UNIVERSITY OF THESSALY
SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL
& COMPUTER ENGINEERING

Νευρο-Ασαφής Υπολογιστική (Neuro-Fuzzy Computing)
Fall Semester 2025-2026
Δημήτριος Κατσαρός (prof. Dimitrios Katsaros)

Problem-set 1st: **Teamwork** (2-persons)

Announcement: Friday, October 31, 2025
Delivery deadline: Sunday, November 30, 2025

SECTION 0: Warming-up with linear algebra, calculus, and automatic control

Problem-01

Plot the contour lines of the following function: $f(x,y) = x^4 + y^4 - 4xy + 1$. Then, find and characterize the real (local) minima/maxima of this function (Show your analytic calculations).

Problem-02

Execute two iterations of the Gradient Descent to the function $f(x,y)=6x^2-4xy+4y^2$ with initial point $\mathbf{x}_0=(3,2)$. Show your analytic calculations.

SECTION 1: Introduction to neural networks

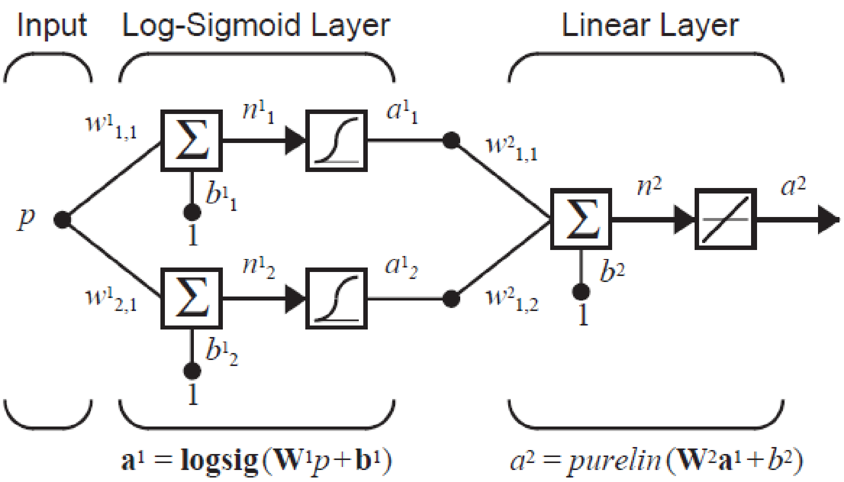
Problem-03

Express the derivative dS/dx, denoted as S', of the following activation functions S in terms of the original function S, i.e., determine φ such that S' = φ(S,x). [The first three functions comprise established activation functions, with c=1 known as logsig, tansig, and Google's Swish, respectively.]

- $$S = \frac{1}{1+e^{-x}}$$
- $$S = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$
- $$S = \frac{x}{1+e^{-x}}$$
- $$S = x \times \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$

Problem-04

Consider the following neural network



with $w^1_{1,1}=-2$, $w^1_{2,1}=-1$, $b^1_1=-2$, $b^1_2=-0.2$, $w^2_{1,1}=1.5$, $w^2_{1,2}=2.5$, $b^2=1$.

Sketch the following responses (plot the indicated variable versus p for $-2 < p < 2$).

A. a^1_1

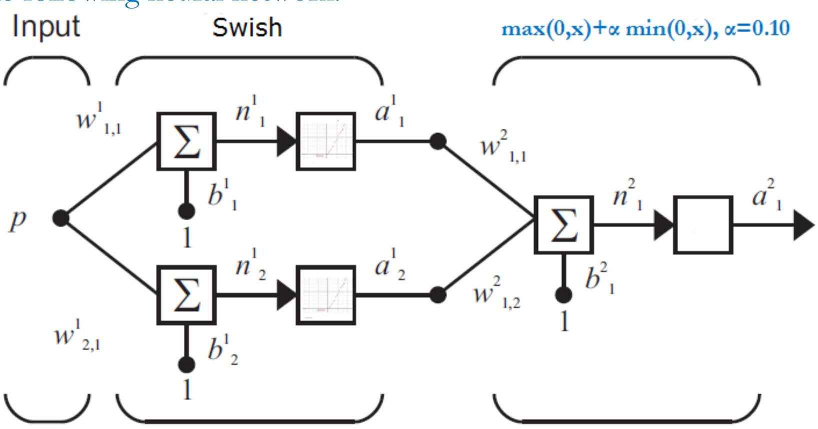
B. a^1_2

C. a^2

Then, change the *logsig* activation function with the *ReLU* activation function and sketch again the aforementioned responses.

