sigmoid 求导推导

$$y = \left(\frac{1}{e^{-x}+1}\right)' \qquad \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

$$= \frac{-1 * (e^{-x}+1)'}{(e^{-x}+1)^2} \qquad (e^{-x}+1)' = (e^{-x})' + 1' = (-x)' \cdot e^{-x} = -e^{-x}$$

$$= \frac{e^{-x}}{(e^{-x}+1)^2}$$

$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}}$$

$$= \frac{1}{1+e^{-x}} \cdot \frac{1+e^{-x}-1}{1+e^{-x}}$$

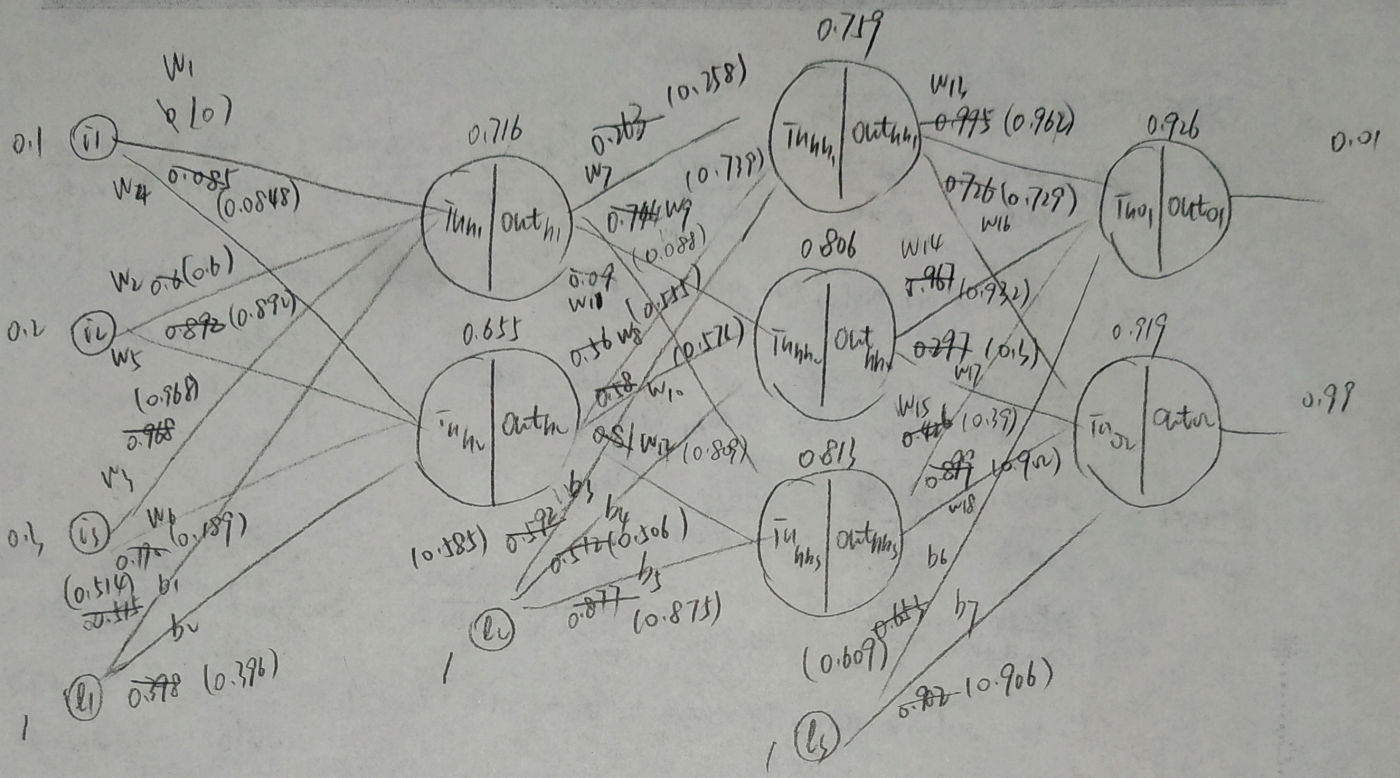
$$= \frac{1}{1+e^{-x}} \left(1 - \frac{1}{1+e^{-x}}\right)$$

$$= y(1-y)$$

装……订……线

$$loss = \frac{1}{2}(0.01 - 0.926)^2 + \frac{1}{2}(0.99 - 0.919)^2$$

$$= 0.421$$



$$in_{h_1} = i_1 * w_1 + i_2 * w_2 + i_3 * b_3 + 1 * b_1 = 0.1 * 0 + 0.2 * 0.6 + 0.1 * 0.515 + 1 * 0.396 = 0.4756$$

$$out_{h_1} = \frac{1}{e^{-in_{h_1}} + 1} = \frac{1}{e^{-0.4756} + 1} = 0.716$$

$$in_{h_2} = i_1 * w_4 + i_2 * w_5 + i_3 * w_6 + 1 * b_2 = 0.1 * 0.085 + 0.2 * 0.868 + 0.1 * 0.889 + 1 * 0.858 = 0.6419$$

$$out_{h_2} = \frac{1}{e^{-in_{h_2}} + 1} = \frac{1}{e^{-0.6419} + 1} = 0.655$$

$$in_{hh_1} = out_{h_1} * w_7 + out_{h_2} * w_8 + 1 * b_4 = 0.716 * 0.744 + 0.655 * 0.872 + 1 * 0.962 = 1.467$$

$$out_{hh_1} = \frac{1}{e^{-in_{hh_1}} + 1} = \frac{1}{e^{-1.467} + 1} = 0.759$$

$$in_{hh_2} = out_{h_1} * w_9 + out_{h_2} * w_{10} + 1 * b_4 = 0.716 * 0.744 + 0.655 * 0.872 + 1 * 0.962 = 1.425$$

$$out_{hh_2} = \frac{1}{e^{-in_{hh_2}} + 1} = \frac{1}{e^{-1.425} + 1} = 0.806$$

$$in_{hh_3} = out_{h_1} * w_{11} + out_{h_2} * w_{12} + 1 * b_5 = 0.716 * 0.609 + 0.655 * 0.889 + 1 * 0.877 = 1.471$$

$$out_{hh_3} = \frac{1}{e^{-in_{hh_3}} + 1} = \frac{1}{e^{-1.471} + 1} = 0.813$$

$$in_{o_1} = out_{hh_1} * w_{13} + out_{hh_2} * w_{14} + out_{hh_3} * w_{15} + 1 * b_6 = 0.759 * 0.995 + 0.806 * 0.967 + 0.813 * 0.297 + 1 * 0.653 = 2.534$$

$$out_{o_1} = \frac{1}{e^{-in_{o_1}} + 1} = \frac{1}{e^{-2.534} + 1} = 0.926$$

$$in_{o_2} = out_{hh_1} * w_{16} + out_{hh_2} * w_{17} + out_{hh_3} * w_{18} + 1 * b_7 = 0.759 * 0.726 + 0.806 * 0.932 + 0.813 * 0.39 + 1 * 0.902 = 2.423$$

$$out_{o_2} = \frac{1}{e^{-in_{o_2}} + 1} = \frac{1}{e^{-2.423} + 1} = 0.919$$

$$coef_{o1} = \frac{\partial E_{total}}{\partial out_{o1}} \cdot \frac{\partial out_{o1}}{\partial in_{o1}} = -(o1 - out_{o1}) \cdot out_{o1}(1 - out_{o1})$$

$$= -(0.01 - 0.916) \cdot 0.916 (1 - 0.916) = 0.0628$$

$$b6 = b6 - \alpha \cdot coef_{o1} = 0.655 - 0.7 * 0.0628 = 0.609$$

$$w13 = w13 - \alpha \cdot coef_{o1} \cdot out_{hh1} = 0.995 - 0.7 * 0.0628 * 0.759 = 0.962$$

$$w14 = w14 - \alpha \cdot coef_{o1} \cdot out_{hh2} = 0.867 - 0.7 * 0.0628 * 0.806 = 0.932$$

$$w15 = w15 - \alpha \cdot coef_{o1} \cdot out_{hh3} = 0.416 - 0.7 * 0.0628 * 0.813 = 0.39$$

$$coef_{o2} = \frac{\partial E_{total}}{\partial out_{o2}} \cdot \frac{\partial out_{o2}}{\partial in_{o2}} = -(o2 - out_{o2}) \cdot out_{o2}(1 - out_{o2})$$

$$= -(0.99 - 0.919) \cdot 0.919 (1 - 0.919) = -0.005$$

$$b7 = b7 - \alpha \cdot coef_{o2} = 0.902 - 0.7 * -0.005 = 0.906$$

$$w16 = w16 - \alpha \cdot coef_{o2} \cdot out_{hh1} = 0.716 - 0.7 * -0.005 * 0.759 = 0.729$$

$$w17 = w17 - \alpha \cdot coef_{o2} \cdot out_{hh2} = 0.297 - 0.7 * -0.005 * 0.806 = 0.3$$

$$w18 = w18 - \alpha \cdot coef_{o2} \cdot out_{hh3} = 0.888 - 0.7 * -0.005 * 0.813 = 0.902$$

$$coef_{hh1} = \frac{\partial E_{total}}{\partial out_{hh1}} \cdot \frac{\partial out_{hh1}}{\partial in_{hh1}} = (coef_{o1} * w13 + coef_{o2} * w16) * out_{hh1}(1 - out_{hh1})$$

$$= (0.0628 * 0.962 + -0.005 * 0.729) * 0.759 (1 - 0.759) = 0.01$$

$$b3 = b3 - \alpha \cdot coef_{hh1} = 0.592 - 0.7 * 0.01 = 0.585$$

$$w7 = w7 - \alpha \cdot coef_{hh1} \cdot out_{h1} = 0.263 - 0.7 * 0.01 * 0.716 = 0.258$$

$$w8 = w8 - \alpha \cdot coef_{hh1} \cdot out_{h2} = 0.56 - 0.7 * 0.01 * 0.655 = 0.555$$

$$coef_{hh2} = \frac{\partial E_{total}}{\partial out_{hh2}} \cdot \frac{\partial out_{hh2}}{\partial in_{hh2}} = (coef_{o1} * w14 + coef_{o2} * w17) * out_{hh2}(1 - out_{hh2})$$

$$= (0.0628 * 0.932 + -0.005 * 0.3) * 0.806 (1 - 0.806) = 0.009$$

$$b4 = b4 - \alpha \cdot coef_{hh2} = 0.512 - 0.7 * 0.009 = 0.506$$

$$w9 = w9 - \alpha \cdot coef_{hh2} \cdot out_{h1} = 0.744 - 0.7 * 0.009 * 0.716 = 0.739$$

$$w10 = w10 - \alpha \cdot coef_{hh2} \cdot out_{h2} = 0.58 - 0.7 * 0.009 * 0.655 = 0.576$$

$$coef_{hh3} = \frac{\partial E_{total}}{\partial out_{hh3}} \cdot \frac{\partial out_{hh3}}{\partial in_{hh3}} = (coef_{o1} * w15 + coef_{o2} * w18) * out_{hh3}(1 - out_{hh3})$$

$$= (0.0628 * 0.39 - 0.005 * 0.902) * 0.813 * (1 - 0.813) = 0.003$$

$$b5 = b5 - \alpha \cdot coef_{hh3} = 0.877 - 0.7 * 0.003 = 0.875$$

$$w11 = w11 - \alpha \cdot coef_{hh3} * out_{h1} = 0.09 - 0.7 * 0.003 * 0.716 = 0.088$$

$$w12 = w12 - \alpha \cdot coef_{hh3} * out_{h2} = 0.81 - 0.7 * 0.003 * 0.655 = 0.809$$

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial in_{h_1}} \cdot \frac{\partial in_{h_1}}{\partial w_1}$$

$$= \left(\frac{\partial E_{total}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial in_{o_1}} \cdot \frac{\partial in_{o_1}}{\partial out_{h_1}} + \frac{\partial E_{total}}{\partial out_{o_2}} \cdot \frac{\partial out_{o_2}}{\partial in_{o_2}} \cdot \frac{\partial in_{o_2}}{\partial out_{h_1}} \right) \frac{\partial out_{h_1}}{\partial in_{h_1}} \cdot \frac{\partial in_{h_1}}{\partial w_1}$$

$$= \Big(-(o_1 - out_{o_1})\, out_{o_1}(1-out_{o_1}) \cdot w_5 + -(o_2 - out_{o_2}) \cdot out_{o_2} \cdot (1-out_{o_2}) \cdot w_7 \Big)$$
$$out_{h_1}(1-out_{h_1}) \cdot i_1$$

