Name: Waqar Kaleem Khan

Enrollment No: 01-249191-013

Assignment No: 4

Date: 30-Apr-2020

Teacher: Dr Imran Siddiqi

Class: MSDS(3-A)

Course: Deep Learning

Name : Waqar kaleem Khan
Enrolment # 01-249191-013
Class : MS(DS) 3-A
Assignment # 04
Date: 30/04/2020
Teacher : Dr. Imran Siddiqi

Task 1: Regularization

Regularization Technique — weight update eq

(1) No-Regularization — $w = w - \eta \frac{dJ_0}{dw}$

Updating W_1 weights

Sol:-

1st iteration (values are given in Assignm)

$w_{11} = 0.2 - (1)(0.2)$

$w_{11} = 0$

2nd iteration:-

$w_{12} = 0 - (1)(0.1)$

$w_{12} = -0.1$

3rd iteration:-

$w_{13} = -0.1 - (1)(0.3)$

$= -0.1 - 0.3$

$= -0.4$

4th iteration

$w_{14} = -0.4 - (1)(0.5)$

$w_{14} \Rightarrow -0.4 - 0.5$

$w_{14} \Rightarrow -0.9$

Fifth iteration

$w_{15} = -0.9 - (1)(0.1)$

$w_{15} = -0.9 - 0.1$

$w_{15} = -1$

Now updating w_2 weights

1st iteration

$w_{21} = -0.2 - (1)(0.2)$

$w_{21} = -0.4$

2nd iteration

$w_{22} = -0.4 - (1)(0.8)$

$w_{22} = -0.4 - 0.8$

$w_{22} = -1.2$

3rd iteration

$w_{23} = -1.2 - (1)(0.2)$

$w_{23} = -1.2 - 0.2$

$w_{23} = -1.4$

4th iteration

$w_{24} = -1.4 - (1)(0.5)$

$w_{24} = -1.4 - 0.5$

$w_{24} = -1.9$

5th iteration

$w_{25} = -1.9 - (1)(0.1)$

$w_{25} = -1.9 - 0.1$

$w_{25} = -2.0$

2) Regularization Technique
L1 Regularization

weight update equation
$w = w - \eta \frac{\lambda}{m} \text{sign}(w) - \eta \frac{dJ_0}{dw}$

Sol

updating W_1 weights

1st iteration:-

$w_{11} = 0.2 - (1 * 0.5/10 * (+1)(0.2)$
$w_{11} = 0.2 - (1 * 0.05) - 0.2$
$w_{11} = 0.2 - 0.05 - 0.2$
$w_{11} = -0.05$

2nd iteration:-

$w_{12} = -0.05 - (1 * 0.5/10 * (-1)) - 1(0.1$
$w_{12} = -0.05 - (1 * 0.05) - 0.1$
$w_{12} = -0.05 + 0.05 - 0.1$
$w_{12} = -0.1$

3rd iteration

$w_{13} = -0.1 - (1 * 0.5/10 * (-1)) - 1(0.3)$
$w_{13} = -0.1 - (1 * -0.05) - 0.3$
$w_{13} = -0.1 + 0.05 - 0.3$
$w_{13} = -0.35$

4th iteration

$w_{14} = -0.35 - (1 * 0.5/10 * (-1)) - 1(0.5)$
$w_{14} = -0.35 - (1 * -0.05) - 0.5$
$w_{14} = -0.35 + 0.05 - 0.5$
$w_{14} = -0.8$

5th iteration

$w_{15} = -0.8 - (1 * 0.5/10 * (-1)) - 1(0.1)$

$w_{15} = -0.8 - (1 * 0.05) - 0.1$

$w_{15} = -0.8 + 0.05 - 0.1$

$w_{15} = -0.85$

Now updating weights w_2 (L_1 Regularization)

