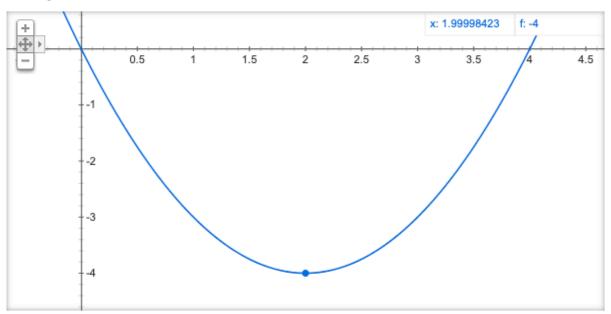
lab

จงทดลอง ปรับค่า w, loop, alpha เพื่อหา w ที่ทำให้ มีค่า Loss (Error) น้อยที่สุด โดยให้ plot กราฟค่า w และ loss ประกอบการศึกษา เพื่อตอบคำถามต่อไปนี้

1. ควรกำหนดค่า w เท่าไหร่

ควรกำหนดค่า w เข้าใกล้กับ 2 ให้มากที่สุด เนื่อจากทำให้เกิดค่า loss น้อยที่สุด

Graph for x^2-4*x

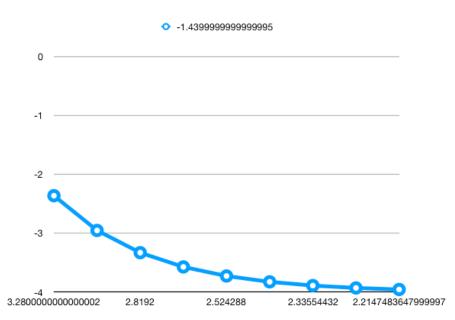


More info

2. ค่า loop มาก/น้อย มีผลกับ w และ loss อย่างไร

begin1

3.6	-1.439999999999995
3.2800000000000002	-2.3615999999999993
3.024	-2.9514239999999994
2.8192	-3.32891136
2.65536	-3.5705032703999997
2.524288	-3.725122093056
2.4194304	-3.82407813955584
2.33554432	-3.8874100093157375
2.2684354559999997	-3.927942405962072
2.2147483647999997	-3.9538831398157264

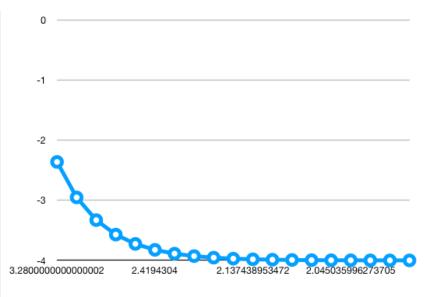


รูปที่2 w = 4 loop = 20 alpha = 0.1

begin1

3.6	-1.439999999999995
3.2800000000000002	-2.3615999999999993
3.024	-2.9514239999999994
2.8192	-3.32891136
2.65536	-3.5705032703999997
2.524288	-3.725122093056
2.4194304	-3.82407813955584
2.33554432	-3.8874100093157375
2.2684354559999997	-3.927942405962072
2.2147483647999997	-3.9538831398157264
2.17179869184	-3.9704852094820646
2.137438953472	-3.981110534068521
2.1099511627776	-3.9879107418038533
2.08796093022208	-3.9922628747544664
2.070368744177664	-3.995048239842858
2.0562949953421312	-3.9968308734994293
2.045035996273705	-3.9979717590396344
2.0360287970189637	-3.998701925785366
2.028823037615171	-3.9991692325026342
2.023058430092137	-3.999468308801686

-1.4399999999999995



จากรูปที่ 1 มี <u>loop = 10</u> จะมีค่า w คือ 2.2147483647999997 และค่า loss คือ -3.9538831398157264 ส่วนรูปที่ 2 มี <u>loop = 20</u> จะมีค่า w คือ 2.023058430092137 และค่า loss คือ -3.999468308801686

