

HW1. XOR / Back propagation

조주현

박형민 교수님

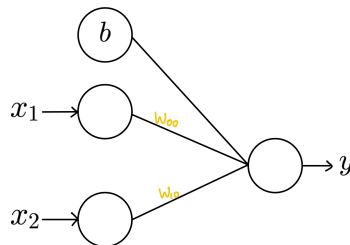
2021년 10월 15일

HW1의 목표는 아래와 같은 table을 가지는 XOR gate를 back propagation으로 학습하는 neural network를 직접 구현하고, 1-layer model과 2-layer 모델의 결과를 비교하여 다른 점을 확인하는 것이다.

[XOR gate 표]

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	0

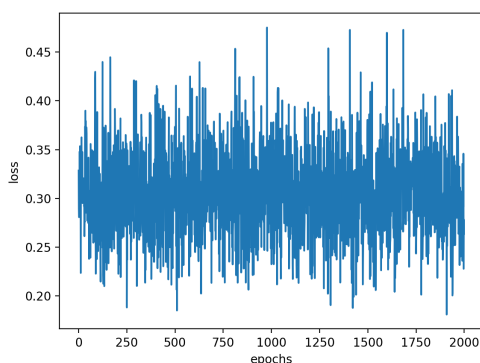
1. 1-layer model



해당 model에서 $y = \sigma(\text{net}_j)$, $\text{net}_j = \sum_{i=0}^1 x_i w_{ji} + b$ 이다. Weight (W)와 bias (b)의 초기값은 랜덤값으로 저장하

였고, activation function은 sigmoid를, training error는 binary cross entropy 함수를 이용하였다. Back propagation은 error값을 줄이는 방향으로 W를 업데이트를 하여 네트워크를 학습한다.

Epoch = 20000, learning rate = 0.1 로 학습을 시작했다. 원하는 예측 결과는 위 Table 1. Y값에 가깝게 나와야 하지만, 실제 결과를 확인해보면 모두 0.4와 0.5 사이 값으로 예측된 것을 확인할 수 있다. 이는 training 결과가 어느 한 값으로 수렴하지 않았다는 것을 의미한다. error값들을 이용해 그래프를 그려보아도 error가 0으로 수렴하지 않는 것을 확인할 수 있다.

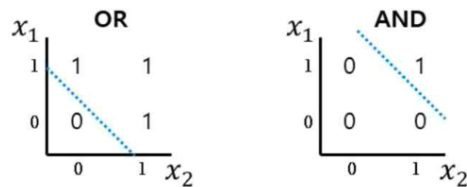


Layer1 error graph

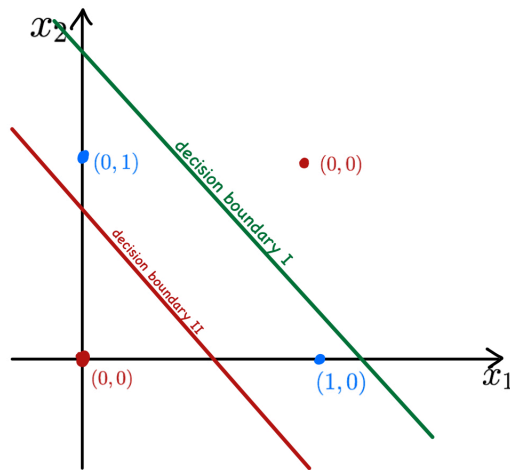
```
input:  1 1 , answer:  0 , pred: 0.4556
input:  1 0 , answer:  1 , pred: 0.4458
input:  0 1 , answer:  1 , pred: 0.4957
input:  0 0 , answer:  0 , pred: 0.4858
```

Layer1 XOR prediction

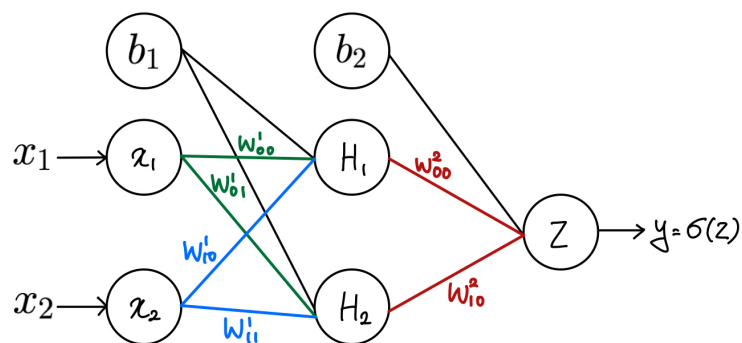
위와 같은 결과가 나온 이유는 다음과 같다. 1 layer model은 결국 한번의 선형 구분으로 4개의 점을 분류하는 것이다. Network에서 y값이 나오는 식을 생각해 보면 $w_{00}x_1 + w_{10}x_2 + b$ 가 나오는데, 이는 결국 x_1 축과 x_2 축에서의 직선 함수와 같다. 아래 그림처럼 OR gate나 AND gate(OR: 값이 서로 다르면 1인 gate, AND : 값이 모두 1로 같으면 1인 gate)의 경우는 하나의 선형 구분으로 분류가 가능하다.



하지만 XOR gate 같은 경우는 하나의 구분선으로 0과 1을 나눌 수 없다. 따라서 2개의 layer가 필요하다.



2. 2-layer model

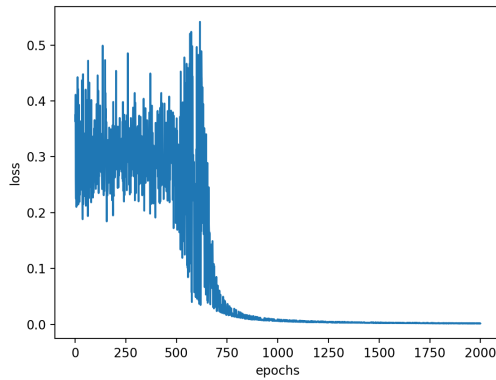


이 모델의 경우 input과 output 사이에 hidden layer가 존재하고, hidden layer의 activation function은 sigmoid이다. 따라서 비선형인 분류도 가능할 것이라 예측할 수 있다. 이 model에서의 변수는 아래와 같이 표현된다.

$$h_j = \sigma(\text{net}_j), \quad \text{net}_j = \sum_{i=0}^1 w_{ji}^1 x_i + b_1$$

$$y = \sigma(\text{net}_k), \quad \text{net}_k = \sum_{j=0}^1 w_{kj}^2 h_j + b_2$$

Activation function과 cost function은 1-layer model과 동일하며, learning rate와 epoch도 동일하게 설정하였다. 다만 이 때 training마다 loss값이 일정하게 수렴하지 않는 문제가 생겨 weight와 bias의 초기값이 커서 생기는 문제라 판단하여 초기값을 10으로 나누어 주었다. 이후 layer-1 model도 동일하게 10으로 나눠 주었다(물론 난수로 설정하긴 했지만, 둘의 조건의 동일해야 비교하는데 의의가 있다고 판단했다). Training 결과는 아래와 같다.



Layer 2 error graph

```
input:  1 1 , answer:  0 , pred: 0.0034
input:  1 0 , answer:  1 , pred: 0.9964
input:  0 1 , answer:  1 , pred: 0.9953
input:  0 0 , answer:  0 , pred: 0.0039
```

Layer 2 XOR prediction

학습 이후, 두개의 모델 weight 값을 txt file로 저장하여 본 과제의 파이썬 코드와 함께 첨부하였다.