1st iteration

$w_{21} = -0.2 - (1 * 0.5/10) * (-1)) - 1(0.2)$

$w_{21} = -0.2 - (1 * 0.05) - 0.2$

$w_{21} = 0.2 + 0.05 - 0.2$

$w_{21} = -0.35$

2nd iteration

$w_{22} = -0.35 - (1 * 0.5/10 * (-1)) - 1(0.8)$

$w_{22} = -0.35 - (1 * 0.05) - 0.8$

$w_{22} = -0.35 + 0.05 - 0.8$

$w_{22} = -1.1$

3rd iteration

$w_{23} = -1.1 - (1 * 0.5/10 * (-1)) - 1(0.2)$

$w_{23} = -1.1 (1 * -0.05) - 0.2$

$w_{23} = -1.1 + (0.05) - 0.2$

$w_{23} = -1.25$

4th iteration

$w_{24} = -1.25 - (1 * 0.5/10 * (-1) - 1(0.5)$

$w_{24} = -1.25 - (1 * -0.05) - 0.5$

$w_{24} = -1.25 + 0.05 - 0.5$

$w_{24} = -1.7$

5th iteration

$w_{25} = -1.7 - (1 * 0.5/10 * (-1)) - 1(0.1)$

$w_{25} = -1.7 - (1 * -0.05) - 0.1$

$w_{25} = -1.7 + 0.05 - 0.1$

$w_{25} = -1.75$

Regularization Technique — L_2 Regularization

weight update eq

$w = w - \eta \frac{\lambda}{m} w - \eta \frac{dj_0}{dw}$

Sol:-

Updating weights for w_{11}

1st iteration:-

$w_{11} = 0.2 - (1 * 0.5/10 * 0.2) - 1(0.2)$

$w_{11} = 0.2 - (1 * 0.05 * 0.2) - 0.2$

$w_{11} = 0.2 - 0.01 - 0.2$

$w_{11} = -0.01$

2nd iteration

$w_{12} = -0.01 - (1 * 0.5/10 * -0.01) - 1(0.1)$

$w_{12} = -0.01 - (1 * 0.05 * -0.01) - 0.1$

$w_{12} = -0.01 + 0.0005 - 0.01$

$w_{12} = -0.1095$

3rd iteration

$w_{13} = -0.1095 - (1 * 0.5/10 * 0.1095) - 1(0.3)$

$w_{13} = -0.1095 - (1 * 0.05) - 0.3$

$w_{13} = -0.1095 - 0.0145 - 0.3$

$w_{13} = -0.424$

4th iteration.

