1) a) $\frac{\partial L(t)}{\partial z_j^L(t)} \overset{!}{=} -(z_j^L(t) - z_j^*(t))$

with $L(t) = -\frac{1}{2}\sum_j (z_j^L(t) - z_j^*(t))^2$

we can reformulate

$\frac{\partial L(t)}{\partial z_j^L(t)}$ as $\frac{\partial(-\frac{1}{2}\sum_j (z_j^L(t) - z_j^*(t))^2)}{\partial z_j^L(t)}$

this is equal to

$$\frac{\partial(-\frac{1}{2}\sum_j (z_j^L(t)^2 - 2 \cdot z_j^L(t) \cdot z_j^*(t) + z_j^*(t)))}{\partial z_j^L(t)}$$

$$\Rightarrow -\frac{1}{2}(2 z_j^L(t) - 2 z_j^*(t))$$

$$\Rightarrow -(z_j^L(t) - z_j^*(t))$$

b) $\frac{\partial z_j^{l+1}(t)}{\partial z_m^l(t)} \overset{!}{=} \varphi'(v_j^{l+1}(t)) \cdot w_{jm}^{l+1}$ for $l < L$

with $z_j^{l+1}(t) = \sum_i (w_{jm}^{l+1} \cdot \varphi(v_i^{l+1}(t)))$

we can reformulate

$\frac{\partial z_j^{l+1}(t)}{\partial z_m^l(t)}$ as $\frac{\partial(\sum_i (w_{jm}^{l+1} \cdot \varphi(v_i^{l+1}(t))))}{\partial z_m^l(t)}$

$$\frac{\partial z_j^{l+1}(t)}{\partial z_m^l(t)} \quad \text{as} \quad \frac{\partial(\sum_j (w_{jm} \cdot \varphi(v_j^{l+1}(t))))}{\partial z_m^l(t)}$$

with $l < L$ we obtain:

$$\Rightarrow w_{jm}^{l+1} \cdot \varphi'(v_j^{l+1}(t))$$

c) $$\frac{\partial z_j^l}{\partial w_{ki}} \stackrel{!}{=} \varphi'(v_j^l(t)) \cdot z_i^{l-1}(t) \cdot K_{jk}$$

$$\hookrightarrow K_{jk} = \begin{cases} 1 & \text{for } j = k \\ 0 & \text{otherwise} \end{cases}$$

we can write

$$\frac{\partial z_j^l}{\partial w_{ki}} \quad \text{as} \quad \frac{\partial z_j^l}{\partial z_j^{l-1}} \cdot \frac{\partial z_j^{l-1}}{\partial w_{ki}}$$

from b) we know that

$$\frac{\partial z_j^l}{\partial z_j^{l-1}} = w_{ji}^l \cdot \varphi'(v_j^l(t))$$

and we therefore obtain

$$\frac{\partial z_j^l}{\partial w_{ki}} = \frac{\partial(w_{ji}^l \cdot \varphi'(v_j^l(t)) \cdot z_i^{l-1})}{\partial w_{ki}}$$

This is 0 for $w_{ki} \neq w_{ji} \rightarrow$ for $j \neq k$

For $j = k$ we obtain

$$\varphi'(v_j^l(t)) \cdot z_i^{l-1}$$

This can be described by the Kronecker Delta K to get:

$$\frac{\partial z_j^l}{\partial W_{ki}} = \varphi'(v_j^l(t)) \cdot z_i^{l-1} \cdot K_{jk}$$

where $K = \begin{cases} 1 & \text{for } j=k \\ 0 & \text{otherwise} \end{cases}$

d) $$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^L} \overset{!}{=} \delta_k^L(t) \cdot z_i^{L-1}(t) \quad (\text{for } l = L)$$

we can reformulate $\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^L}$ as

$$\underbrace{\frac{\partial \mathcal{L}(t)}{\partial z_k^L(t)}}_{a)} \cdot \underbrace{\frac{\partial z_k^L(t)}{\partial W_{ki}^L}}_{c)}$$ and use the

calculated solutions from a) & c). Therefore,

$$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^L} = -\left(z_k^L(t) - z_k^*(t)\right) \cdot \varphi'(v_k^L(t)) \cdot z_i^{L-1}(t) \cdot K_{kk}$$

since $K_{kk}$ is 1 we can ignore it.

