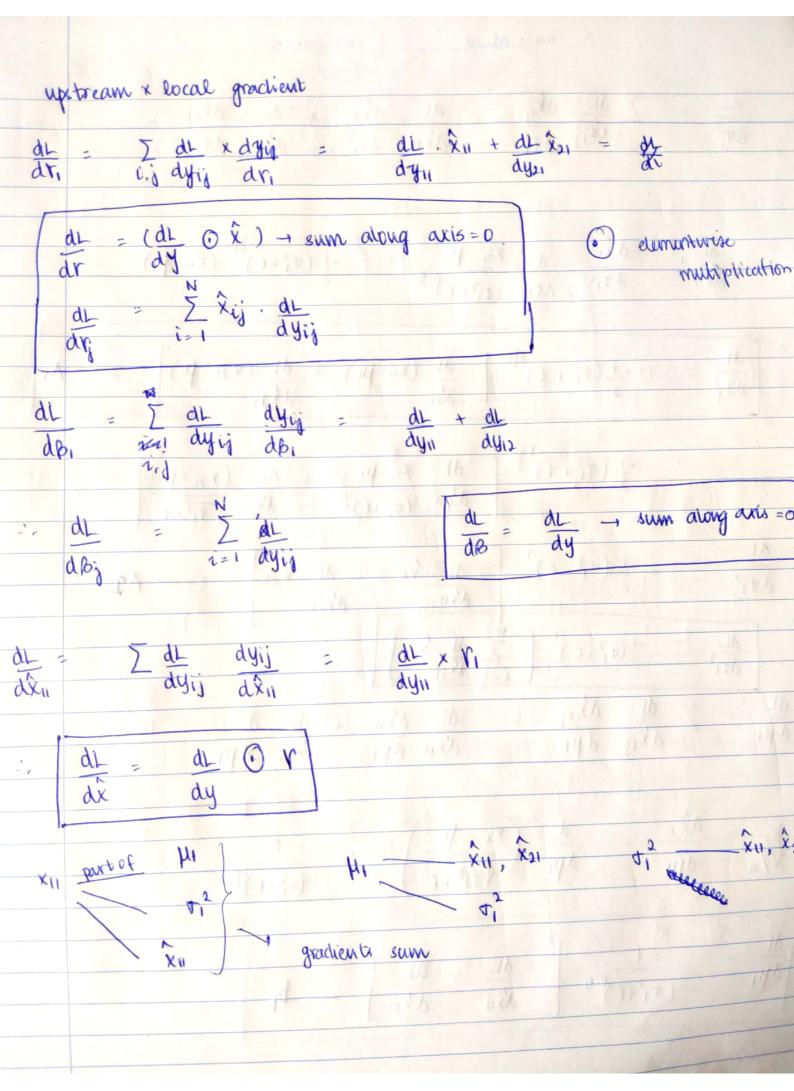
	Bo	atch Norm	nalization:	: Back-propagation
		X S X S X S X S X S X S X S X S X S X S		$\frac{1+x_{21}}{2}$ , $\frac{x_{12}+x_{22}}{2}$ , $\frac{x_{13}+x_{23}}{2}$ $\frac{N}{2}$ $\frac{1}{N}$ $\frac{1}{N}$ $\frac{1}{N}$ $\frac{1}{N}$ $\frac{1}{N}$ $\frac{1}{N}$ $\frac{1}{N}$ $\frac{1}{N}$
4,5	2 ((			$\frac{1}{2} \left[ (x_{12} - \mu_2)^2 + (x_{22} - \mu_2)^2 \right],$ $\frac{1}{2} \left[ (x_{12} - \mu_2)^2 + (x_{22} - \mu_2)^2 \right]$
٠, ١	1	N	μ <sub>1</sub> ) <sup>2</sup> , <u>ι</u>	$\frac{1}{N} \sum_{i=1}^{N} (x_{i2} - \mu_2)^2 - \dots$
X.	=	XII-HI	100x+0	×18-μ3 √ σ3 <sup>2</sup> + €
		12/27 E	123 F(	
4 =	= 1, x11 + p1		r2 x12 + B	B2 13 x13 + B3
			V2×22 + (	β2 V8 x29 + β3
dl =	<u>dr</u> dyn	dL dyn	dL dy13	
	dyn	dl dys	dL dy13	



