Backpropagation Intuitior

Note: [4:39, the last term for the calculation for z_1^3 (thre
a_2^2 instead of a_1^2. 6:08 - the equation for cost(i) is incorre
for the log() function, and the second term should be $(1$
$y - a^{(4)}$ is incorrect and should be $\delta^{(4)} = a^{(4)} - y$.]

Recall that the cost function for a neural network is:

$$J(\Theta) = -\frac{1}{m}\sum_{t=1}^{m}\sum_{k=1}^{K} y_k^{(t)} \log(h_\Theta(x^{(t)}))_k + (1 - y_k^{(t)}) \log$$

If we consider simple non-multiclass classification (k = 1
computed with:

$$cost(t) = y^{(t)}\ \log(h_\Theta(x^{(t)})) + (1 - y^{(t)})\ \log(1 - h_\Theta(x$$

Intuitively, $\delta_j^{(l)}$ is the "error" for $a_j^{(l)}$ (unit j in layer l). Mo
the derivative of the cost function:

$$\delta_j^{(l)} = \frac{\partial}{\partial z_j^{(l)}} cost(t)$$

Recall that our derivative is the slope of a line tangent tc
slope the more incorrect we are. Let us consider the foll
we could calculate some $\delta_j^{(l)}$: