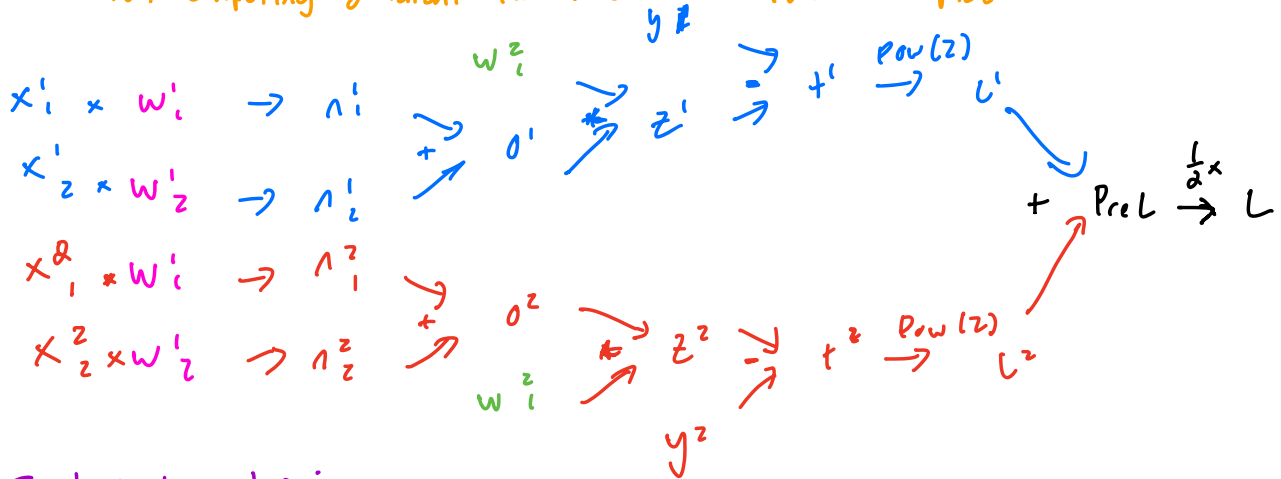


Scenario: Computing gradient for a batch of two examples



Topological order:

$$\left[ x_1^1, w_1^1, n_1^1, x_2^1, w_2^1, n_2^1, o_1^1, w_1^1, z_1^1, y_1^1, t_1^1, l_1^1, x_2^2, n_1^2, x_2^2, n_2^2, o_2^2, z_2^2, y_2^1, t_2^2, l_2^2, \text{PreL}, L \right]$$

Backprop:

$$\frac{dL}{dL} = 1$$

$$\frac{dL}{d\text{PreL}} = \frac{1}{2}$$

$$\frac{dL}{dL^1} = \frac{1}{2}$$

$$\frac{dL}{dt^1} = 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dz^1} = 1 \times (2t^1 \times \frac{1}{2})$$

$$\frac{dL}{dy^1} = -1 \times (2t^1 \times \frac{1}{2})$$

$$\frac{dL}{do^1} = w_1^1 \times 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dn_2^1} = w_1^1 \times 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dx_2^1} = w_1^1 \times w_2^1 \times 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dn_1^1} = w_1^1 \times 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dx_1^1} = w_1^1 \times w_2^1 \times 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dL_1} = \frac{1}{2}$$

$$\frac{dL}{dt^1} = 2t^1 \times \frac{1}{2}$$

$$\frac{dL}{dO'} = W^2_1 \times 2 + \frac{1}{2}$$

$$\frac{dL}{dn^{\frac{1}{2}}_2} = W^2_1 \times 2 + \frac{1}{2}$$

$$\frac{dL}{dy'} = -1 \times (z + \frac{1}{2})$$

$$\frac{dL}{dz'} = 1 \times (z + \frac{1}{2})$$

$$\frac{dL}{W^{\frac{1}{2}}_2} = x^2_2 \times \frac{dL}{n^{\frac{1}{2}}_2}$$

$$+ \frac{1}{2} \times \frac{dL}{n^{\frac{1}{2}}_2}$$

$$\frac{dL}{W^{\frac{1}{2}}_1} = x^2_1 \times \frac{dL}{n^{\frac{1}{2}}_1} +$$

$$+ \frac{1}{2} \times \frac{dL}{n^{\frac{1}{2}}_1}$$

$W^2_1$  is used in both forward passes,

$\therefore$  gradient computed for

weights in a batch is the sum of the gradients computed for each independent forward pass and corresponding backward pass

$$\frac{dL}{dW^2_1} = \frac{dz^1}{dW^2_1} \times \frac{dL}{dz^1} \rightarrow 0' \times \frac{dL}{dz^1} +$$

$$\frac{dz^2}{dW^2_1} \times \frac{dL}{dz^2} \rightarrow 0^2 + \frac{dL}{dz^2}$$

↳ Note: In actuality for this loss it is the average across all forward pass since the loss is averaged

by  $\frac{1}{2}$  and in turn the gradient computed for each pass is multiplied by  $\frac{1}{2}$

$$\frac{dL}{dz^1} = \frac{1}{2} \times 2 +$$