

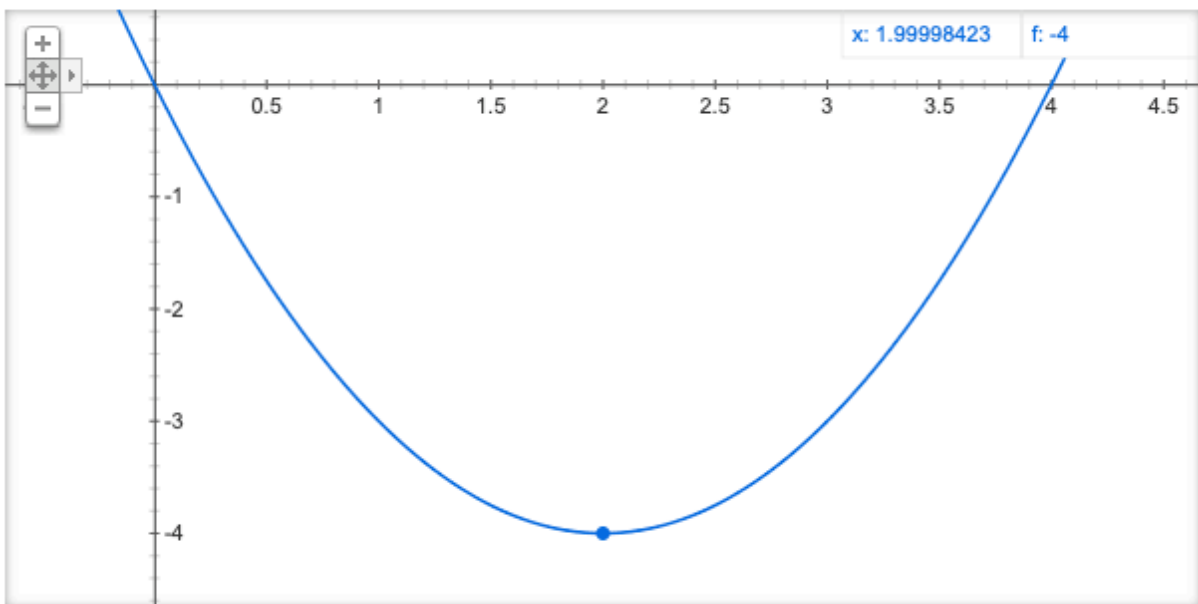
lab

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 เนื่องจากทำให้เกิดค่า loss น้อยที่สุด

Graph for $x^2 - 4x$



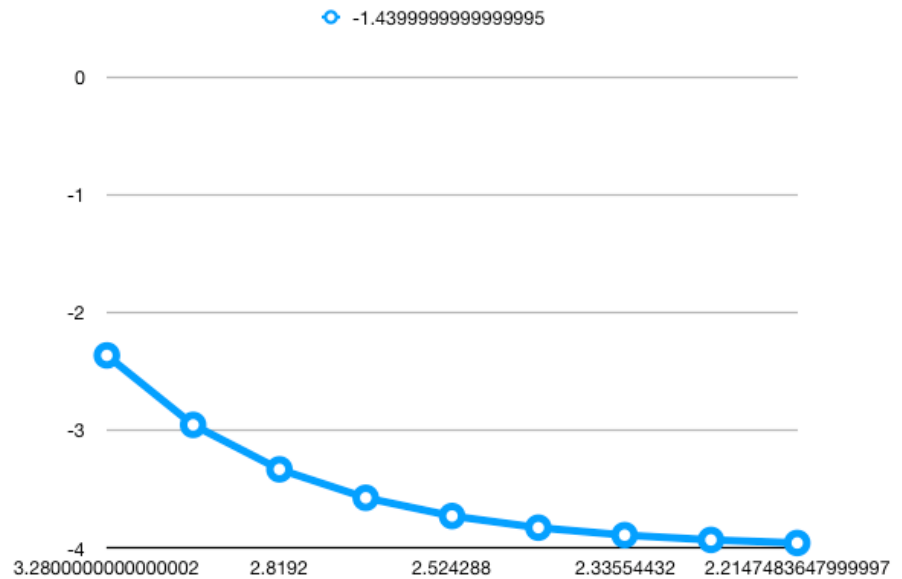
[More info](#)

