Deep Learning Exercise 07 Loss function

- $f(\mu,\sigma) = -\dfrac{(t-\mu)^2}{2\sigma^2} - \log(\sqrt{2\pi}\sigma)$

a) $$\frac{\partial f}{\partial u} = \frac{2t-2u}{2\sigma^2} = \frac{(t-u)}{\sigma^2}$$

$$\frac{\partial f}{\partial \sigma} = \frac{-(t-u)^2}{2} \cdot (-2)\cdot\sigma^{-3} - \left[\frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \sqrt{2\pi}\right]$$

$$= \frac{(t-u)^2}{\sigma^3} - \frac{1}{\sigma}$$

b) $\sigma(z) = \sqrt{1+z^2}$ * nonlinear activation function

$$\frac{\partial\sigma(z)}{\partial z} = \frac{1}{2}\cdot(1+z^2)^{-\frac{1}{2}}\cdot 2z$$

$$= z\cdot(1+z^2)^{-\frac{1}{2}}$$

$$= z\cdot(1+z^2)^{\frac{1}{2}\cdot(-1)}$$

$$= z\cdot\frac{1}{\sigma}$$

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial\sigma}\cdot\frac{\partial\sigma}{\partial z}$$

$$= \left(\frac{(t-u)^2}{\sigma^3} - \frac{1}{\sigma}\right)\cdot\frac{z}{\sigma}$$

$$= \left(\frac{(t-u)^2}{\sigma^4} - \frac{1}{\sigma^2}\right)\cdot z$$

c) $$\left|\frac{\partial f}{\partial z}\right| = \left|\left[\frac{(t-u)^2}{\sigma^4} - \frac{1}{\sigma^2}\right]\cdot z\right|$$

$\sigma \geq 1$ and $\sigma \geq |z|$ $\Rightarrow |z| \leq 1$

$$\Rightarrow \left|\frac{\partial f}{\partial z}\right| \leq \left|\left[\frac{(t-u)^2}{\sigma^4} - \frac{1}{\sigma^2}\right]\right|\cdot|z|$$

$$\leq \left[\left[\frac{(t-u)^2}{\sigma^2} - 1\right]\cdot\frac{1}{\sigma^2}\right]\cdot 1$$

because $\sigma \geq 1 \Rightarrow \sigma^2 \geq 1 \Rightarrow \frac{1}{\sigma^2} \leq 1$

$$\Rightarrow \left|\frac{\partial f}{\partial z}\right| \leq \left|\left[\left[\frac{(t-u)^2}{\sigma^2} - 1\right]\cdot 1\right]\right|$$

$$= \left|\frac{(t-u)^2 - \sigma^2}{\sigma^2}\right|$$

$$\leq \left|(t-u)^2 - \sigma^2\right|$$

rewrite σ as $(1+\varepsilon)$, $\varepsilon \geq 0$

$$\left|\frac{\partial f}{\partial z}\right| \leq \left|(t-u)^2 - (1+\varepsilon)^2\right|$$

$$= \left|(t-u)^2 - 1 - 2\varepsilon - \varepsilon^2\right|$$

becaus $(1+\varepsilon)^2 \geq 1$

$$\Rightarrow \left|\frac{\partial f}{\partial z}\right| \leq \left|(t-u)^2 - 1\right|$$

$$\leq (t-u)^2 + 1$$

d)

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(-\frac{|t_i - u_i|^2}{2\sigma_i^2} - \log(\sqrt{2\pi}\cdot\sigma_i)\right)$$

$$\frac{\partial f}{\partial u} = \frac{(t-u)}{\sigma^2}$$

$$\frac{\partial f}{\partial z} = \left(\frac{(t-u)^2}{\sigma^4} - \frac{1}{\sigma^2}\right)\cdot z \quad \text{, with } \sigma = \sigma(z)$$

$$f_i(u_i, z_i) = \frac{-|t_i - u_i|^2}{2\sigma_i^2} - \log(\sqrt{2\pi}\cdot\sigma_i)$$

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N} f_i(u_i(\theta), z_i(\theta))$$

$$\frac{\partial f_i}{\partial \theta} = \frac{\partial f_i}{\partial u_i}\cdot\frac{\partial u_i}{\partial \theta} + \frac{\partial f_i}{\partial z_i}\cdot\frac{\partial z_i}{\partial \theta}$$

$$\frac{\partial J}{\partial \theta} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\partial f_i}{\partial u_i}\cdot\frac{\partial u_i}{\partial \theta} + \frac{\partial f_i}{\partial z_i}\cdot\frac{\partial z_i}{\partial \theta}\right)$$

(e) $\theta = (w, v)$, $u_i = w^T x_i$, $z_i = v^T \cdot x_i$

$$\frac{\partial u_i}{\partial w} = x_i \quad , \quad \frac{\partial z_i}{\partial v} = x_i$$

$$\frac{\partial u_i}{\partial v} = 0 \quad , \quad \frac{\partial z_i}{\partial w} = 0$$

$$\frac{\partial f_i}{\partial z_i} = \left(\frac{(t_i - u_i)^2}{\sigma_i^4} - \frac{1}{\sigma_i^2}\right)\cdot z_i$$

$$\frac{\partial f_i}{\partial u_i} = \frac{t_i - u_i}{\sigma_i^2}$$

$$\frac{\partial J}{\partial w} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial f_i}{\partial w_i} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial f_i}{\partial u_i}\cdot\frac{\partial u_i}{\partial w}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{(t_i - u_i)}{\sigma_i^2}\cdot x_i = \frac{1}{N}\sum_{i=1}^{N}\frac{(t_i - w^T x_i)}{\sigma_i^2}\cdot x_i$$

$$\frac{\partial J}{\partial v} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial f_i}{\partial v} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial f_i}{\partial z_i}\cdot\frac{\partial z_i}{\partial v}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{(t_i - u_i)^2}{\sigma_i^4} - \frac{1}{\sigma_i^2}\right) z_i x_i$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{(t_i - u_i)^2}{\sigma_i^4} - \frac{1}{\sigma_i^2}\right)\cdot(v^T \cdot x_i)\, x_i$$