Chapter theory 1 solution

Artem Puzanov

2016-06-04

1 Sigmoid neurons simulating perceptrons

1.1 part I

Intuition: Each neuron has a binary activation function. Multiplication of weights and biases by a positive constant is identical to multiplying an inequality by a constant. Such a multiplication does not change the inequality, therefore all perceptron values shall stay the same.

Mathematical notation:

- a_j^l - value of activation function of j-th neuron in l-th layer of our network
- C - the multiplier constant
- w_{ji} - weight of connection from i-th neuron from the $(l-1)$-th layer to the j-th neuron in the l-th layer
- b_{jl} - bias from j-th neuron in the l-th layer
- x_i - output from the i-th neuron in the $(l-1)$-th layer

$$a_j^l = \begin{cases} 1 & \text{if } \sum_i w_{ji}x_i + b_j > 0 \\ 0 & \text{if } \sum_i w_{ji}x_i + b_j \leq 0 \end{cases}$$

For each neuron, we have three possible cases:

1. $\sum_i w_{ji}x_i + b_j > 0$
2. $\sum_i w_{ij}x_i + b_j = 0$
3. $\sum_i w_{ji}x_i + b_j < 0$

Given $C > 0$, it's easy to see that multiplication by C changes none of the conditions

Now, there is a catch: x_i is actually a_i^{l-1}. However, as the choice of the l was non specific, we can apply the same logic for a_i^{l-1}. By moving backwards through $(l-1)$, $(l-2)$, $(l-3)$ and so on, we can see that no a is changed. Therefore there are no changes in the network activation values, which is what we wanted to prove.

1.2 part II

Intuition: The activation function of the sigmoid neuron limits either at 0 or at 1 when z is multiplied by a constant. Limit of the particular neuron depends on whether $wx + b > 0$ or $wx + b < 0$

Mathematical notaion

$$z_j = \sum_i w_i x_i + b_j$$

$$Cz_j = C \sum_i w_i x_i + Cb_j$$

Let's take arbtirary neuron a^l_j. For perceptron, if $z > 0$ (and therefore $a = 1$) than $Cz > 0$ and $a(Cz) = 1$. If it's a sigmoid neuron, than if

$$C \to \infty$$

than

$$1/(1 + e^{-Cz}) \to 1$$

If $z < 0$ (and therefore $a = 0$) than $Cz < 0$. For sigmoid neuron it's

$$C \to \infty$$

than

$$1/(1 + e^{-Cz}) \to 0$$

Therefore if $C \to \infty$ than for all $z < 0$ or $z > 0$ sigmoid neurons emulate perceptrons.

If $wx + b = 0$ than perceptron activation value is 0, and activation value for this sigmoid neuron is $1/(1 + e^0) = 1/2$, which is different from perceptron activation value. Therefore this sigmoid neuron shall not emulate it's perceptron counterpart