Problem-05

Consider the following neural network:



Sketch the following responses (plot the indicated variable versus p for $-2 < p < 2$).

i. n^1_1

ii. a^1_1

iii. n^1_2

iv. a^1_2

v. n^2

vi. a^2

Initialization is as follows:

$w^1_{1,1} = -0.27, w^1_{2,1} = -0.41, b^1_1 = -0.48, b^1_2 = -0.13, w^2_{1,1} = 0.09, w^2_{1,2} = -0.17, b^2 = 0.48$

SECTION 2: Working with ADALINE neural networks

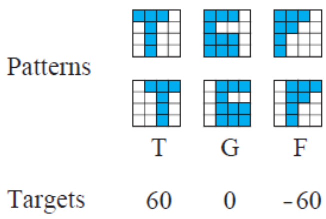
Problem-06

Suppose that you are given the following seven reference patterns and their categories: Class I consists of: $p_1=\{0, 0\}$, $p_2=\{0, 1\}$, Class II consists of: $p_3=\{1, 0\}$, $p_4=\{2, 0\}$, and Class III consists of: $p_5=\{-1, -1\}$, $p_6=\{0, -2.5\}$, $p_7=\{1.5, -1.5\}$. The probability of each vector $\mathbf{p}_1$, $\mathbf{p}_2$, is 0.25, and the probability of each vector $\mathbf{p}_3$, $\mathbf{p}_4$, $\mathbf{p}_5$, $\mathbf{p}_6$, and $\mathbf{p}_7$ is 0.1.

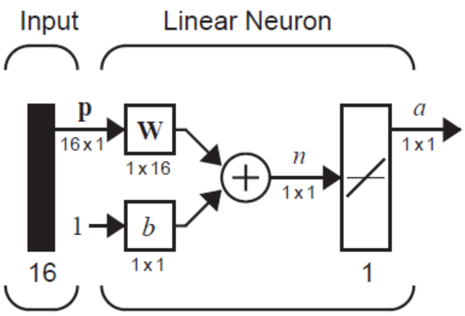
A. Select appropriate target (category) values.
B. Draw the network diagram for an ADALINE network with no bias that could be trained on these patterns.
C. Sketch the contour plot of the mean square error performance index.
D. Show the optimal decision boundary (for the weights that minimize mean square error), and verify that it separates the patterns into the appropriate categories.
E. Find the maximum stable learning rate for the LMS algorithm. Change the target values to opposite values, and see how this change affected the maximum stable learning rate?

Problem-07

Repeat the work of Widrow and Hoff on a pattern recognition problem from their classic 1960 paper (link-4 on the course's Webpage). They wanted to design a recognition system that would classify the six patterns shown below:



These patterns represent the letters T, G and F, in an original form on the top and in a shifted form on the bottom. The targets for these letters (in their original and shifted forms) are +60, 0 and -60, respectively. The objective is to train a network so that it will classify the six patterns into the appropriate T, G or F groups.
The blue squares in the letters will be assigned the value +1, and the white squares will be assigned the value -1. First we convert each of the letters into a single 16-element vector. We choose to do this by starting at the upper left corner, going down the left column, then going down the second column, etc.
Learning rate $\alpha=0.03$. Present the training patterns in a random sequence. You should use an ADALINE of the following topology:



$a = purelin(\mathbf{W}\mathbf{p}+b)$

You are required to draw a plot of the sum square error versus training steps. Each step is defined as the presentation of one input pattern to the neural network.
[Your plot should look similar to that in Fig.5 of Widrow-Hoff original paper.]

Problem-08

Suppose that we have the following two reference patterns and their targets:

$$\left\{\mathbf{p}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, t_1 = \begin{bmatrix} -1 \end{bmatrix}\right\}, \left\{\mathbf{p}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, t_2 = \begin{bmatrix} 1 \end{bmatrix}\right\}.$$

The vectors are equiprobable. We want to train an ADALINE network without a bias on this data set.

A. Sketch the contour plot of the mean square error performance index.

B. Sketch the optimal decision boundary.

C. Sketch the trajectory of the LMS algorithm on your contour plot. Assume a very small learning rate, and start with initial weights $\mathbf{W}(0) = [0 \;\; 1]$.

SECTION 3: Basics of MLPs

Problem-09

Write code to implement the backpropagation algorithm for a 1-S^1-1 MLP network (logsigmoid-linear). Write the program using matrix operations, as we did in the class lecture. Choose the initial weights and biases to be random numbers uniformly distributed between -0.5 and 0.5, and train the network to approximate the function

$$g(p)=1+\sin[p(\pi/3)] \text{ for } -2 \leq p \leq 2.$$

Use S^1= 2, S^1= 6, S^1= 10 and S^1= 20. Experiment with several different values for the learning rate α, and use several different initial conditions. Discuss the convergence properties of the algorithm as the learning rate changes.

Problem-10

The standard steepest descent backpropagation algorithm, which is summarized in the slide entitled "Summary of backpropagation algorithm" in Lecture-06, was designed to minimize the performance function that was the sum of squares of the network errors, as given in the last equation of slide 17 of Lecture-06. Suppose that we want to change the performance function to the sum of the fourth powers of the errors (e^4) plus the sum of the squares of the weights and biases in the network. Show how the equations in the slide entitled "Summary of backpropagation algorithm" will change for this new performance function. (You don't need to rederive any steps which are already derived in our lectures and do not change.)

Useful information:

The deadline is strict. It is possible to get an extension (up to 4 days), but you need to get the approval of the instructor, and this is going to cost a 10% penalty to the final grade of this Problem-Set. Delivery of the solutions' pdf (typeset or very easy to read handwritten) is done by email to dkatsar@uth.gr. The subject of the message should be **strictly**: CE418-Problem set 01: AEM1-AEM2

Symbol interpretation:

It requires "algorithmic" and/or mathematical thinking.

It requires the development of code (in any language/platform you wish). The final deliverable should contain: a) The solution of the exercise, and b) The implementation source code.