MA 225 MULTIVARIABLE CALCULUS
APPLICATION OF DOT PRODUCT TO AI

SIU-CHEONG LAU

Deep learning is a crucial technique in Artificial Intelligence. It is based on a model of the neural network in human brain.

A neuron is modeled by the following ***multivariable function***. Let's say the neuron receiving three input signals $x_1, x_2, x_3 \in \mathbb{R}$. We group these as a ***vector***

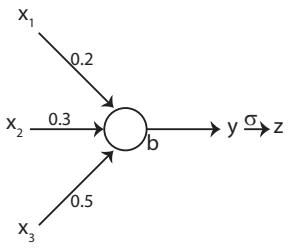
$$\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

The neuron assigns a fixed 'weight' to each input, representing the relative importance of the input. The weights are grouped into a vector $\vec{w} \in \mathbb{R}^3$. The neuron also has a 'threshold value' $b \in \mathbb{R}$. Then the neuron applies the ***dot product*** and compares with the threshold value:

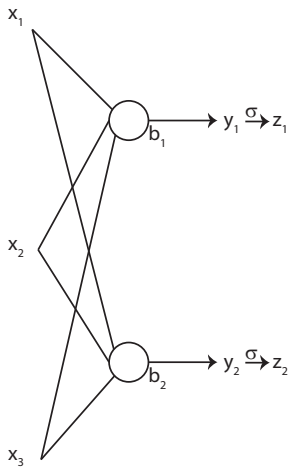
$$y = \vec{w} \cdot \vec{x} - b.$$

(Figure 1a.)

Question A: Let $\vec{w} = (0.2, 0.3, 0.5)$ and $b = 0.1$. What is the output y for the input $x_1 = 1, x_2 = 0, x_3 = 1$?



(A) A neuron model.



(B) A layer consisting of two neurons.

A neural network consists of several layers of neurons. Let's consider one layer. Suppose the layer has two neurons, which just perform the operation

$$\vec{y} = (y_1, y_2) = (\vec{w}_1 \cdot \vec{x} - b_1,\ \vec{w}_2 \cdot \vec{x} - b_2).$$

(Figure 1b.)

Question B: Suppose $\vec{w}_1 = (0.2, 0.3, 0.5)$ and $\vec{w}_2 = (0, 0.6, 0.4)$, $b_1 = 0.1$, $b_2 = 0.6$. What is the output $\vec{y}$ for the input $x_1 = 1, x_2 = 0, x_3 = 1$?

Thus such a layer of neuron is simply a ***linear projection*** $\mathbb{R}^3 \to \mathbb{R}^2$.

In the end of the operation, each neuron applies a function to the value y, say the 'sigmoid function'

$$\sigma(y) = \frac{1}{1+e^{-10y}}$$

which is an approximation of a step function (Figure 2).

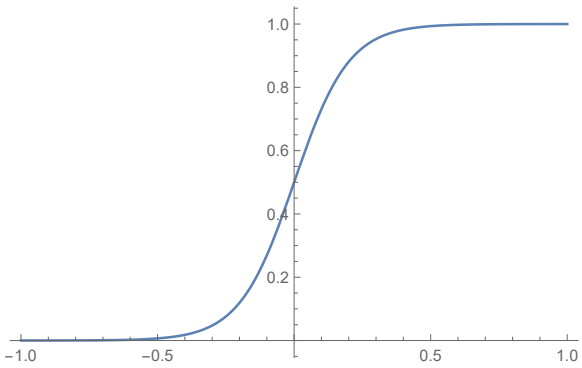


FIGURE 2. The 'sigmoid function'.

Question C: For the above layer of neurons, what is the final output $\vec{z} = (\sigma(y_1), \sigma(y_2))$? Please give the answer up to two decimal places by using a calculator.

The art of self-learning (by the neurons) is to adjust the weights $\vec{w}$ and thresholds b to get closer to given 'correct answers'. Directional derivatives and chain rule that we will learn in this course are the key ingredients in deep learning.

Question D: Suppose the correct answer (for the particular input $\vec{x} = (1,0,1)$) is $\vec{z}_{\text{correct}} = (1,0)$. What is the ***squared distance***

$$\|\vec{z} - \vec{z}_{\text{correct}}\|^2$$

(where $\|\vec{v}\| := \sqrt{\vec{v} \cdot \vec{v}}$ for a vector $\vec{v}$)? (Again use a calculator to give the answer up to two decimal places.) Suppose we change $\vec{w}_1$ to $(0.1, 0.3, 0.6)$. How does it change $\|\vec{z} - \vec{z}_{\text{correct}}\|^2$? Does it result in a better answer or not?