$w_{14} = (-0.424 - (1 * 0.5/10 * -0.424)) - 1(0.5)$

$w_{14} = -0.424 - (1 * -0.05 * -0.424) - 0.5$

$w_{14} = -0.424 + 0.0212 - 0.5$

$w_{14} = -0.9028$

5th iteration:-

$w_{15} = (-0.9028 - (1 * 0.5/10 * 0.9028)) - 1(0.1)$

$w_{15} = -0.9028 - (1 * 0.05 * 0.9028) - 0.1$

$w_{15} = -0.9028 + 0.04514 - 0.1$

$w_{15} = -0.95766$

$w_{15} = -0.9577$

Now updating weights for w_2

1st iteration:-

$w_{21} = -0.2 - (1 * 0.5/10 * -0.2) - 1(0.2)$

$w_{21} = 0.2 - (1 * 0.05 * -0.2) - 0.2$

$w_{21} = 0.2 + 0.01 - 0.2$

$w_{21} = 0.01$

2nd iteration

$w_{22} = 0.01 - (1 * 0.5/10 * 0.1) - 1(0.8)$

$w_{22} = 0.01 - (1 * 0.05 * 0.1) - 0.8$

$w_{22} = 0.01 - 0.0005 - 0.8$

$w_{22} = 0.7895$

3rd iteration

$w_{23} = -0.7895 - (1 * 0.5/10 * 0.7895) - 1(0.2)$

$w_{23} = -0.7895 - (1 * 0.05 * 0.7895) - 0.2$

$w_{23} = -0.7895 + 0.039475 - 0.3$

$w_{23} = -1.0289$

4th iteration

$w_{24} = -1.0289 - (1 * 0.5/10 * 1.0289) - 1(0.5)$

$w_{24} = -1.0289 - (1 * 0.05 * 1.0289) - 0.5$

$w_{24} = -1.0289 + 0.05144 - 0.5$

$w_{24} = -1.4775$

5th iteration

$w_{25} = -1.4775 - (1 * 0.5/10 * -1.4775) - 1(0.1)$

$w_{25} = -1.4775 - (1 * 0.05 * -1.4775) - 0.1$

$w_{25} = -1.4775 + 0.07387 - 0.1$

$w_{25} = -1.503625$

No Regularization, L1 Regularization, L2 Regularization all iteration W_1 Results are below in table

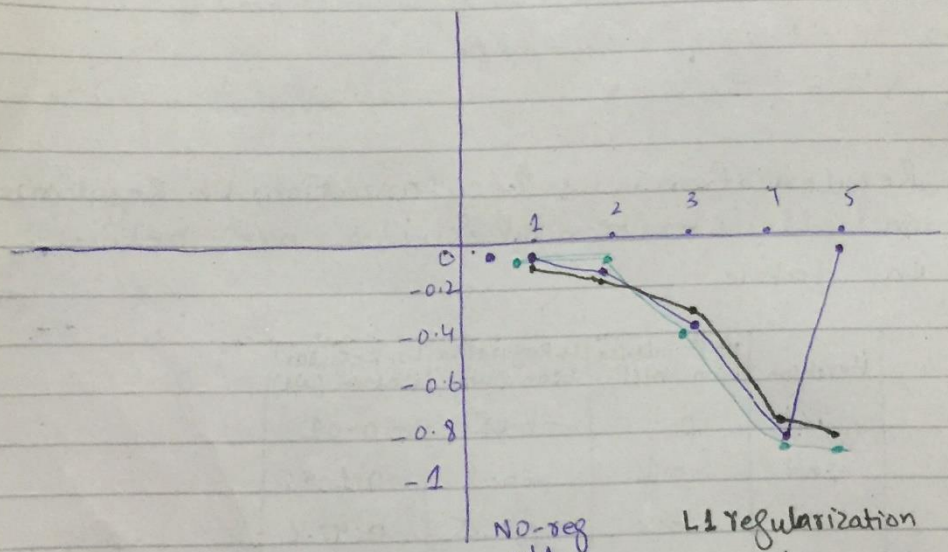
iteration	No Regularization (W_1)	L1 Regularization (W_1)	L2 Regularization (W_1)
1st	0	-0.05	-0.01
2nd	-0.1	-0.1	-0.1095
3rd	-0.4	-0.35	-0.424
4th	-0.9	-0.8	-0.9028
5th	-0.1	-0.85	-0.957

