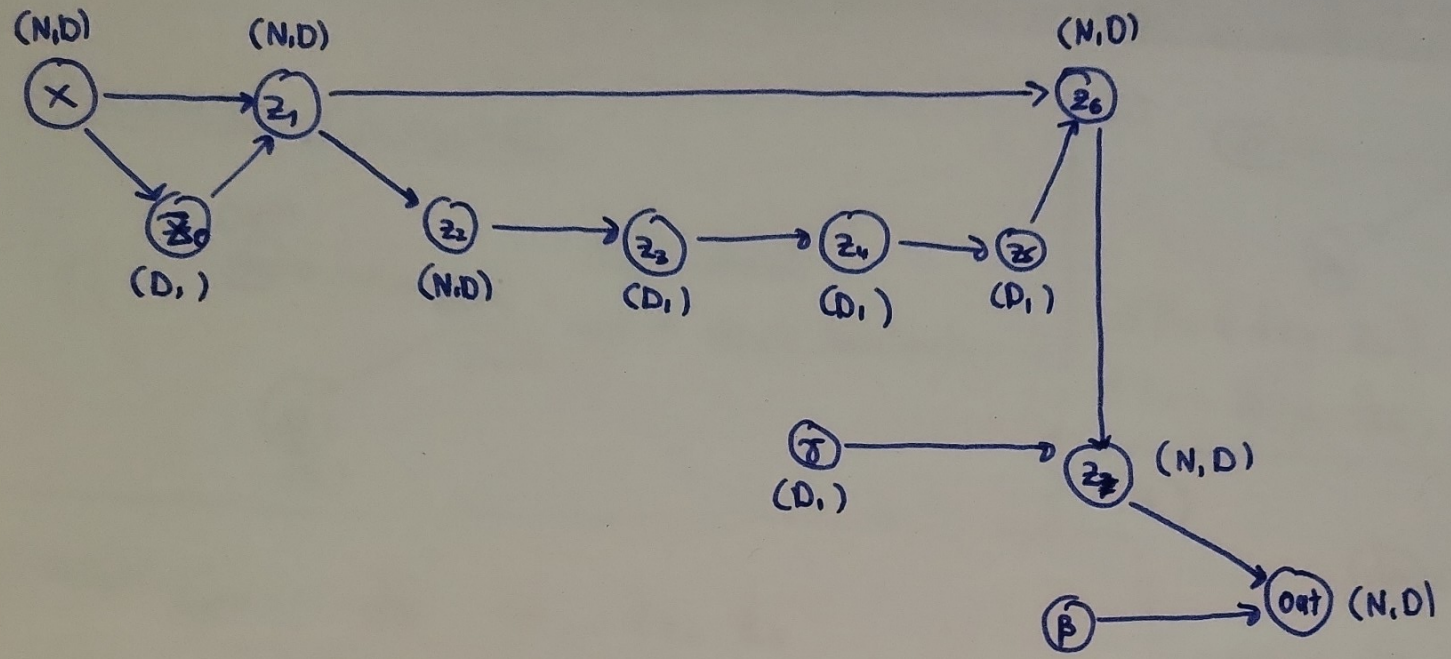


BATCHNORM FORWARD

$\bar{x}_0 = \frac{1}{N}\sum x_i$

$z_1 = x - \bar{x}_0$

$z_2 = z_1^2$

$z_3 = \frac{1}{N}\sum z_{2,i}$

$z_4 = \sqrt{z_3 + \varepsilon}$

$z_5 = \frac{1}{z_4}$

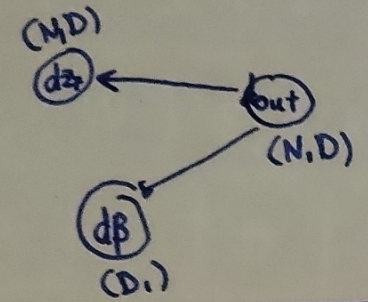
$z_6 = z_1 \cdot z_5$

$z_7 = z_6 \cdot \gamma$

$out = \beta + z_7$

BATCHNORM BACKWARD

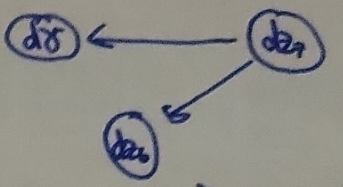
1)



$dz_4 = dout$

$d\beta = dout.sum(axis=0)$

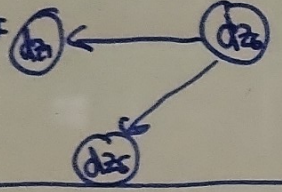
2)



$d\gamma = (dz_7 \cdot z_6).sum(axis=0)$

$dz_6 = \gamma \cdot dz_7$

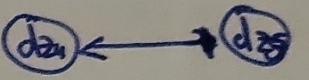
3)



$dz_1 = dz_6 \cdot z_5$

$dz_5 = (z_1 \cdot dz_6).sum(axis=0)$

4)

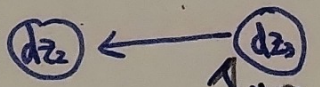


$dz_4 = dz_5 \cdot \frac{-1}{z_4^2}$

5)

$dz_3 = dz_4 \cdot \frac{1}{z}$

6)



$dz_2 = \frac{1}{N} \cdot dz_3$

7) $dz_1 \pm 2z_1 \cdot dz_2$

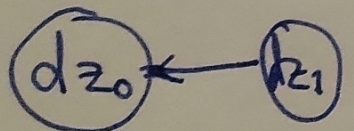
(i.e. $dz_1 = dz_0 \cdot z_5 + 2z_1 \cdot dz_2$)

1 NxD

$dx = \frac{1}{N} \cdot \gamma \cdot dz_0 +$

$+ dz_1$

8)



$dz_0 = -(dz_1).sum(axis=0)$