Name Date 13/08/2025 14:01

Batch normalisation backprobg

h prebn → bnraw

(x)

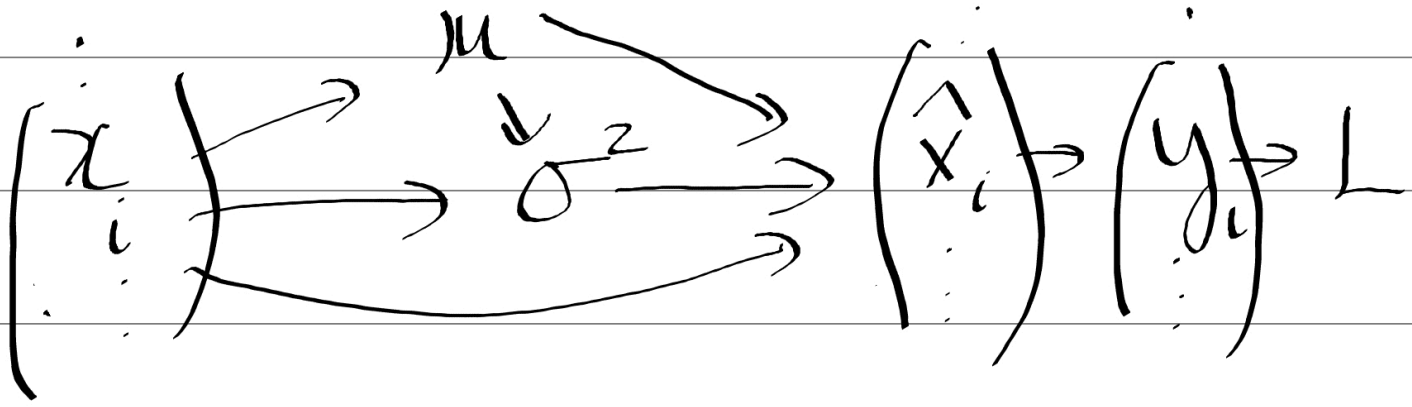
(y)

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu)^2$$

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

$$y_i = \gamma x_i + \beta$$



$$\frac{\partial L}{\partial \hat{x}_i} = \frac{\partial L}{\partial y_i}\frac{\partial y_i}{\partial \hat{x}_i} = \frac{\partial L}{\partial y_i}\cdot\gamma$$

$$\frac{\partial L}{\partial x_i} = \frac{\partial L}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial x_i} + \frac{\partial L}{\partial \mu}\frac{\partial \mu}{\partial x_i}$$

$$+ \frac{\partial L}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial x_i}$$

$$\frac{\partial L}{\partial \mu} = \frac{\partial L}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial \mu} + \frac{\partial L}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial \mu}$$

$$\frac{\partial L}{\partial \sigma^2} = \frac{\partial L}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial \sigma^2}$$