

Graph 3 — Lecture 16 Exercise

- Here are the equations of a simple layered graph

$$L = \frac{1}{2}(y - a^2)^2$$

$$a^2 = \sigma(z^2)$$

$$z^2 = w_1^2 a_1^1 + w_2^2 a_2^1$$

$$a_1^1 = \sigma(z_1^1)$$

$$a_2^1 = \sigma(z_2^1)$$

$$z_1^1 = w_{1,1}^1 x_1 + w_{1,2}^1 x_2$$

$$z_2^1 = w_{2,1}^1 x_1 + w_{2,2}^1 x_2$$

- Here, the superscripts indicate the layers, so that for example a_2^1 is the second activation variable at layer 1 and $w_{2,1}^1$ is a layer 1 weight from x_1 to z_2^1 .
 - The lone exception is the equation for L where the outermost 2 superscript is in fact “square”. Sorry for the weirdness.
- Please submit a pdf file showing the three derivatives:

$$\frac{\partial L}{\partial z^2} = \dots$$

$$\frac{\partial L}{\partial z_1^1} = \dots$$

$$\frac{\partial L}{\partial w_{1,1}^1} = \dots$$

The third should be recursively defined in terms of the second and the second should be recursively defined in terms of the first.