Scanned with CamScanner

$$\frac{dL}{d\tau_{1}^{2}} = \frac{2}{dL} \frac{dk_{1}}{dx_{1}^{2}} = \frac{dL}{dx_{1}} \frac{dx_{1}}{dx_{1}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{dx_{1}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{dx_{1}^{2}}$$

$$= \frac{dL}{dx_{1}} \frac{dx_{1}}{dx_{1}} = \frac{d}{dx_{1}} \frac{x_{1} - \mu_{1}}{(x_{1}^{2} + e)^{3/2}} = \frac{-1}{3} (x_{1}^{2} + e)^{3/2} (x_{1} - \mu_{1})$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1} - \mu_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} (x_{2} - \mu_{1}) \right] - A_{1}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1} - \mu_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} (x_{2} - \mu_{1}) \right] - A_{1}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1} - \mu_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} \frac{(x_{2} - \mu_{1})}{(x_{2}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{(x_{2}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{(x_{2}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{(x_{2}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{(x_{1}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} + \frac{dL}{dx_{2}} \frac{dx_{2}}{(x_{1}^{2} + e)^{3/2}} - A_{2}$$

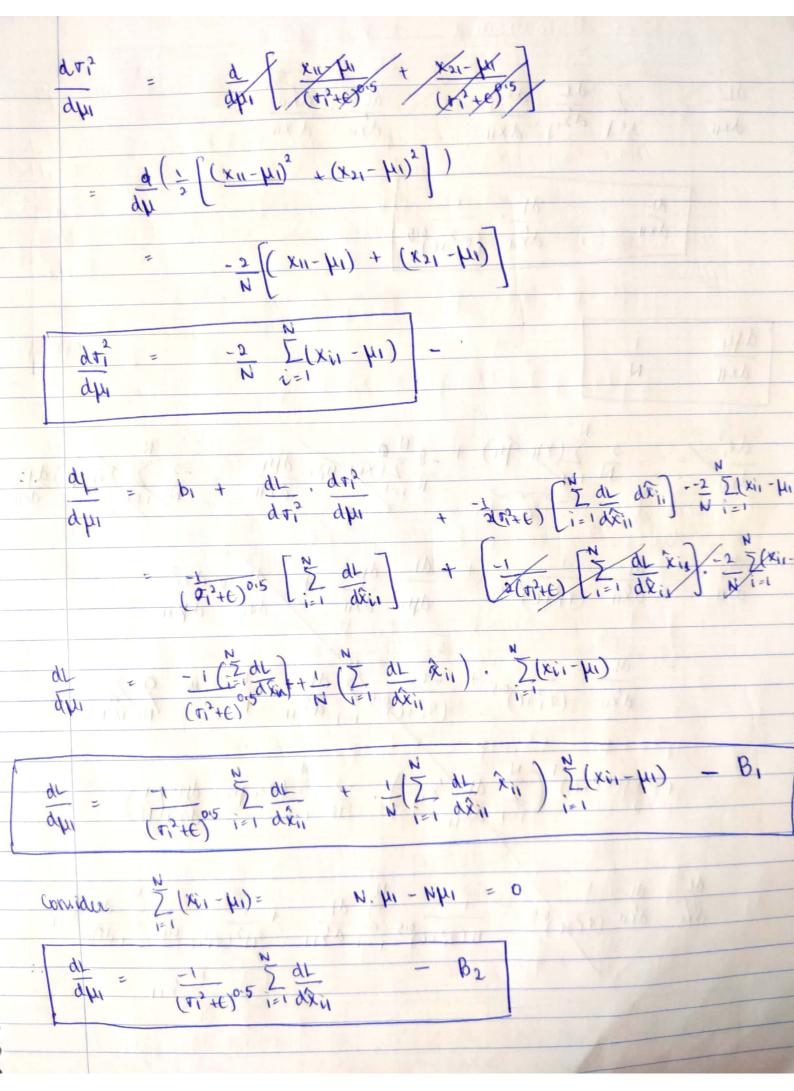
$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} - \frac{dL}{dx_{2}} \frac{x_{2}}{(x_{1}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} - \frac{dL}{dx_{2}} \frac{x_{2}}{(x_{1}^{2} + e)^{3/2}} - A_{2}$$

