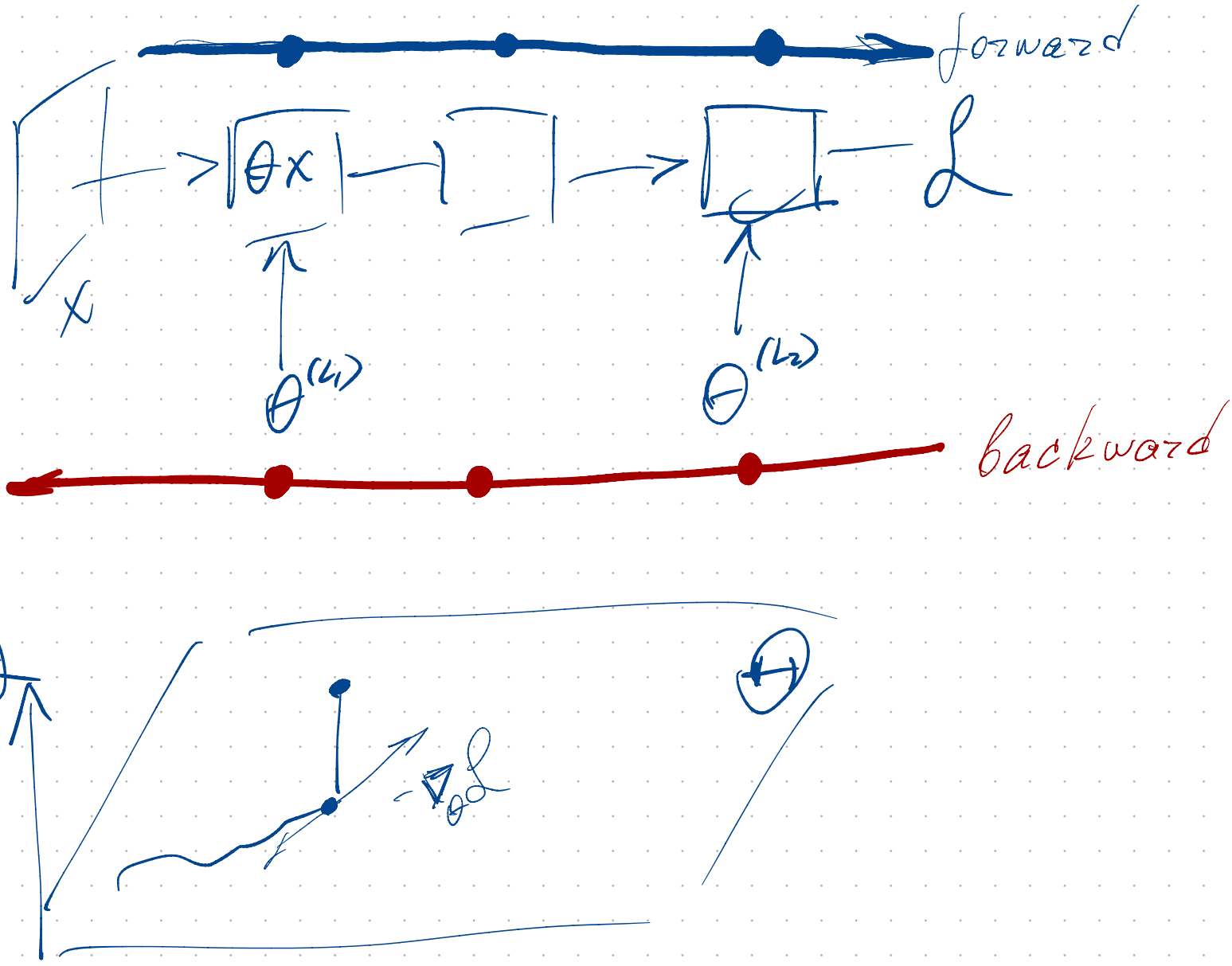
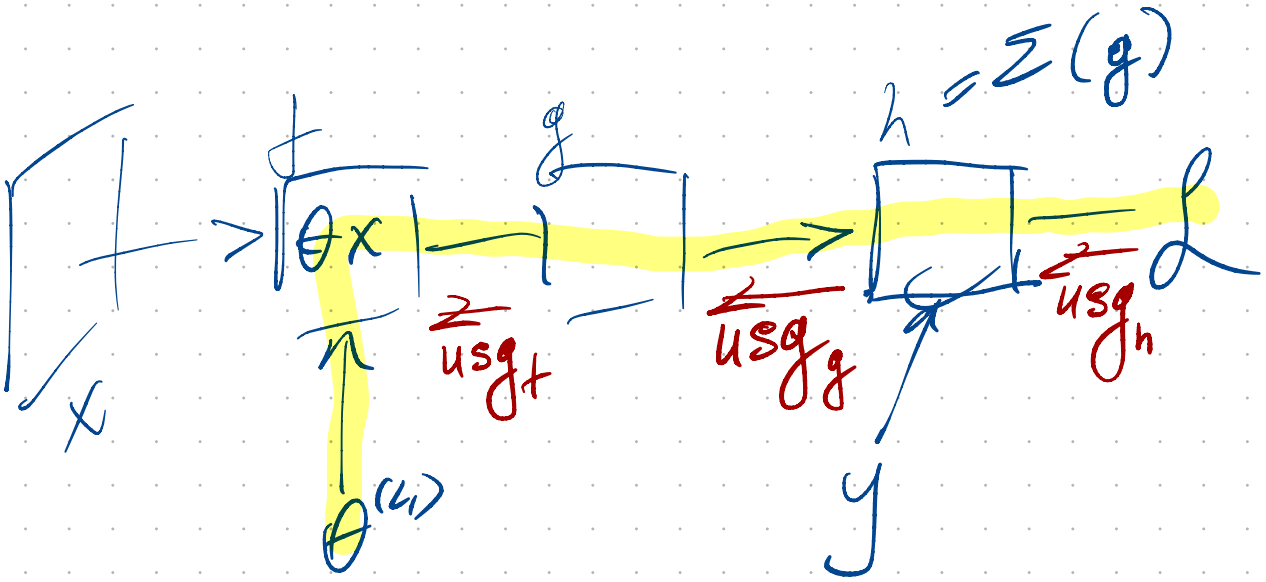

forward
θx
x
$\theta^{(1)}$
$\theta^{(2)}$
$\mathcal{L}$
backward

$\mathcal{J}$
$-\nabla_\theta \mathcal{L}$



$$\ell = h(y, g(f(x, \theta)))$$

$$\frac{\partial \ell}{\partial \theta} = \frac{\partial h}{\partial g} \frac{\partial g}{\partial f} \frac{\partial f}{\partial \theta}$$

f: $\frac{\partial f}{\partial \theta}$ $\frac{\partial f}{\partial x}$ $\frac{\partial \ell}{\partial \theta}$ $\frac{\partial \ell}{\partial x}$

$$\frac{\partial \ell}{\partial \theta} = \text{usg}_f \frac{\partial f}{\partial \theta}$$