上层 上层 此层 此层 上层 上层 本层

$$= \boxed{\left(\delta_{o_1} \cdot w_5 + \delta_{o_2} \cdot w_7 \right) \cdot out_{h_1}(1-out_{h_1}) \cdot i_1}$$
$$\delta_{h_1}$$

$$\frac{\partial E_{total}}{\partial w_2} = \frac{\partial E_{total}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial in_{h_1}} \cdot \frac{\partial in_{h_1}}{\partial w_2}$$

$$= \left(\frac{\partial E_{total}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial in_{o_1}} \cdot \frac{\partial in_{o_1}}{\partial out_{h_1}} + \frac{\partial E_{total}}{\partial out_{o_2}} \cdot \frac{\partial out_{o_2}}{\partial in_{o_2}} \cdot \frac{\partial in_{o_2}}{\partial out_{h_1}} \right) \cdot \frac{\partial out_{h_1}}{\partial in_{h_1}} \cdot \frac{\partial in_{h_1}}{\partial w_2}$$

$$= \Big(-(o_1 - out_{o_1}) \cdot out_{o_1}(1-out_{o_1}) \cdot w_5 + -(o_2 - out_{o_2}) \cdot out_{o_2} \cdot (1-out_{o_2}) \cdot w_7 \Big)$$
$$out_{h_1}(1-out_{h_1}) \cdot i_2$$

$$= \boxed{\left(\delta_{o_1} \cdot w_5 + \delta_{o_2} \cdot w_7 \right) \cdot out_{h_1}(1-out_{h_1}) \cdot i_2}$$
$$\delta_{h_1}$$

$$\frac{\partial E_{total}}{\partial w_3} = \frac{\partial E_{total}}{\partial out_{h_2}} \cdot \frac{\partial out_{h_2}}{\partial in_{h_2}} \cdot \frac{\partial in_{h_2}}{\partial w_3}$$

$$= \left(\frac{\partial E_{total}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial in_{o_1}} \cdot \frac{\partial in_{o_1}}{\partial out_{h_2}} + \frac{\partial E_{total}}{\partial out_{o_2}} \cdot \frac{\partial out_{o_2}}{\partial in_{o_2}} \cdot \frac{\partial in_{o_2}}{\partial out_{h_2}} \right) \frac{\partial out_{h_2}}{\partial in_{h_2}} \cdot \frac{\partial in_{h_2}}{\partial w_3}$$

$$= \Big(-(o_1 - out_{o_1}) \cdot out_{o_1}(1-out_{o_1}) \cdot w_6 + -(o_2 - out_{o_2}) out_{o_2}(1-out_{o_2}) \cdot w_8 \Big)$$
$$out_{h_2}(1-out_{h_2}) \cdot i_1$$

$$= \boxed{\left(\delta_{o_1} \cdot w_6 + \delta_{o_2} \cdot w_8 \right) \cdot out_{h_2}(1-out_{h_2}) \cdot i_1}$$
$$\delta_{h_2}$$

$$\frac{\partial E_{total}}{\partial w_4} = \frac{\partial E_{total}}{\partial out_{h_2}} \cdot \frac{\partial out_{h_2}}{\partial in_{h_2}} \cdot \frac{\partial in_{h_2}}{\partial w_4}$$

$$= \left(\frac{\partial E_{total}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial in_{o_1}} \cdot \frac{\partial in_{o_1}}{\partial out_{h_2}} + \frac{\partial E_{total}}{\partial out_{o_2}} \cdot \frac{\partial out_{o_2}}{\partial in_{o_2}} \cdot \frac{\partial in_{o_2}}{\partial out_{h_2}} \right) \frac{\partial out_{h_2}}{\partial in_{h_2}} \cdot \frac{\partial in_{h_2}}{\partial w_4}$$

$$= \Big(-(o_1 - out_{o_1}) \cdot out_{o_1}(1-out_{o_1}) \cdot w_6 + -(o_2 - out_{o_2}) \cdot out_{o_2}(1-out_{o_2}) \cdot w_8 \Big)$$
$$out_{h_2}(1-out_{h_2}) \cdot i_2$$

$$= \boxed{\left(\delta_{o_1} \cdot w_6 + \delta_{o_2} \cdot w_8 \right) \cdot out_{h_2}(1-out_{h_2}) \cdot i_2}$$
$$\delta_{h_2}$$

