Math 6373 – Spring 2024 Name: SOLUTION

QUIZ #1

(1) [6 Pts]

(a) Determine the number of parameters of a feedforward neural network with the following architecture:

- Input layer: 2 neurons.
- 11 Hidden layers: 9 neurons each layer.
- Output layer: 1 neurons.

(b) Modify the network architecture above by changing the number of hidden layers to obtain a feedforward neural network containing at least 390 but no more than 410 parameters.

Solution:

(b) Since each pair of layers contributes $(N_\ell + 1) * N_{\ell+1}$ **parameters, observing that 11 hidden layers create 10 pairs of hidden layers, we have**

$$\textbf{Total parameters} \ = 3 * 9 + 10(10 * 9) + 10 * 1 = 937$$

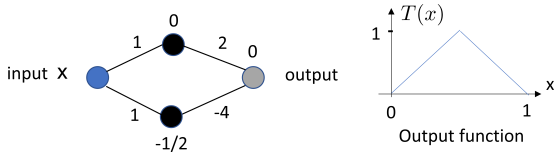
(c) We choose 5 hidden layers. Then we have

$$\textbf{Total parameters} \ = 3 * 9 + 4 * (10 * 9) + 10 * 1 = 397$$

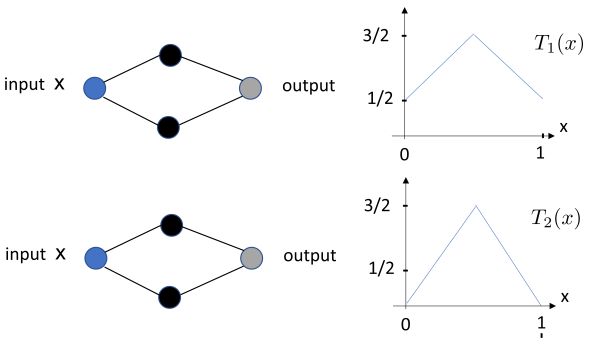
(2) [4 Pts] In class I have shown that a shallow neural network with ReLU activation function implementing the following function

$$T(x) = \begin{cases} 2x & \text{if } 0 \le x < \frac{1}{2} \\ 2(1-x) & \text{if } \frac{1}{2} \le x \le 1 \end{cases} \qquad x \in [0,1]$$

is obtained by writing as $T(x) = 2(x-0)_+ - 4(x-\frac{1}{2})_+$ which gives the architecture below



Modifying the example above write networks implementing the functions below



Solution:
The function T_1 is obtained by adding 1/2 to the function T

$$T_1(x) = T(x) + 1/2$$

Hence the only change to the original NN needed to generate the new NN implementation is the last bias at the output layer, which must be 1/2 (it was 0 for T).

For the function T_2, the slope of the first linear component is now 3 rather than 2. For the second linear component, the slope is -3, hence we need to set the coefficient -6 to compensate for the first linear component. Hence:

$$T_2(x) = 3(x-0)_+ - 6(x-\frac{1}{2})_+$$

An alternative argument to determine the coefficient of the second component is to impose the condition $T_2(1) = 0$. Hence the changes to the original NN needed to generate the new NN implementation are the wights to the output layer, which must be 3 and -6, where in T is was 2 and -4, respectively.