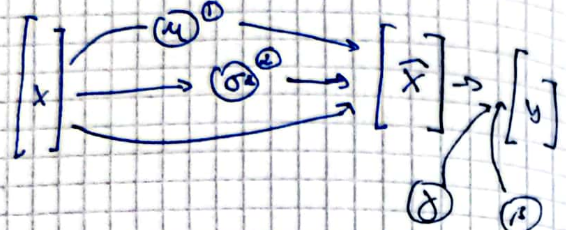
$$\mu = \frac{1}{m}\sum x_i \qquad \text{Have: } \frac{dL}{dy_i}$$

$$\sigma^2 = \frac{1}{m-1}\sum (x_i-\mu)^2 \qquad \text{Find: } \frac{dL}{dx_i}$$

$$\hat{x}_i = \frac{x_i-\mu}{\sqrt{\sigma^2+e}}$$

$$y_i = \gamma \hat{x}_i + \beta$$



① $$\frac{dL}{d\hat{x}_i} = \frac{dL}{dy_i}\cdot\gamma$$

② $$\frac{dL}{d\sigma^2} = \frac{dL}{d\hat{x}_i}\cdot\frac{d\hat{x}_i}{d\sigma^2}$$ $\Big\|$ $\hat{x}$ - vector $[n:1]$; σ^2 - const $\Big\|$

$$\frac{d\hat{x}_i}{d\sigma^2} = \sum_i \frac{d}{d\sigma^2}\left[\frac{x_i-\mu}{\sqrt{\sigma^2+e}}\right] = \sum_i (x_i-\mu)\cdot(\sigma^2+e)^{-\frac{3}{2}}\cdot\left(-\frac{1}{2}\right)$$

$$\frac{dL}{d\sigma^2} = -\frac{1}{2}\gamma\sum_i\left[(x_i-\mu)(\sigma^2+e)^{-\frac{3}{2}}\cdot\frac{dL}{dy_i}\right]$$

③ $$\frac{dL}{d\mu} = \frac{dL}{d\sigma^2}\cdot\frac{d\sigma^2}{d\mu} + \sum_i \frac{dL}{d\hat{x}_i}\cdot\frac{d\hat{x}_i}{d\mu}$$

(3.1) (3.2)

3.1 $$\frac{d\sigma^2}{d\mu} = \frac{1}{m-1}\cdot -2\sum_i (x_i-\mu) \rightsquigarrow \sum_i (x_i-\mu) = \sum_{i}^{m} x_i - m\cdot\mu = 0$$

3.2 $$\frac{d\hat{x}_i}{d\mu} = -1\cdot(\sigma^2+e)^{-\frac{1}{2}}$$

3: $$\frac{dL}{d\mu} = \gamma\sum\frac{dL}{dy_i}\cdot\left(-(\sigma^2+e)^{-\frac{1}{2}}\right)$$

④ $$\frac{dL}{dx_i} = \sum\frac{d\sigma^2}{dx_i} + \sum\frac{d\mu}{dx_i} + \sum\frac{d\hat{x}_i}{dx_i}$$

(4.1) (4.2) (4.3)

4.1 $$\frac{d\sigma^2}{dx_i} = \frac{2}{m-1}\cdot\sum_i (x_i-\mu)$$?

4.2 $$= -\frac{1}{m}\cdot\gamma\sum_i\frac{dL}{dy_i}\cdot(\sigma^2+e)^{-\frac{1}{2}}$$

4.3 $$\frac{d\hat{x}_i}{dx_i} = \frac{1}{\sqrt{\sigma^2+e}}\cdot\frac{dL}{dy_i}\cdot\gamma$$

$$\left[-\frac{1}{m}\cdot\gamma\cdot\sum_i\frac{dL}{dy_i}\cdot(\sigma^2+e)^{-\frac{1}{2}}\right] +$$

$$+\left[\sum_j\frac{dL\cdot\gamma}{\sqrt{\sigma^2+e}\cdot dy_j}\right] + \sum_j (x_j-\mu)\cdot(\sigma^2+e)^{-\frac{1}{2}} = \hat{x}_i$$

$$+\left[\frac{2}{m-1}\cdot\sum_{k}^{m}(x_k-\mu)\cdot -\frac{1}{2}\gamma\sum_i(x_i-\mu)(\sigma^2+e)^{-\frac{3}{2}}\cdot\frac{dL}{dy_i}\right]$$

$$\Big\|\ \hat{x}_i = \frac{x_i-\mu}{\sqrt{\sigma^2+e}}\ \Big\|$$

$$(\sigma^2+e)^{-\frac{1}{2}}\cdot\gamma\left[-\frac{1}{m}\cdot\sum_i\frac{dL}{dy_i} + \sum_i\frac{dL}{dy_i}\right.$$

$$(\sigma^2+e)^{-\frac{1}{2}}\cdot\gamma\left[-\frac{1}{m}\cdot\sum_j\frac{dL}{dy_j} + \sum_j\frac{dL}{dy_j} +\right.$$

$$+\left(\left(-\frac{1}{m-1}\underbrace{\sum_j (x_j-m)(\sigma^2+e)^{-\frac{1}{2}}}_{\sum_j \hat{x}_j}\right)\cdot\left(\underbrace{\sum_k (x_k-m)(\sigma^2+e)^{-\frac{1}{2}}}_{\hat{x}_i}\cdot\frac{dL}{dy_i}\right)\right.$$

$$= (\sigma^2+e)^{-\frac{1}{2}}\cdot\gamma\left[-\frac{1}{m}\cdot\sum_j\frac{dL}{dy_j} + \sum_j\frac{dL}{dy_j} + \left(\frac{\sum_j\hat{x}_j}{m-1}\cdot(-1)\cdot\sum_j\hat{x}_j\cdot\frac{dL}{dy_i}\right)\right.$$