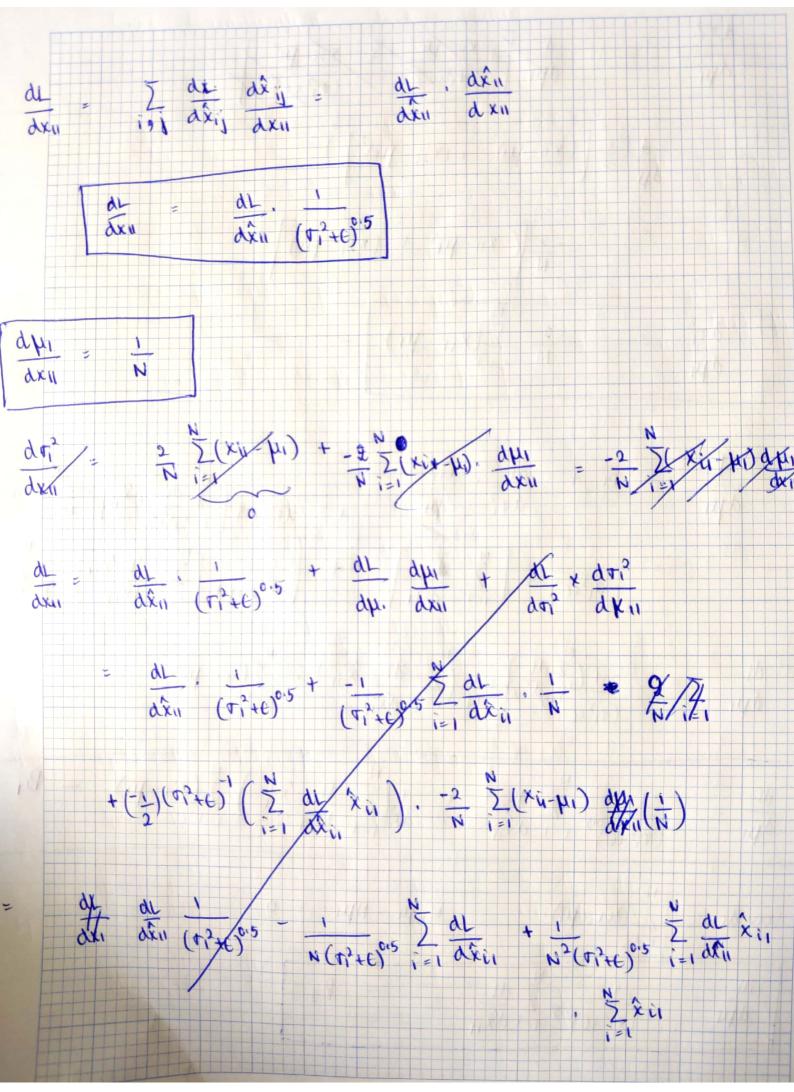
$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} - \frac{dL}{dx_{2}} \frac{x_{2}}{(x_{1}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{x_{1}}{(x_{1}^{2} + e)^{3/2}} - \frac{dL}{dx_{2}} \frac{x_{2}}{(x_{1}^{2} + e)^{3/2}} - A_{2}$$

$$= \frac{-1}{3} (x_{1}^{2} + e)^{3/2} \left[ \frac{dL}{dx_{1}} \frac{$$



Scanned with CamScanner



Scanned with CamScanner

