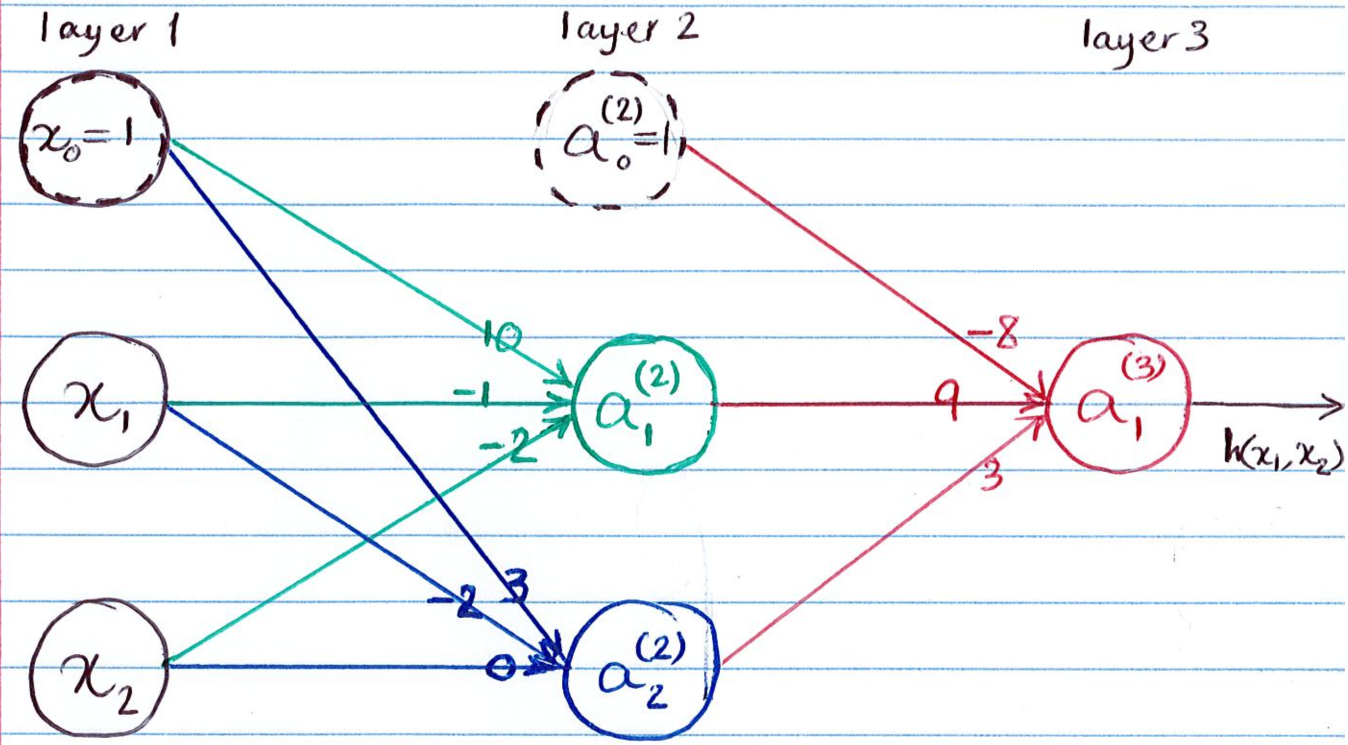
forward propagation for neural networks — example 1

Given the following neural network:



where $\Theta^{(1)} = \begin{bmatrix} \Theta_{10}^{(1)} = 10 & \Theta_{11}^{(1)} = -1 & \Theta_{12}^{(1)} = -2 \\ \Theta_{20}^{(1)} = 3 & \Theta_{21}^{(1)} = -2 & \Theta_{22}^{(1)} = 0 \end{bmatrix}_{2\times 3}$

and $\Theta^{(2)} = \begin{bmatrix} \theta_{10}^{(2)} = -8 & \theta_{11}^{(2)} = 9 & \theta_{12}^{(2)} = 3 \end{bmatrix}_{1\times 3}$

what is $h(x_1, x_2)$ when $\begin{bmatrix} x_0 = 1 \\ x_1 = 2 \\ x_2 = 3 \end{bmatrix} = X$?