Formulating $-\left(z_k^L(t) - z_k^*(t)\right) \cdot \varphi'(v_k^L(t))$ as $\delta_k^L(t)$

we obtain the solution

$$\frac{\partial \mathcal{L}(t)}{\dots} = \delta_k^L(t) \cdot z_i^{L-1}(t)$$

$$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}} = \delta_k'(t) \cdot z_i(t)$$

e) $$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^{l}} \overset{!}{=} \delta_k^{l}(t) \cdot z_i^{l-1}(t) \quad \text{for } l = L-1$$

$(\rightarrow L = l+1)$

we can extend $\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^{l}}$ to

$$\underbrace{\frac{\partial \mathcal{L}(t)}{\partial z_j^{l+1}(t)}}_{a)} \cdot \underbrace{\frac{\partial z_j^{l+1}(t)}{\partial z_j^{l}(t)}}_{b)} \cdot \underbrace{\frac{\partial z_j^{l}(t)}{\partial W_{ki}^{l}}}_{c)}$$

Again, using previous solutions (this is possible because $l = L-1$ and therefore $l+1 = L$):

$$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^{l}} = -\left(z_j^{L}(t) - z_j^{*}(t)\right) \cdot \varphi'(v_j^{L}(t)) \cdot W_{kj}^{L} \ldots$$

$$\cdot\, \varphi'(v_j^{l}(t)) \cdot z_i^{l-1}(t) \cdot \underbrace{K_{kk}}_{1}$$

with $\delta_j^{L}(t) = -\left(z_j^{L}(t) - z_j^{*}(t)\right) \cdot \varphi'(v_j^{L}(t))$ [see d)]

we obtain

$$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^{l}} = \delta_j^{L}(t) \cdot W_{kj}^{L} \cdot \varphi'(v_j^{l}(t)) \cdot z_i^{l-1}(t)$$

$$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^{l}} = \delta_j^L(t) \cdot W_{kj}^L \cdot \varphi'(v_j^l(t)) \cdot z_i \; (t)$$

where we can set

$$\delta_k^l(t) = \delta_j^L(t) \cdot W_{kj}^L \cdot \varphi'(v_j^l(t)) = \delta_j^{l+1}(t) \cdot W_{kj}^{l+1} \cdot \varphi'(v_j^l(t))$$

to obtain

$$\frac{\partial \mathcal{L}(t)}{\partial W_{ki}^{l}} = \delta_K^l(t) \cdot z_i^{l-1}(t)$$ which relies on the error terms beforehand, whereas $\delta_j^L(t)$ is the first one.

f) $$\frac{\partial \mathcal{L}(t)}{\partial b_j^l} \stackrel{!}{=} \delta_j^l(t)$$

To maintain previous solutions, we can introduce the bias $b_j^l$ as weight to the additional node $z_0^{l-1}$ with constant output 1. This means we reformulate the activation function of form

$$b_j^l + \sum_{i=1} w_{ji}^l \cdot z_i^{l-1} \quad \text{to} \quad \sum_{i=0} w_{ji}^l \cdot z_i^{l-1}.$$

Now our previous solutions hold.

Now our previous

$\frac{\partial \mathcal{L}(t)}{\partial b_j^l}$ is now, similar to e), because $b_j \in w_{ij}$:

$$\frac{\partial \mathcal{L}(t)}{\partial b_j^l} = \delta_j^l(t) \cdot z_i^{l-1}(t)$$

where $z_i^{l-1}(t)$ represents our additional node with constant output 1. Therefore,

$$\frac{\partial \mathcal{L}(t)}{\partial b_j^l} = \delta_j^l(t)$$