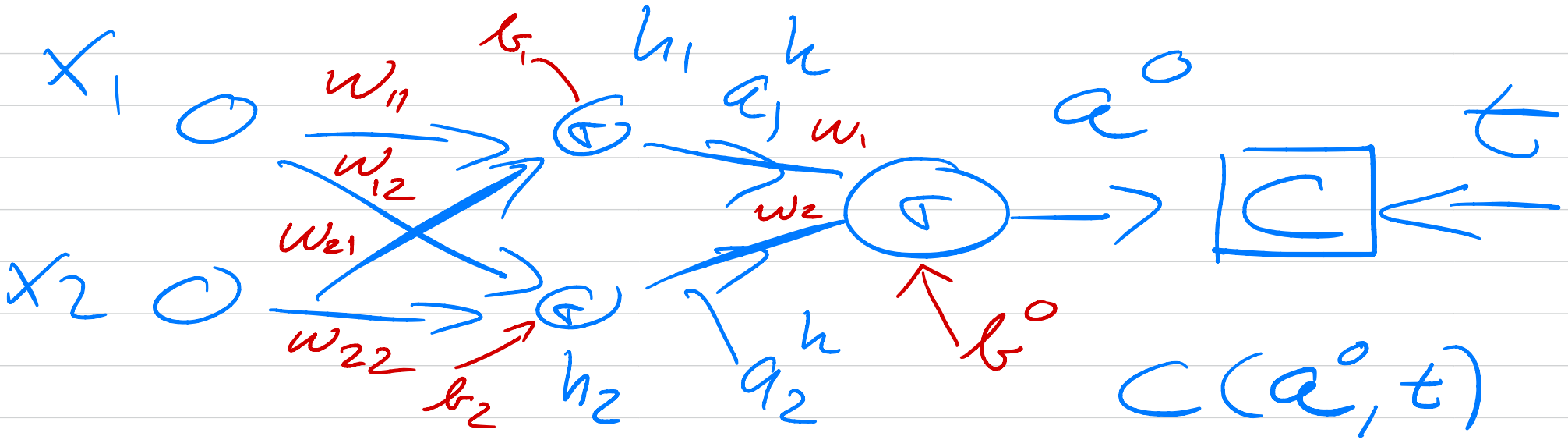
Fys-stk3155/4155, the back propagation part for project 2 and exercises weeks 43 and 44



$$a_1^h = \sigma(z_1^h) = \sigma(w_{11}x_1 + w_{21}x_2 + b_1)$$
$$a_2^h = \sigma(z_2^h) = \sigma(w_{12}x_1 + w_{22}x_2 + b_2)$$
$$a^o = \sigma(z^o) = \sigma(w_1 a_1^h + w_2 a_2^h + b^o)$$

Back propagation,

here $C(w, b) = C(\theta)$

$$= -\sum_i \left[t_i \log a_i^o + (1 - t_i) \log(1 - a_i^o) \right]$$

$\sigma(z) = 1/(1+e^{-z})$

$$\frac{\partial C}{\partial w^L_{ij}} = \frac{\partial C}{\partial a^L_i}\frac{\partial a^L_i}{\partial z^L_i}\frac{\partial z^L_i}{\partial w^L_{ij}}$$ chain rule

L = output (o)

$\frac{\partial C}{\partial a^L_i}\frac{\partial a^L_i}{\partial z^L_i} = \delta^L_i$

$\frac{\partial z^L_i}{\partial w^L_{ij}} = a^{L-1}_j$

autograd or hand coding

(autograd or hand code) derivative of activation function, here Sigmoid

$\sigma'(z) = \sigma(z)(1-\sigma(z))$

crossentropy

$$\frac{a^L_i - t_i}{a^L_i(1-a^L_i)} \Rightarrow \frac{a^L_i - t_i}{a^L_i(1-a^L_i)} \times \sigma'(z)$$

$$= a^L_i - t_i \quad !$$

Back propagation

$$\delta_j^\ell = \sum_k \delta_k^{\ell+1} w_{kj}^{\ell+1} \sigma'(z_j^\ell)$$

Gradient descent

$$w_{jk}^\ell \leftarrow w_{jk}^\ell - \eta \delta_j^\ell a_k^{\ell-1}$$

$$b_j^\ell \leftarrow b_j^\ell - \eta \delta_j^\ell$$

↳ gradients

convenient to start with plain gradient descent.