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

รูปที่ 1 $w = 4$ $\text{loop} = 10$ $\alpha = 0.1$

begin1

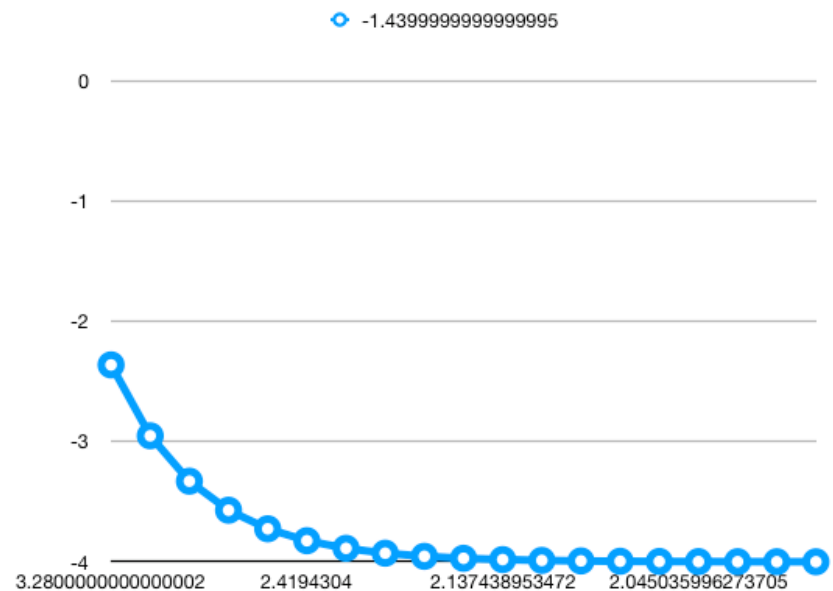
3.6	-1.4399999999999995
3.2800000000000002	-2.3615999999999993
3.024	-2.9514239999999994
2.8192	-3.32891136
2.65536	-3.5705032703999997
2.524288	-3.725122093056
2.4194304	-3.82407813955584
2.33554432	-3.8874100093157375
2.2684354559999997	-3.927942405962072
2.2147483647999997	-3.9538831398157264



รูปที่2 $w = 4$ $\text{loop} = 20$ $\alpha = 0.1$

begin1

3.6	-1.4399999999999995
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



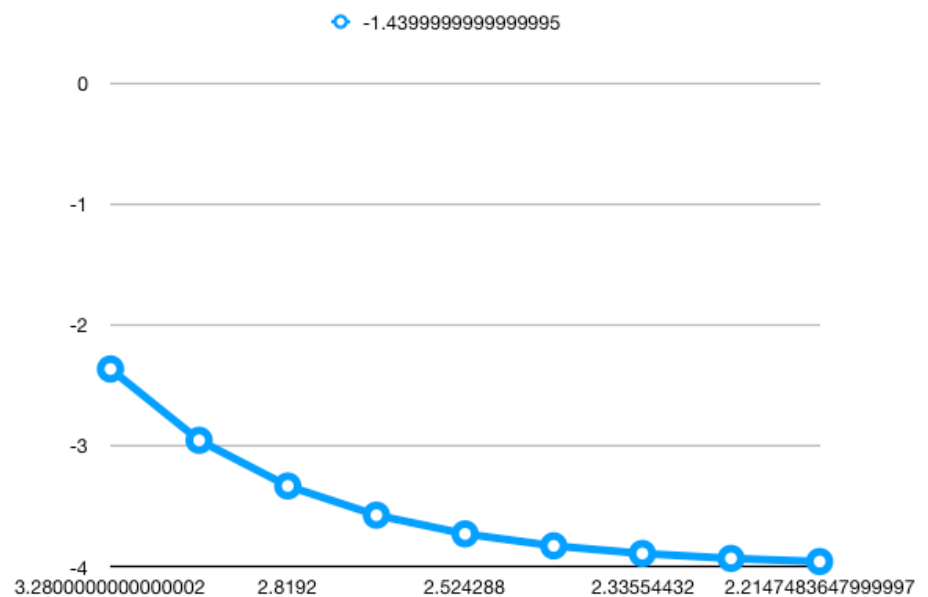
จากรูปที่ 1 มี $\text{loop} = 10$ จะมีค่า w คือ 2.2147483647999997 และค่า loss คือ -3.9538831398157264 ส่วนรูปที่ 2 มี $\text{loop} = 20$ จะมีค่า w คือ 2.023058430092137 และค่า loss คือ -3.999468308801686

