

Backpropagation with softmax output + logistic loss fn

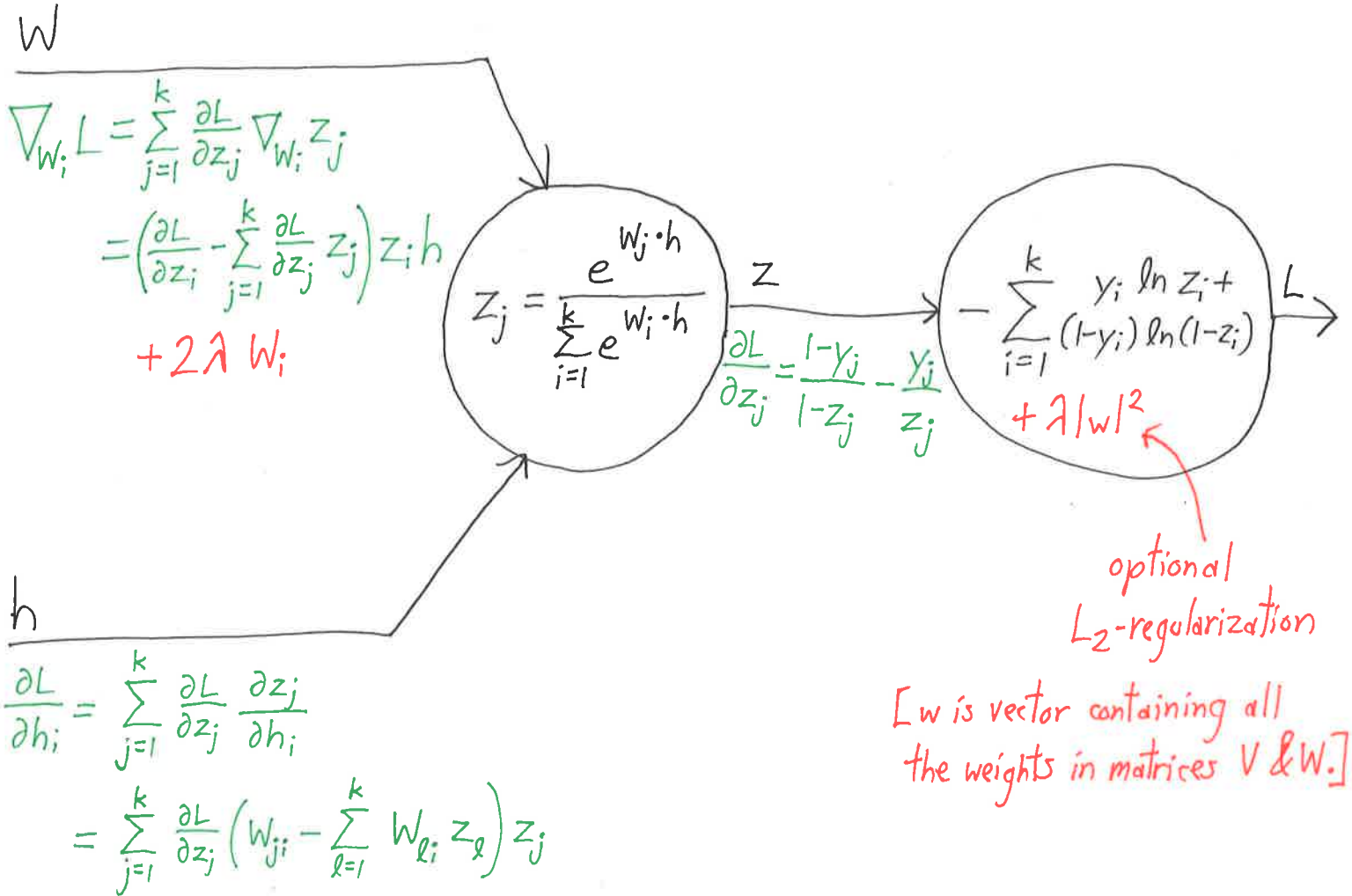
Softmax output is $z_j(t) = \frac{e^{t_j}}{\sum_{i=1}^k e^{t_i}}$

$\epsilon(0, 1)$ $\sum_j z_j = 1$

$t_i = W_i \cdot h$

$$\frac{dz_j}{dt_j} = z_j(1-z_j)$$

$$\frac{dz_j}{dt_i} = -z_i z_j \quad i \neq j$$



Derivatives of inputs to hidden units h are computed same way as previously.