DSC 190 - Homework 07

Due: Wednesday, May 18

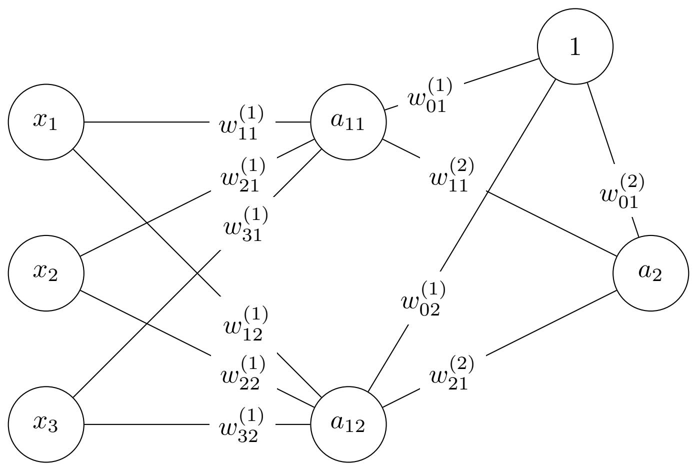
Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Unless otherwise noted by the problem's instructions, show your work or provide some justification for your answer. Homeworks are due via Gradescope at 11:59 PM.

Problem 1.

In lecture it was said that a neural network with linear activation functions is a linear prediction function, meaning that its decision boundary will also be linear. If we wish to have a non-linear decision boundary, we must introduce non-linearities with, for instance, non-linear activation functions.

In this problem, we'll see concretely that a neural network with linear activations is again linear.

Consider the neural network shown below.



The inputs x_1, x_2, and x_3 are numbers. Each $w_{ij}^{(k)}$ is a scalar weight. a_{ij} denotes the output of a neuron. Remember that when linear activations are used, the output of a neuron is simply the weighted sum of its inputs. So for instance:

$$a_{11} = w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2 + w_{31}^{(1)} x_3 + w_{01}^{(1)}$$

The node labeled 1 is the bias input. a_2 is the output of the neural network overall.

a) Write the output of the network, a_2, as an expression involving only the inputs x_1, x_2, x_3 and the weights, $w_{ij}^{(k)}$. a_i should not appear in your expression.

Solution: We have

$$a_{11} = w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2 + w_{31}^{(1)} x_3 + w_{01}^{(1)}$$
$$a_{12} = w_{12}^{(1)} x_1 + w_{22}^{(1)} x_2 + w_{32}^{(1)} x_3 + w_{02}^{(1)}$$

Therefore:

$$
\begin{aligned}
a_2 &= w_{11}^{(2)} a_{11} + w_{21}^{(2)} a_{12} + w_{01}^{(2)} \\
&= w_{11}^{(2)} \left(w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2 + w_{31}^{(1)} x_3 + w_{01}^{(1)} \right) \\
&\quad + w_{21}^{(2)} \left(w_{12}^{(1)} x_1 + w_{22}^{(1)} x_2 + w_{32}^{(1)} x_3 + w_{02}^{(1)} \right) \\
&\quad + w_{01}^{(2)}
\end{aligned}
$$

b) Show that the output of the network can be written

$$a_2 = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3,$$

where w_0, w_1, w_2, and w_3 are scalars that depend only on the weights in the original network. By showing this, you're proving that the network above is equivalent to a much simpler linear perceptron.

Solution: Starting from the result of the last subproblem:

$$
\begin{aligned}
a_2 &= w_{11}^{(2)} \left(w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2 + w_{31}^{(1)} x_3 + w_{01}^{(1)} \right) \\
&\quad + w_{21}^{(2)} \left(w_{12}^{(1)} x_1 + w_{22}^{(1)} x_2 + w_{32}^{(1)} x_3 + w_{02}^{(1)} \right) \\
&\quad + w_{01}^{(2)}
\end{aligned}
$$

Grouping the terms involving x_1, x_2, and x_3 separately:

$$
\begin{aligned}
&= \left(w_{11}^{(2)} w_{11}^{(1)} + w_{21}^{(2)} w_{12}^{(1)} \right) x_1 \\
&\quad + \left(w_{11}^{(2)} w_{21}^{(1)} + w_{21}^{(2)} w_{22}^{(1)} \right) x_2 \\
&\quad + \left(w_{11}^{(2)} w_{31}^{(1)} + w_{21}^{(2)} w_{32}^{(1)} \right) x_3 \\
&\quad + w_{01}^{(2)} \\
&= w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3
\end{aligned}
$$

where

$$
\begin{aligned}
w_1 &= w_{11}^{(2)} w_{11}^{(1)} + w_{21}^{(2)} w_{12}^{(1)} \\
w_2 &= w_{11}^{(2)} w_{21}^{(1)} + w_{21}^{(2)} w_{22}^{(1)} \\
w_3 &= w_{11}^{(2)} w_{31}^{(1)} + w_{21}^{(2)} w_{32}^{(1)} \\
w_0 &= w_{01}^{(2)}
\end{aligned}
$$