Exercises - Chapter 1

Patrick Browne

June 2023

Sigmoid neurons simulating perceptrons, part I
To show that the behavior of a network of perceptrons does not change when multiplied by a constant c, let us work through the algebra. Suppose we have a neural network of perceptrons with weights $w = [w_1 w_2 ... w_n]$, neurons $x = [x_1 x_2 ... x_n]$ and bias b. The behavior of the network is equal to $w \cdot x + b$, that is,

$$w_1 x_1 + w_2 x_2 + ... + w_n x_n + b$$

Multiplying the network by a constant c, we have

$$cw \cdot x + cb = c \cdot w_1 x_1 + c \cdot w_2 x_2 + ... + c \cdot w_n x_n + c \cdot b$$

Our function to determine whether the neuron fires involves comparing $w \cdot x + b$ to 0, checking whether it is greater than 0 (firing) or not (not firing). We can divide this equation by the constant c and the behavior remains the same.

$$\frac{1}{c} \cdot cw \cdot x + cb > 0 \cdot \frac{1}{c}$$

$$w \cdot x + b > 0$$

We are left with the original behavior of the network. This makes intuitive sense, as scaling a network by a constant factor, multiplying everything in it by some constant, should not change its decision to fire or not. It should only change the magnitude of the result, i.e. how close $w \cdot x + b$ is to 0.

Sigmoid neurons simulating perceptrons, part II
Supposing the same setup as the previous problem, we further assume the inputs to our network $x = [x_1 x_2 ... x_n]$ to be fixed. We also assume that $w \cdot x + b \neq 0$ for input x to any perceptron in our network. Furthermore, suppose we multiply our network by some arbitrary positive constant c and replace all perceptrons with sigmoid neurons (our outputs are now governed by the sigmoid function σ). Let us calculate the behavior as $c \to \infty$:

$$z = (cw \cdot x + cb)$$
$$\lim_{c \to \infty} \frac{1}{1 + e^{-z}}$$
$$= \lim_{c \to \infty} \frac{1}{1 + e^{-c(w \cdot x + b)}}$$

As c tends to ∞, z tends to either ∞ or $-\infty$, depending on the values of w, x and b (which can only result in an output of 0 or 1 in the perceptron model). Thus, σ tends to either

$$\frac{1}{1 + e^{-}\infty} = \frac{1}{1 + 0} = 1$$

or

$$\frac{1}{1 + e^{\infty}} = \frac{1}{\infty} = 0$$

Though this is rather informal, I think you could use the perceptron function definition of $w \cdot x + b$ to split the limit into two cases, then solve for each case as I have just done.

In the case where $w \cdot x + b = 0$, we have a bit of an issue. The limit becomes

$$\lim_{c \to \infty} \frac{1}{1 + e^{-c \cdot 0}}$$
$$= \lim_{c \to \infty} \frac{1}{1 + 1}$$
$$= \lim_{c \to \infty} \frac{1}{2} = \frac{1}{2}$$

Since we simplify $e^{-c \cdot 0}$ to 1 before taking the limit, we get a non- integer value, which does not work with our perceptron model. Taking this to the next layer, if a value relies on that neuron as a threshold (for example, if we have two neurons with weights -2 and bias 3), then we can get a value of 0 and a rejection when, in another case, we would have gotten an acceptance (or vice versa).