UNIVERSITY OF THESSALY
SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL
& COMPUTER ENGINEERING

Νευρο-Ασαφής Υπολογιστική
Χειμερινό Εξάμηνο 2020-2021
Δημήτριος Κατσαρός

Σειρά προβλημάτων: 1η: ΟΜΑΔΙΚΕΣ (2-ΑΤΟΜΩΝ) ΕΡΓΑΣΙΕΣ

Ημέρα ανακοίνωσης: Friday, October 30, 2020
Προθεσμία παράδοσης: Κυριακή, Νοέμβριος 08, 2020

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

Problem-01

Compute the eigenvalues and the corresponding eigenvectors of the following matrix:

$$\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$$

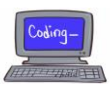
Show all of your computations (e.g., calculation of the characteristic polynomial), and not simply the final computed eigenpairs.

Problem-02

Plot the contour lines of the following function: $f(x,y) = x^3 - 4x - y^2$. Then, find and characterize the real (local) minima/maxima of this function (Show your analytic calculations).

Problem-03

Execute three iterations of the Gradient Descent to the function $f(x,y)=x^2+2x+y^2+4$ with initial point $\mathbf{x_0}=(2,1)$. Show your analytic calculations using the optimal λ in each iteration.



Problem-04

Consider the discrete dynamical system:

$$x_{n+1} = f(x_n) = (\mu - x^2) \times x$$

for x_0 values in [0,1.7], $\mu > 0$. Many dynamical systems transition into chaos as we increase a control or gain parameter such as μ.

A. Select $(\mu)=(2)$ and $(\mu)=(3)$ and $(\mu)=(4)$ with $x_0=0.3$, to generate $x_1,\ldots,x_{100}$. Plot the three trajectories. Are they aperiodic (chaotic) or periodic? How many fixed points exhibit each trajectory?

B. Repeat the above computations for the pair $(\mu)=(3.0)$ with $x_0=1.1$ and $(\mu)=(3.01)$ with $x_0=1.1$. Does a difference in 0.01 in the control parameter μ significantly affect the overall shape of the discrete trajectory? [Investigate whether we fall into chaos].

Problem-05

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-06

Consider the logistic activation function:

$$S(x) = \frac{1}{1+e^{-cx}}, \quad c > 0.$$

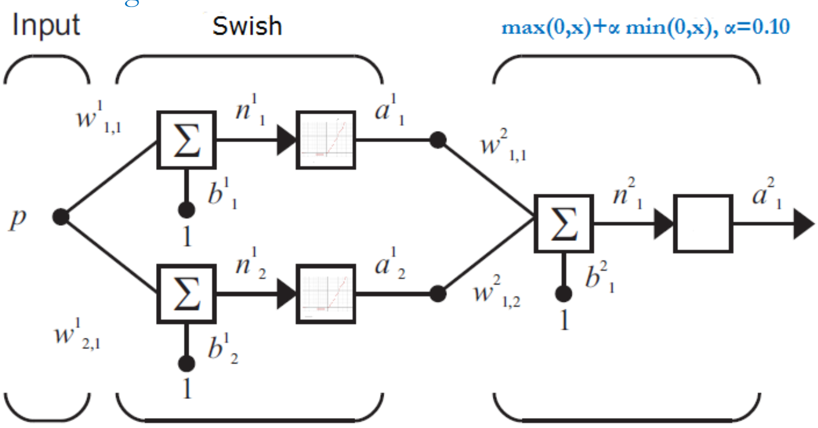
A. Solve this activation function S(x) for the activation x.
B. Investigate whether x strictly increases with S.
C. Explain why in general the "inverse function" x increases with S, if S'>0.

Problem-07

Compute the derivative of the activation function: **S(x)=max(0,x)+α min(0,x)**. Then, plot the function and its derivative into the same plot, for -∞ < x < ∞. Answer the previous questions for α=0.01 and for α=0.25.

Problem-08

Consider the following neural network:



Sketch the following responses (plot the indicated variable versus p for -3 < p < 3).

i. n_1^1

ii. a_1^1

iii. n_2^1

iv. a_2^1

v. n^2

vi. a^2

Initialization is as follows:
$w_{1,1}^1 = -0.27, w_{2,1}^1 = -0.41, b_1^1 = -0.48, b_2^1 = -0.13, w_{1,1}^2 = 0.09, w_{1,2}^2 = -0.17, b^2 = 0.48$
[Οι activation functions στο πρώτο layer είναι η Swish, ενώ στο δεύτερο είναι αυτή του Problem-07 με α=0.10.]

