ISE 442 - Stochastic Models in Operations Research (Spring 2025)
Y. Hu

MIAMI

Homework 3 - Due Apr. 2 Midnight

Important: Please follow the restrictions and requirements described in the Homework Policy.

1. **[40 points] Reinforcement Learning.** Answering the following question and justify the reasons of your choice.

   (A) [10 points] Which of the following is true regarding model-free reinforcement learning? Select all options that apply.

   - ☐ Our goal is to learn the state transitions and rewards.
   - ☐ Our goal is to approximate policies and/or value functions without the underlying transition information.
   - ☐ Agents learn purely from experience without prior knowledge.
   - ☐ Model-free reinforcement learning typically exhibits higher sample efficiency.

   (B) [10 points] Which of the following are true regarding value function approximation in reinforcement learning? Select all options that apply.

   - ☐ We approximate the reward function.
   - ☐ We minimize an error involving the value function.
   - ☐ We maximize the value of the function approximation.
   - ☐ We approximate the value function.

   (C) [10 points] Which of the following are possible consequences of decreasing the step size in stochastic gradient descent? Select all that apply.

☐ More steps to algorithm completion.

☐ Fewer steps to algorithm completion.

☐ Larger gradient values.

☐ No change to gradient values.

(D) [10 points]Which of the following are possible consequences of increasing the number of neural network layers? Select all that apply.

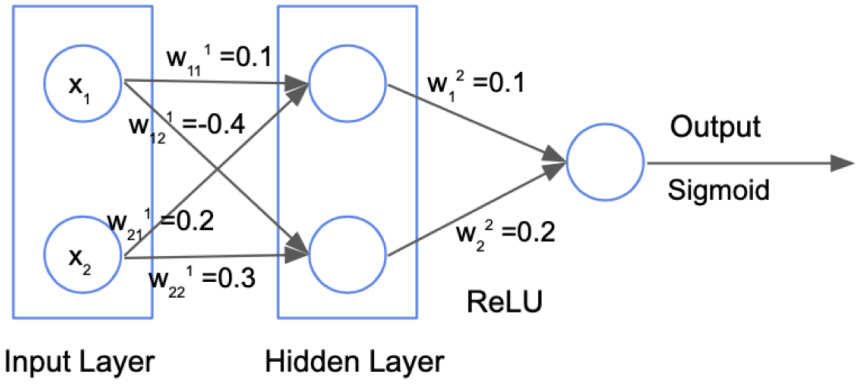
☐ Increased ability to represent an underlying function.

☐ Increased training time.

☐ Decreasing the amount of data required to train the network.

☐ Increasing the number of parameters of the network.

2. **[60 points] Backpropagation.** Consider the following two-layer neural network:



You will use the ReLU function as the activation function at the hidden layer, and the sigmoid

activation function at the output layer. Thus, the overall neural network is

$$f(x) = \sigma\left(W_2^\top \operatorname{ReLU}\left(W_1 x\right)\right) \quad \text{where} \quad W_1 = \begin{bmatrix} 0.1 & 0.2 \\ -0.4 & 0.3 \end{bmatrix}, \quad W_2 = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}.$$

The input vector is $x = [5, 4]^\top$. All weights are displayed on the image (the superscript denotes the layer). Suppose that the loss function is $\mathcal{L}(\text{output}) = \text{output}$.

- **[30 points] Written parts.** Compute the output of this neural network, and the gradient of this loss with respect to each of the weights. Be sure to show details.
- **[30 points] Programming parts.** Construct this neural network in Python. Build a forward function to get output, and a backpropagation function to get gradients of each weight. **Print out** the outputs and gradients, given the input vector x and weights displayed in the figured.

(As a reminder, the ReLU function and its derivative are:

$$\operatorname{ReLU}(z) = \begin{cases} z & \text{if } z \geq 0 \\ 0 & \text{otherwise} \end{cases}$$
$$\partial_z \operatorname{ReLU}(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the sigmoid function and its derivative are:

$$\sigma(z) = \frac{1}{1 + \exp(-z)} \qquad)$$
$$\partial_z \sigma(z) = \sigma(z) \times (1 - \sigma(z)).$$