table 1 Results of W_1

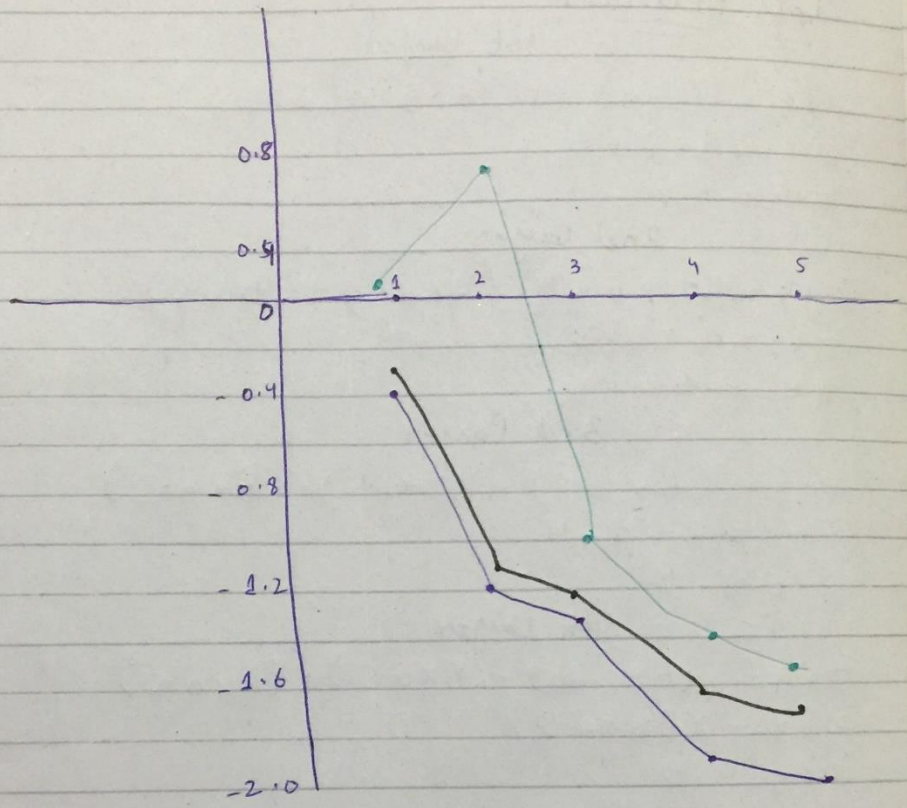
W2 Results

Iteration	W2 No Regularization	L1 Regularization (W2)	L2 Regularization (W2)
1st	-0.4	-0.35	0.01
2nd	-1.2	-1.1	0.78
3rd	-1.4	-1.25	-1.0289
4th	-1.9	-1.7	-1.477
5th	-2.0	-1.75	-1.503

Updated W1 graph



updated w2 graph



NO Regularization L1 Reg L2 Reg
W2 W2 W2

Task 2 : Batch Normalization.

Top Neuron :-

1st Layer

$z_{t_1} = (2*0.5) + (4*0.5) + (6*0.5)$

$z_{t_1} = 1 + 2 + 3$

$z_{t_1} = 6$

2nd Layer

$z_{t_2} = (4*0.5) + (4*0.5) + (4*0.5)$

$z_{t_2} = 2 + 2 + 2$

$z_{t_2} = 6$

3rd Layer

$z_{t_3} = (6*0.5) + (2*0.5) + (8*0.5)$

$z_{t_3} = 3 + 1 + 4$

$z_{t_3} = 8$

4th Layer

$z_{t_4} = (6*0.5) + (8*0.5) + (6*0.5)$

$z_{t_4} = (3 + 4 + 3)$

$z_{t_4} = 10$

5th Layer

$z_{t_5} = (2*0.5) + (2*0.5) + (2*0.5)$

$z_{t_5} = (1 + 1 + 1)$

$z_{t_5} = 3$

Batch Mean for Top Neuron (μ)

$\mu = 1/5 (6 + 6 + 8 + 10 + 3)$

$\mu = 33/5$

$\mu = 6.6$

Batch variance for Top Neuron (δ) $(z_{t_n} - \mu)$

$z_{t_1} = (6 - 6.6) \Rightarrow (0.6)^2$

$z_{t_1} = 0.36$

$z_{t2} = (6-6.6)^2$
$z_{t2} = (-0.6)^2$
$z_{t2} = (0.36)$

$z_{t3} = (8-6.6)^2$
$z_{t3} = (1.4)^2$
$z_{t3} = (1.96)$

$z_{t4} = (10-6.6)^2$
$z_{t4} = (3.4)^2$
$z_{t4} = 11.56$

$z_{t5} = (3-6.6)^2$
$z_{t5} = (-3.6)^2 =$
$z_{t5} = 12.96$

$\delta^2 = 1/5\,(0.36+0.36+1.96+11.56+12.96)$
$\delta^2 = 1/5\,(27.2)$
$\delta^2 = 5.44$

Now z-norm for the top neuron.

