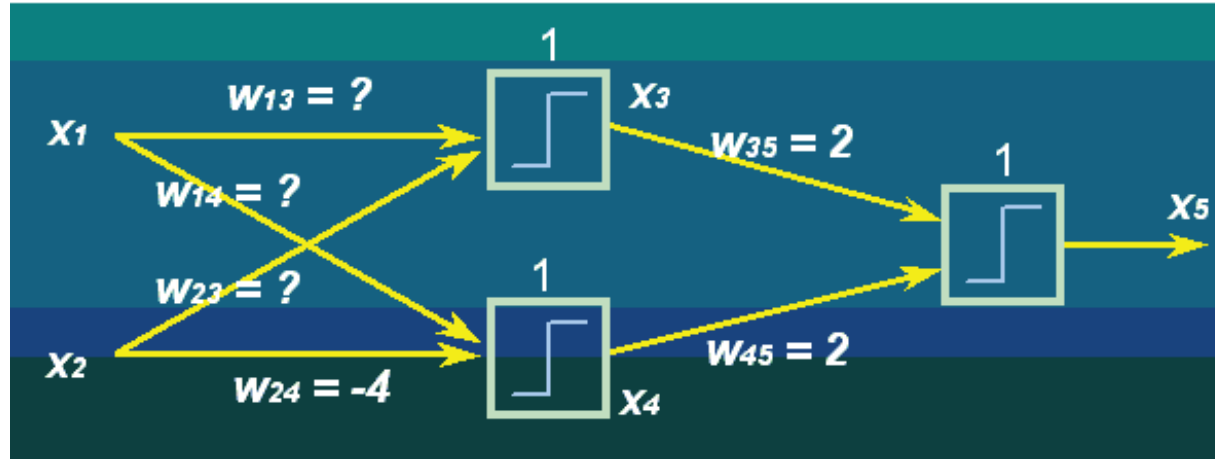
Q.6: What is the difference between logistic regression and Softmax?

Q.7: Write the **hinge loss** expression and in which classifier it is usually used.

Q.8: The common regularization in machine learning are based on L2 or L1 norm on the weights, write the expressions for these two norms.

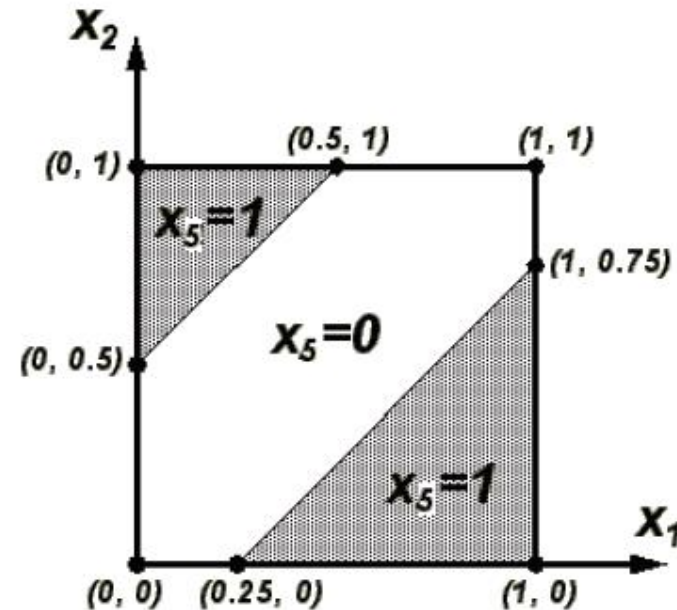
Q.9: In DNN we use auto-encoders, what are the usefulness & applications of developing autoencoders?

Q10. An 2-2-1 MLP has the following configuration:



Suppose that the above MLP has the following decision boundary for XOR problem:

Find the missing weights of the MLP.



Q11. Consider sigmoid function $f(NE T_j)$:

Find its gradient.

$$f(NE T_j) = \frac{1}{1 + e^{-\lambda NE T_j}}$$

Q.12 What is kernel function? What is the kernel trick?

Q.13 Given 2-dimensional vectors $\mathbf{x}_i = [x_1, x_2]^T$ for $i=1, 2, \dots, N$, let the kernel be a polynomial $k(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j)^2$

show that $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ where $\phi(\mathbf{x}_i) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T$