Problem-09

We have two categories of vectors. Category I consists of:

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \right\}$$

Category II consists of:

$$\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \right\}$$

A. Design (the architecture and find the weights and bias) a single-neuron perceptron network to recognize these two categories of vectors.
B. Draw the network diagram.
C. Sketch the decision boundary.
D. If we add the following vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to Category I, will your network classify it correctly? Demonstrate by computing the network response.
E. Can your weight matrix and bias be modified so your network can classify this new vector correctly (while continuing to classify the other vectors correctly)? Explain.

SECTION 2: Working with ADALINE neural networks

Problem-10

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-11

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

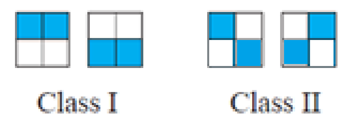
$$\left\{\mathbf{p_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, t_1 = 1\right\}, \left\{\mathbf{p_2} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, t_2 = -1\right\}$$

Suppose that the probability of vector $\mathbf{p_1}$ is 0.75 and that the probability of vector $\mathbf{p_2}$ is 0.25. We want to train an ADALINE network with a bias. We thus have three parameters to find $w_{1,1}$, $w_{1,2}$ and b_1.

A. Find the mean square error and the maximum stable learning rate.
B. Write a MATLAB M-file to implement the LMS algorithm for this problem. Take 40 steps of the algorithm for a stable learning rate. Use the zero vector as the initial guess. Sketch the final decision boundary.
C. Take 40 steps of the algorithm after setting the initial values of all parameters to 1. Sketch the final decision boundary.
D. Compare the final parameters and the decision boundaries from (B) and (C). Explain your results.

Problem-12

In the following figure two classes of patterns are given.



A. Use the LMS algorithm to train an ADALINE network to distinguish between patterns belonging to Class I and Class II (we want the network to identify horizontal and diagonal lines).
B. Can you explain why the ADALINE network might have difficulty with this problem, if it really has a difficulty?
C. [Try with/without biases, different initial values, and different target values. Record your observations and findings.]

SECTION 3: Introduction to fuzzy logic

Problem-13

Let P(x) and Q(x) be fuzzy truth functions, each of which can only give truth values of 0, 0.5 and 1. That is, for all x, P(x) is in the set {0, 0.5, 1} and Q(x) is in the set {0, 0.5, 1}.
Recall the Kleene-Dienes definition of implication: "a → b" is equivalent to

"(not a) or b"

Compute the truth table for the fuzzy statement "(P(x) and (P(x) → Q(x))) → Q(x)".

How does this compare for the same truth table in crisp logic (where P(x) and Q(x) can only be "true" or "false")?

Problem-14

Dimitris and Fany go to park if it is a *beautiful day* and it is not too *hot*, or if it isn't *raining*. Assuming that:

- It is a *beautiful day* with 0.7 degree
- It is *hot* with 0.3 degree
- It is *raining* with 0.6 degree

with which degree Dimitris and Fany will go to park?

Χρηστικές πληροφορίες:
Η προθεσμία παράδοσης είναι αυστηρή. Είναι δυνατή η παροχή παράτασης (μέχρι 3 ημέρες), αλλά μόνο αφού δώσει ο διδάσκων την έγκρισή του και αυτή η παράταση στοιχίζει 10% ποινή στον τελικό βαθμό της συγκεκριμένης Σειράς Προβλημάτων. Η παράδοση γίνεται με email στο dkatsar@e-ce.uth.gr του αρχείου λύσεων σε μορφή pdf (typeset). Το subject του μηνύματος **αυστηρά** πρέπει να είναι: CE418-Problem set 01: AEM1-AEM2

Ερμηνεία συμβόλων:


Δεν απαιτεί την χρήση υπολογιστή ή/και την ανάπτυξη κώδικα.

Απαιτεί την χρήση του Web για ανεύρεση πληροφοριών ή διεξαγωγή πειράματος.

Απαιτεί την ανάπτυξη κώδικα σε όποια γλώσσα προγραμμαστιμού ή Matlab. Το παραδοτέο θα περιέχει:

- Την λύση της άσκησης
- Τον πηγαίο κώδικα υλοποίησης