$z_{t1} = 6-6.6/\sqrt{5.44+0.01}$
$z_{t1} = -0.6/2.33$
$z_{t1} = -0.26$

$z_{t2} = 6-6.6/\sqrt{5.44+0.01}$
$z_{t2} = -0.6/2.33$
$z_{t2} = -0.26$

$z_{t3} = 8 - 6.6 / \sqrt{5.44 + 0.01}$

$z_{t3} = 1.4 / 2.33$

$z_{t3} = 0.6$

$z_{t4} = 10 - 6.6 / \sqrt{5.44 + 0.01}$

$z_{t4} = 3.4 / 2.33$

$z_{t4} = 1.46$

$z_{t5} = 3 - 6.6 / \sqrt{5.44 + 0.01}$

$z_{t5} = -3.6 / \sqrt{5.44 + 0.01}$

$z_{t5} = -3.6 / 2.33$

$z_{t5} = -1.5$

Now find $\tilde{z}$ for top neuron

$\tilde{z}_{t1} = 1(-0.26) + 5$

$\tilde{z}_{t1} = -0.26 + 5$

$\tilde{z}_{t1} = 4.74$

Now Apply Relu activation function to $\tilde{z}_{t1}$

$\tilde{z}_{t1} = \text{Relu}(4.74)$

$\tilde{z}_{t1} = 4.74$

$\tilde{z}_{t2} = 1(-0.26) + 5$

$\tilde{z}_{t2} = 0.26 + 5$

$\tilde{z}_{t2} = 4.74$

Apply Relu to $\tilde{z}_{t2}$

$\tilde{z}_{t2} = \text{Relu}(4.74)$

$\tilde{z}_{t2} = 4.74$

$\tilde{z}_{t3} = 1(0.6) + 5$

$\tilde{z}_{t3} = 0.6 + 5$

$\tilde{z}_{t3} = 6.5$

Apply Relu to $\tilde{z}_{t3}$

$\tilde{z}_{t3} = Relu(6.5)$

$\tilde{z}_{t3} = 6.5$

$\tilde{z}_{t4} = 1(1.46) + 5$

$\tilde{z}_{t4} = 1.46 + 5$

$\tilde{z}_{t4} = 6.46$

Apply Relu to $\tilde{z}_{t4}$

$\tilde{z}_{t4} = Relu(6.46)$

$\tilde{z}_{t4} = 6.46$

$\tilde{z}_{t5} = 1(-1.55) + 5$

$\tilde{z}_{t5} = -1.55 + 5$

$\tilde{z}_{t5} = 3.45$

Apply Relu to $\tilde{z}_{t5}$

$\tilde{z}_{t5} = Relu(3.45)$

$\tilde{z}_{t5} = 3.45$

Now Bottom Neuron

$z_{b1} = (2 \times 1) + (4 \times 1) + (6 \times 1)$
$z_{b1} = (2) + (4) + (6)$
$z_{b1} = 12$

$z_{b2} = (4 \times 1) + (4 \times 1) + (4 \times 1)$
$z_{b2} = 4 + 4 + 4$
$z_{b2} = 12$

$z_{b3} = (6 \times 1) + (2 \times 1) + (8 \times 1)$
$z_{b3} = 6 + 2 + 8$
$z_{b3} = 16$

$z_{b4} = (6 \times 1) + (8 \times 1) + (6 \times 1)$
$z_{b4} = 6 + 8 + 6$
$z_{b4} = 20$

$z_{b5} = (2 \times 1) + (2 \times 1) + (2 \times 1)$
$z_{b5} = 2 + 2 + 2$
$z_{b5} = 6$

Now calculating Batch mean for bottom neuron

$\mu = 1/5\,(12 + 12 + 16 + 20 + 6)$
$\mu = 66/5$
$\mu = 13.2$

Now calculating Batch variance for Bottom neuron.