ดังนั้น ถ้ามีจำนวน loop มากจะทำพบบให้ค่า w ที่ทำให้เกินค่า loss น้อยที่สุด มากกว่า ที่มีจำนวน loop น้อย

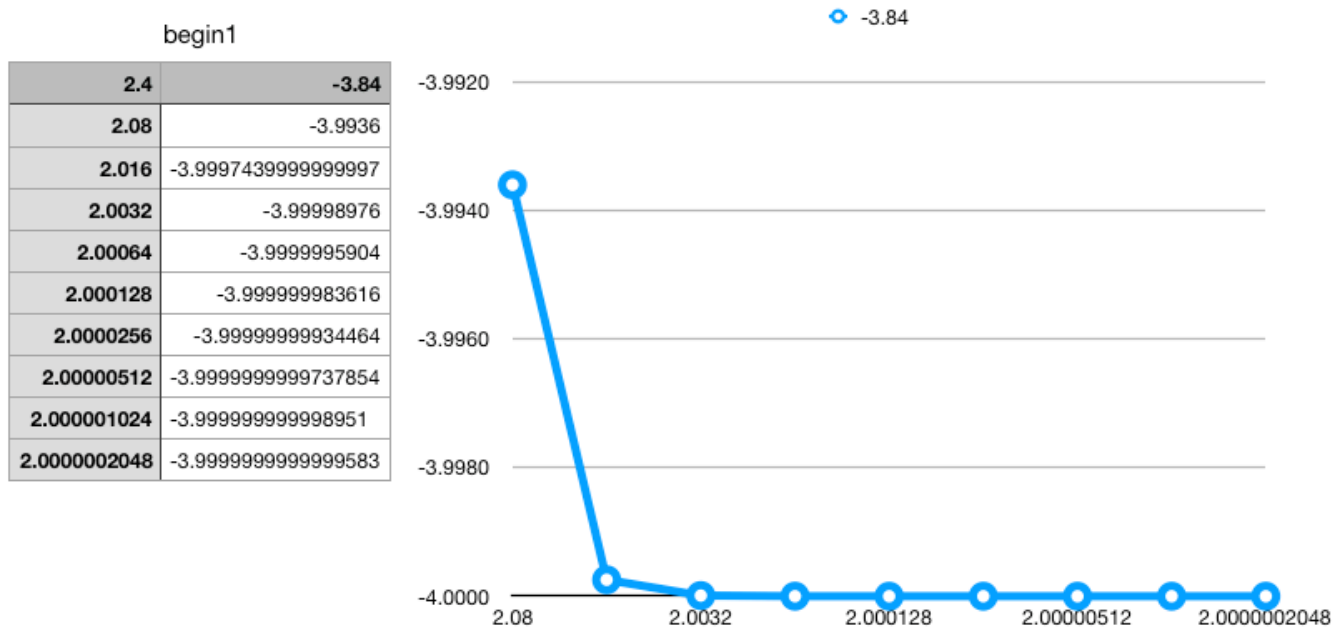
3.alpha มาก/น้อย มีผลกับ w อย่างไร และเพราะอะไร