装……订……线

$$coef_{h1} = \frac{\partial E_{total}}{\partial out_{h1}} \cdot \frac{\partial out_{h1}}{\partial in_{h1}} = \frac{\partial E_{total}}{\partial out_{hh1}} \cdot \frac{\partial out_{hh1}}{\partial in_{hh1}} \cdot \frac{\partial in_{hh1}}{\partial out_{h1}} + \frac{\partial E_{total}}{\partial out_{hh2}} \cdot \frac{\partial out_{hh2}}{\partial in_{hh2}} \cdot \frac{\partial in_{hh2}}{\partial out_{h1}}$$

$$+ \frac{\partial E_{total}}{\partial out_{hh3}} \cdot \frac{\partial out_{hh3}}{\partial in_{hh3}} \cdot \frac{\partial in_{hh3}}{\partial out_{h1}}$$

$$= (coef_{hh1} \cdot w_7 + coef_{hh2} \cdot w_9 + coef_{hh3} \cdot w_{11}) \, out_{h1}(1-out_{h1})$$

$$= (0.01 * 0.258 + 0.009 * 0.739 + 0.003 * 0.088) * 0.716(1-0.716) = 0.002$$

$$b_1 = b_1 - \alpha \cdot coef_{h1} = 0.515 - 0.7 * 0.002 = 0.514$$

$$w_1 = w_1 - \alpha \cdot coef_{h1} \cdot i_1 = 0 - 0.7 * 0.002 * 0.1 = 0$$

$$w_2 = w_2 - \alpha \cdot coef_{h1} \cdot i_2 = 0.6 - 0.7 * 0.002 * 0.2 = 0.6$$

$$w_3 = w_3 - \alpha \cdot coef_{h1} \cdot i_3 = 0.968 - 0.7 * 0.002 * 0.3 = 0.968$$

$$coef_{h2} = \frac{\partial E_{total}}{\partial out_{h2}} \cdot \frac{\partial out_{h2}}{\partial in_{h2}} = \left(\frac{\partial E_{total}}{\partial out_{hh1}} \cdot \frac{\partial out_{hh1}}{\partial in_{hh1}} \cdot \frac{\partial in_{hh1}}{\partial out_{h2}} + \frac{\partial E_{total}}{\partial out_{hh2}} \cdot \frac{\partial out_{hh2}}{\partial in_{hh2}} \cdot \frac{\partial in_{hh2}}{\partial out_{h2}} \right.$$

$$\left. + \frac{\partial E_{total}}{\partial out_{hh3}} \cdot \frac{\partial out_{hh3}}{\partial in_{hh3}} \cdot \frac{\partial in_{hh3}}{\partial out_{h2}} \right) out_{h2}(1-out_{h2})$$

$$= (coef_{hh1} \cdot w_8 + coef_{hh2} \cdot w_{10} + coef_{hh3} \cdot w_{12}) \, out_{h2}(1-out_{h2})$$

$$= (0.01 * 0.555 + 0.009 * 0.576 + 0.003 * 0.809) \cdot 0.655(1-0.655)$$

$$= 0.003$$

$$b_2 = b_2 - \alpha \cdot coef_{h2} = 0.398 - 0.7 * 0.003 = 0.396$$

$$w_4 = w_4 - \alpha \cdot coef_{h2} * i_1 = 0.085 - 0.7 * 0.003 * 0.1 = 0.0848$$

$$w_5 = w_5 - \alpha \cdot coef_{h2} * i_2 = 0.892 - 0.7 * 0.003 * 0.2 = 0.892$$

$$w_6 = w_6 - \alpha \cdot coef_{h3} * i_3 = 0.19 - 0.7 * 0.003 * 0.3 = 0.189$$

装……订……线