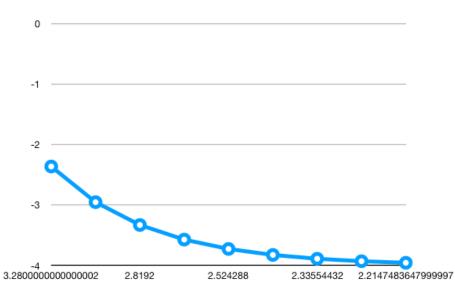
ดังนั้น ถ้ามีจำนวน loop มากจะทำพบให้ค่า w ที่ทำให้เกินค่า loss น้อยที่สุด มากกว่า ที่ มีจำนวน loop น้อย

3.alpha มาก/น้อย มีผลกับ w อย่างไร และเพราะอะไร รูปที่ 1 w = 4 loop = 10 alpha = 0.1

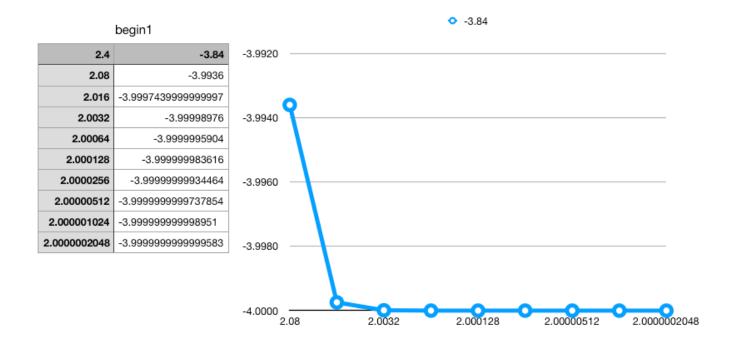
neamni	

3.6	-1.439999999999995
3.2800000000000002	-2.3615999999999993
3.024	-2.9514239999999994
2.8192	-3.32891136
2.65536	-3.5705032703999997
2.524288	-3.725122093056
2.4194304	-3.82407813955584
2.33554432	-3.8874100093157375
2.2684354559999997	-3.927942405962072
2.2147483647999997	-3.9538831398157264

-1.4399999999999995



ฐปที่ 2 w = 4 loop = 10 alpha = 0.4



รูปที่ 3 w = 10 loop = 10 alpha = 0.1

be	gin1			36.96	
8.4	36.96	30.0000			
7.12	22.2144				
6.096	12.777216000000003	22.5000			
5.2768	6.737418239999997	22.5000	٩		
4.62144	2.8719476735999976				
4.0971519999999995	0.398046511103999	15.0000			
3.6777215999999995	-1.1852502328934413		٩		
3.3421772799999996	-2.1985601490518025	7.5000			
3.073741824	-2.847078495393154	7.0000	o,		
2.8589934591999997	-3.2621302370516183			0	
		0.0000			70-0
		-7.5000 7		4.62144	3.3421772799999996

begin1

-1.44000000000000013

2.000512

2.00002048

2.0000008192

2.0	32 -3.897599999999999	
2.0		
	64 -3.9959040000000003	
2.01	-3.99983616	-3.775
2.002	-3.9999934464000004	
2.000	-3.99999737856	
2.00010	-3.9999998951424	-3.85
2.000020	48 -3.999999995805693	
2.0000040	96 -3.99999999983223	Q
2.00000081	92 -3.999999999993285	-3.925

จากรูปทั้ง 4 รูป จะเห็นได้ว่ายิ่งมีค่า alpha มากเท่าไรจะยิ่งทำให้เจอค่า loss ที่น้อยที่สุดได้ เร็วกว่าที่มีค่า alpha น้อย เนื่องจาก sigmoid function คือ w = w - (alpha * gradient(w)) จากการนำค่า alpha คูณกับค่าความชั้นของ w

2.32

2.0128

จะเห็นได้ว่าถ้า alpha = 0.4 เมื่อคูณกับค่าความชั้นของ w แล้ว จะมีค่าสูงมากใน loop ต้นๆ และเมื่อนำค่า w มาลบกับค่า alpha * gradient(w) จะทำให้ค่า w ถัดไปมีค่าน้อยลง แต่ถ้าใช้ alpha = 0.1 ก็ได้เหมือนกัน เพียงแต่ต้องเพิ่มจำนวน loop ให้มากๆ