รูปที่ 1 $w = 4$ $\text{loop} = 10$ $\alpha = 0.1$

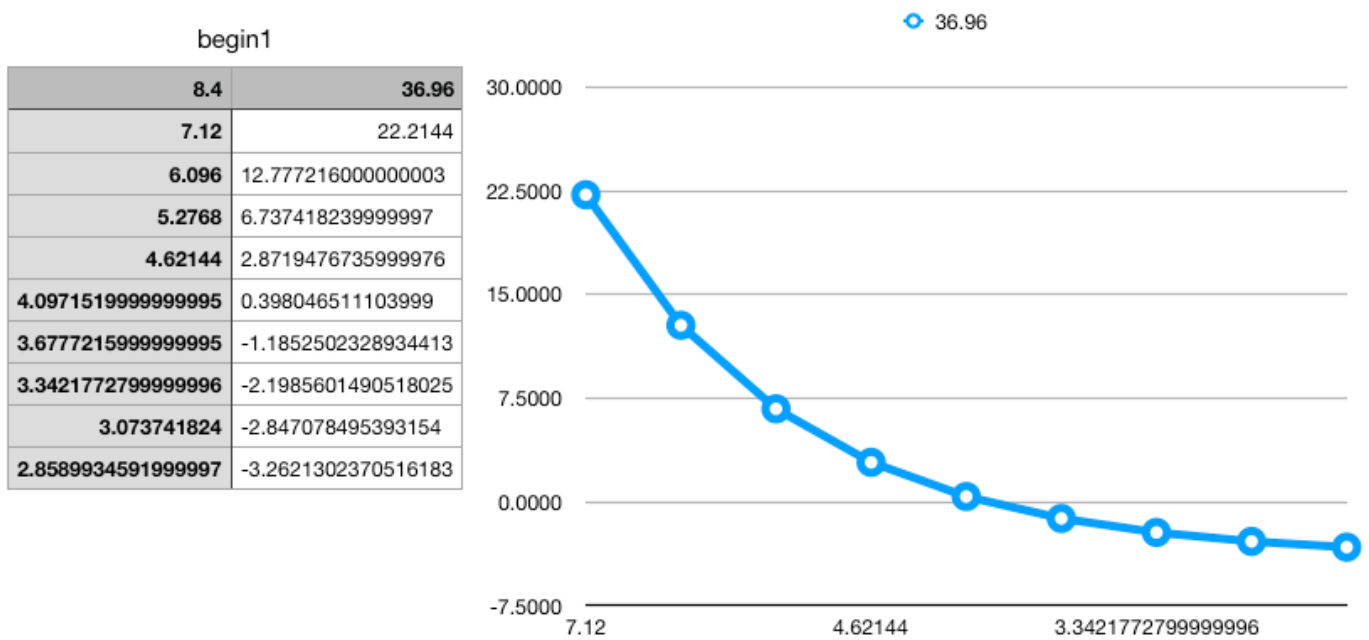
begin1	
3.6	-1.4399999999999995
3.2800000000000002	-2.3615999999999993
3.024	-2.9514239999999994
2.8192	-3.32891136
2.65536	-3.5705032703999997
2.524288	-3.725122093056
2.4194304	-3.82407813955584
2.33554432	-3.8874100093157375
2.2684354559999997	-3.927942405962072
2.2147483647999997	-3.9538831398157264