$z_{b1} = (12 - 13.2)^2 = (-1.2)^2 = 1.44$

$zb_2 = (12 - 13.2)^2 \Rightarrow (-1.2)^2$

$zb_2 = 1.44$

$zb_3 = (16 - 13.2)^2$

$zb_3 = (2.8)^2$

$zb_3 = 7.84$

$zb_4 = (20 - 13.2)^2$

$zb_4 = (6.8)^2$

$zb_4 = 46.24$

$zb_5 = (6 - 13.2)^2$

$zb_5 = (-7.2)^2$

$zb_5 = 51.8$

$\delta^2 = 1/5(1.44 + 1.44 + 7.84 + 46.24 + 51.8)$

$\delta^2 = 1/5(108.8)$

$\delta^2 = 21.76$

z-norm for bottom neuron.

$zb_1 = 12 - 13.2 / \sqrt{21.76 + 0.01}$

$zb_1 = -1.2 / \sqrt{21.7}$

$zb_1 = -1.2 / 4.7$

$zb_1 = -0.26$

$zb_2 = 12 - 13.2 / \sqrt{21.76 + 0.01}$

$zb_2 = -1.2 / 4.7$

$zb_2 = -0.26$

$z_{b3} = 16 - 13.2 / \sqrt{21.76 + 0.01}$

$z_{b3} = 2.8 \sqrt{21.77}$

$z_{b3} = 2.8 / 4.7$

$z_{b3} = 0.6$

$z_{b4} = 20 - 13.2 / \sqrt{21.76 + 0.01}$

$z_{b4} = 6.8 \sqrt{21.77}$

$z_{b4} = 6.8 / 4.7$

$z_{b4} = 1.45$

$z_{b5} = 6 - 13.2 / \sqrt{21.76 + 0.01}$

$z_{b5} = -7.2 / \sqrt{21.77}$

$z_{b5} = -7.2 / 4.7$

$z_{b5} = -1.53$

Now $\tilde{z}$ for Bottom Neuron

$\tilde{z}_{b1} = 3(-0.26) + 8$

$= -0.78 + 8$

$= 7.2$

Apply Relu to $\tilde{z}_{b1}$

$\tilde{z}_{b1} = Relu(7.2)$

$\tilde{z}_{b1} = 7.2$

$\tilde{z}_{b2} = 3(-0.26) + 8$

$\tilde{z}_{b2} = -0.78 + 8$

$z_{b2} = 7.2$

Apply Relu to $\tilde{z}_{b2}$

$\tilde{z}_{b2} = Relu(7.2)$

$\tilde{z}_{b2} = 7.2$

$\tilde{z}_{b3} = 3(0.6) + 8$

$\tilde{z}_{b3} = 1.88 + 8$

$\tilde{z}_{b3} = 9.8$

Apply Relu function to $\tilde{z}_{b3}$

$\tilde{z}_{b3} = Relu(9.8)$

$\tilde{z}_{b3} = 9.8$

$\tilde{z}_{b4} = 3(1.45) + 8$

$\tilde{z}_{b4} = 4.35 + 8$

$\tilde{z}_{b4} = 12.35$

Apply Relu function to $\tilde{z}_{b4}$

$\tilde{z}_{b4} = Relu(12.35)$

$\tilde{z}_{b4} = 12.35$

$\tilde{z}_{b5} = 3(1.53) + 8$

$\tilde{z}_{b5} = 4.59 + 8$

$\tilde{z}_{b5} = 12.59$

Apply Relu function to $\tilde{z}_{b5}$

$\tilde{z}_{b5} = Relu(12.59)$

$\tilde{z}_{b5} = 12.59$

Task 3 Computing Exponentially weighted Average

Batch size	u_i	$(1-\beta)*\beta^i$	$(u_i*(1-\beta)*\beta^i)$
5	20	$(1-0.9)*0.9^0 = 0.1$	$20*0.1 = 2$
4	6	$(1-0.9)*0.9^1 = 0.09$	$6*0.09 = 0.54$
3	15	$(1-0.9)*0.9^2 = 0.081$	$15*0.081 = 1.215$
2	8	$(1-0.9)*0.9^3 = 0.0729$	$8*0.0729 = 0.5832$
1	10	$(1-0.9)*0.9^4 = 0.06561$	$10*0.06561 = 0.6561$

Exp weighted Average = 4.993