Model Representation II

To re-iterate, the following is an example of a neural network:

$$a_1^{(2)} = g(\Theta_{10}^{(1)}\, x_0 + \Theta_{11}^{(1)}\, x_1 + \Theta_{12}^{(1)}\, x_2 + \Theta_{13}^{(1)}\, x_3)$$

$$a_2^{(2)} = g(\Theta_{20}^{(1)}\, x_0 + \Theta_{21}^{(1)}\, x_1 + \Theta_{22}^{(1)}\, x_2 + \Theta_{23}^{(1)}\, x_3)$$

$$a_3^{(2)} = g(\Theta_{30}^{(1)}\, x_0 + \Theta_{31}^{(1)}\, x_1 + \Theta_{32}^{(1)}\, x_2 + \Theta_{33}^{(1)}\, x_3)$$

$$h_\Theta(x) = a_1^{(3)} = g(\Theta_{10}^{(2)}\, a_0^{(2)} + \Theta_{11}^{(2)}\, a_1^{(2)} + \Theta_{12}^{(2)}\, a_2^{(2)} + \Theta_{13}^{(2)}\, a_3^{(2)})$$

In this section we'll do a vectorized implementation of the above functions. We're going to define a new variable $z_k^{(j)}$ that encompasses the parameters inside our g function. In our previous example if we replaced by the variable z for all the parameters we would get:

$$a_1^{(2)} = g(z_1^{(2)})$$

$$a_2^{(2)} = g(z_2^{(2)})$$

$$a_3^{(2)} = g(z_3^{(2)})$$

In other words, for layer j=2 and node k, the variable z will be:

$$z_k^{(2)} = \Theta_{k,0}^{(1)}\, x_0 + \Theta_{k,1}^{(1)}\, x_1 + \cdots + \Theta_{k,n}^{(1)}\, x_n$$

The vector representation of x and z^j is:

$$x = \begin{bmatrix} x_0 \\ x_1 \\ \cdots \\ x \end{bmatrix} \quad z^{(j)} = \begin{bmatrix} z_1^{(j)} \\ z_2^{(j)} \\ \cdots \\ \end{bmatrix}$$