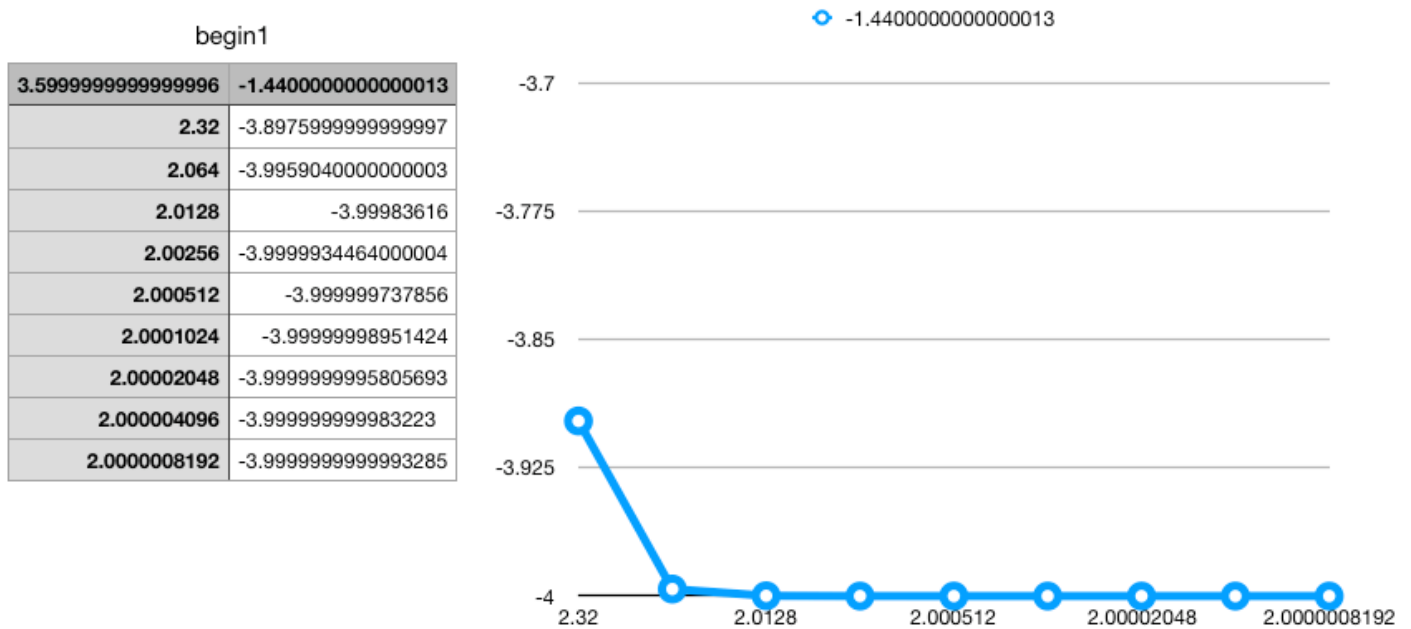
รูปที่ 2 $w = 4$ $\text{loop} = 10$ $\alpha = 0.4$



รูปที่ 3 $w = 10$ $\text{loop} = 10$ $\alpha = 0.1$



รูปที่ 4 $w = 10$ $\text{loop} = 10$ $\alpha = 0.4$



จากรูปทั้ง 4 รูป จะเห็นได้ว่ายิ่งมีค่า α มากเท่าไรจะยิ่งทำให้เจอค่า loss ที่น้อยที่สุดได้เร็วกว่าที่มีค่า α น้อย เนื่องจาก sigmoid function คือ $w = w - (\alpha * \text{gradient}(w))$ จากการนำค่า α คูณกับค่าความชันของ w

สมมติให้ความชันคือ $f(w) = 2*w - 4.0$

ถ้า $\alpha = 0.4$ $w = 10$ $\text{loop} = 3$

$$\text{loop} = 1 \rightarrow w = 10 - (0.4 * (2*10 - 4.0)) = 10 - 6.4 = 3.6$$

$$\text{loop} = 2 \rightarrow w = 3.6 - (0.4 * (2*3.6 - 4.0)) = 3.6 - 1.28 = 2.32$$

$$\text{loop} = 3 \rightarrow w = 2.32 - (0.4 * (2*2.32 - 4.0)) = 2.32 - 0.256 = 2.064$$

ถ้า $\alpha = 0.1$ $w = 10$ $\text{loop} = 3$

$$\text{loop} = 1 \rightarrow w = 10 - (0.1 * (2*10 - 4.0)) = 10 - 1.6 = 8.4$$

$$\text{loop} = 2 \rightarrow w = 8.4 - (0.1 * (2*8.4 - 4.0)) = 8.4 - 1.28 = 7.12$$

$$\text{loop} = 3 \rightarrow w = 7.12 - (0.1 * (2*7.12 - 4.0)) = 7.12 - 1.024 = 6.096$$

จะเห็นได้ว่าถ้า $\alpha = 0.4$ เมื่อคูณกับค่าความชันของ w แล้ว จะมีค่าสูงมากใน loop ต้นๆ และเมื่อนำค่า w มาลบกับค่า $\alpha * \text{gradient}(w)$ จะทำให้ค่า w ถัดไปมีค่าน้อยลง แต่ถ้าใช้ $\alpha = 0.1$ ก็ได้เหมือนกัน เพียงแต่ต้องเพิ่มจำนวน loop ให้มากๆ