propagating from 1_{st} layer to 2_{nd} layer: 2

Let $Z_1^{(2)}$ be the value that goes into the 1_{st} unit of the 2_{nd} layer

and let $Z_2^{(2)}$ be the value that goes into the 2_{nd} unit of the 2_{nd} layer

$$\begin{bmatrix} Z_1^{(2)} \\ Z_2^{(2)} \end{bmatrix} = \Theta^{(1)} * X = \underset{2\times3}{\begin{bmatrix} 10 & -1 & -2 \\ 3 & -2 & 0 \end{bmatrix}} * \underset{3\times1}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}} = \begin{bmatrix} 10(1)+(-1)2+(-2)3 \\ (3)1+(-2)2+(0)3 \end{bmatrix} = \underset{2\times1}{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}$$

$$\begin{bmatrix} a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} = \begin{bmatrix} g(Z_1^{(2)}) \\ g(Z_2^{(2)}) \end{bmatrix} = \begin{bmatrix} \frac{1}{1+e^{-(2)}} \\ \frac{1}{1+e^{-(-1)}} \end{bmatrix} = \underset{2\times1}{\begin{bmatrix} 0.88 \\ 0.27 \end{bmatrix}}$$

Now, we can add $a_0^{(2)}=1$ and get $\begin{bmatrix} a_0^{(2)}=1 \\ a_1^{(2)}=0.88 \\ a_2^{(2)}=0.27 \end{bmatrix}$

So, as $X=\begin{bmatrix} x_0=1 \\ x_1=2 \\ x_2=3 \end{bmatrix}$ propagates from 1_{st} layer to 2_{nd} layer

it will become $a^{(2)}=\begin{bmatrix} a_0^{(2)}=1 \\ a_1^{(2)}=0.88 \\ a_2^{(2)}=0.27 \end{bmatrix}$

propagating from 2nd layer to 3rd layer: 3

Let $Z_1^{(3)}$ be the value that goes into the 1_{st} unit of the 3_{rd} layer

$$\underset{1\times1}{\left[Z_1^{(3)}\right]} = \Theta^{(2)} * a^{(2)} = \underset{1\times3}{\begin{bmatrix} -8 & 9 & 3 \end{bmatrix}} * \underset{3\times1}{\begin{bmatrix} 1 \\ 0.88 \\ 0.27 \end{bmatrix}} = \left[(-8)1 + (9)(0.88) + (3)(0.27)\right] = 0.73$$

$$\left[a_1^{(3)}\right] = \left[g(z_1^{(3)})\right] = \left[\frac{1}{1+e^{-(0.73)}}\right] = 0.67$$

This means $h(x_1 = 2,\ x_2 = 3) = 0.67$

This means $P(y=1) = 0.67$ for input point $\begin{bmatrix} 2 = x_1 \\ 3 = x_2 \end{bmatrix}$

This means input point $\begin{bmatrix} x_1 = 2 \\ x_2 = 3 \end{bmatrix}$ should be classified as belonging to class 1.

probabilities.m

```
function h = probabilities(Theta1, Theta2, X)
%PREDICT Predict the label of an input given a trained neural network
%    p = PREDICT(Theta1, Theta2, X) outputs the predicted label of X given
%    the trained weights of a neural network (Theta1, Theta2)

m = size(X, 2);

a1 = X;
a2 = sigmoid(Theta1 * [ones(1, m); a1]);
a3 = sigmoid(Theta2 * [ones(1, m); a2]);

h = a3;
% =========================================================================

end



function g = sigmoid(z)
    g = 1 ./ (1 + exp(-z));
end
```

Command Window

New to MATLAB? See resources for Getting Started.

```
>> X = [2; 3]

X =

     2
     3

>> Theta1 = [10  -1  -2;     3  -2  0]

Theta1 =

    10    -1    -2
     3    -2     0

>> Theta2 = [-8  9  3]

Theta2 =

    -8     9     3

>> probabilities(Theta1, Theta2, X)

ans =

    0.6757
```