Gradients

Otto and Joost

January 13, 2014

1 h

$$\frac{\partial \log p_\theta(x|z)}{\partial W_3} = \frac{\partial \log p_\theta}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial h}\frac{\partial h}{\partial W_3}$$

$$\frac{\partial h}{\partial W_3} = x(1 - \tanh^2(W_3 x + b_3))$$

$$\frac{\partial z}{\partial h} = \frac{\partial \mu}{\partial h} + \epsilon\frac{\partial \sigma}{\partial h}$$

$$\frac{\partial \mu}{\partial h} = W_4$$

$$\frac{\partial \sigma}{\partial h} = \frac{W_5 e^{W_5 h + b_5}}{2\sqrt{e^{W_5 h + b_5}}} = \frac{1}{2} W_5 \sqrt{e^{W_5 h + b_5}}$$

For y, we call what is after the minus sign in the exponent g.

$$\frac{\partial y}{\partial z} = \frac{\partial y}{\partial g}\frac{\partial g}{\partial z}$$

$$\frac{\partial g}{\partial z} = W_2(W_1(1 - \tanh^2(W_1 z + b_1)))\exp(W_2 \tanh(W_1 z + b_1) + b_2)$$

$$\frac{\partial y}{\partial g} = y(g)(1 - y(g))$$