THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Machine Learning II
Homework #6
Due: *Oct-19-2020*

Homework 6:

- **Show ALL Work, Neatly and in Order.**
- **No** credit for Answers Without Work.
- Submit a single pdf file includes all of your solutions.
- **DO NOT** submit individual files or images.
- For coding questions, submit **ONE** .py file and include your comments.

E.1:

For the network shown below the initial weights and biases are chosen to be

$$w^1(0) = 1, b^1(0) = -2, w^2(0) = 1, b^2(0) = 1$$

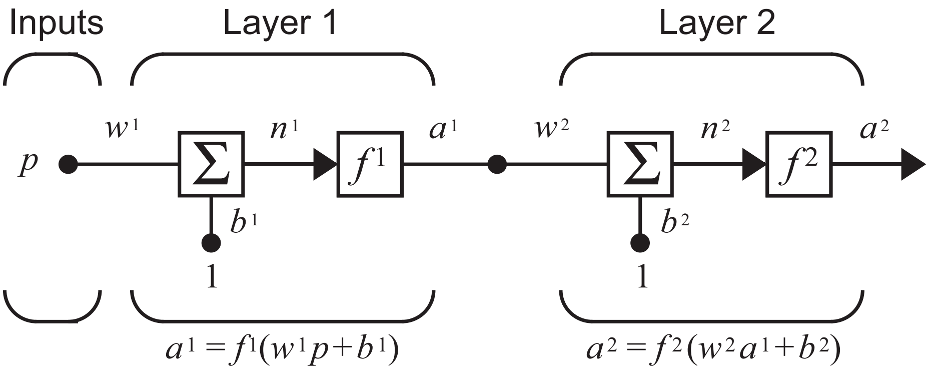
The network transfer functions are

$$f^1(n) = (n)^2, f^2(n) = \frac{1}{n}$$

and an input/target pair is given to be

$$\{p = 1, t = 1\}$$

Perform one iteration of backpropagation with $\alpha = 1$

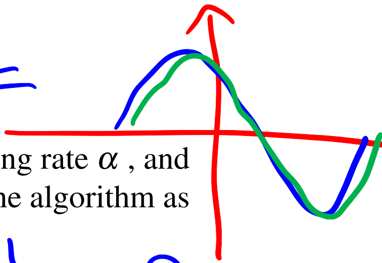


E.1 Python Exercise

Write a Python script to implement the backpropagation algorithm for a $1 - S^1 - 1$ network. Write the program using matrix operations, as in Eq. (11.41) to Eq. (11.47). Choose the initial weights and biases to be random numbers uniformly distributed between -0.5 and 0.5 (using the function rand), and train the network to approximate the function

$$g(p) = e^{-abs(p)} \times sin(\pi p) \text{ for } -2 \leq p \leq 2$$

⇐



Use $S^1 = 2$ and $S^1 = 10$. 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.

1-2-1

1-10-1

p

t

i. Plot the trained networks with the network outputs. Compare them.

ii.Plot squared error for each epochs.

deep learning

⇒

iii. Implement Stochastic gradient approach and repeat part i and ii.

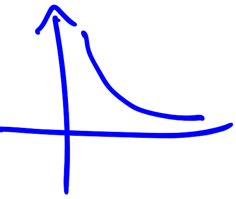
iv. Implement Batch approach (True Gradient) and repeat part i and ii.

v. Write your code in a format that you can enter any number of neurons in hidden layer.

Note: You are not allowed to use any ML packages, you are allowed to use basic python packages (numpy and etc). Please code the summary of backdrop equations (that is the only thing you need).

$R=1$ $S^1 = 2, 10$ $S^2 = 1$ $W^1 = (R\ )$ $b^1 = (R\ )$

$W^2 = (R)$ $b^2 = (R)$



For j in range epoch(len())

For i in range (len(p)): (iteration)

$n^1 = w^1 p[i] + b^1$

$a^1 = f(n^1)$

$n^2 = (w^2 a^2 + b^2)$

$a^2 = f(n^2)$

$e[i] = t[i] - a^2$

$s^2 = -2\dot{F}e$

$s^1 = \dot{F} w^2 s^2$

$w^2 =$ updat

$w^1 =$ upa..

$b^2 =$ updat

$b^1 =$ updat

$MSE = e^T e$

$s^2 a^1$

$s^1 p$

present ⇓ g ⇓ update

stocastic gradient

il

averagy

batch

accumulat gradient iteral

average ⇐