

딥러닝 세미나 2주차
김동협

4강 : 신경망 학습

DSP Lab, Inha University

August , 2021



목차

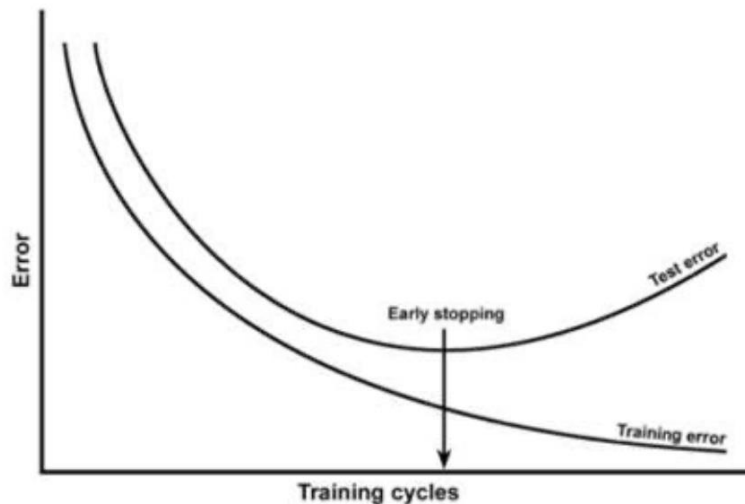
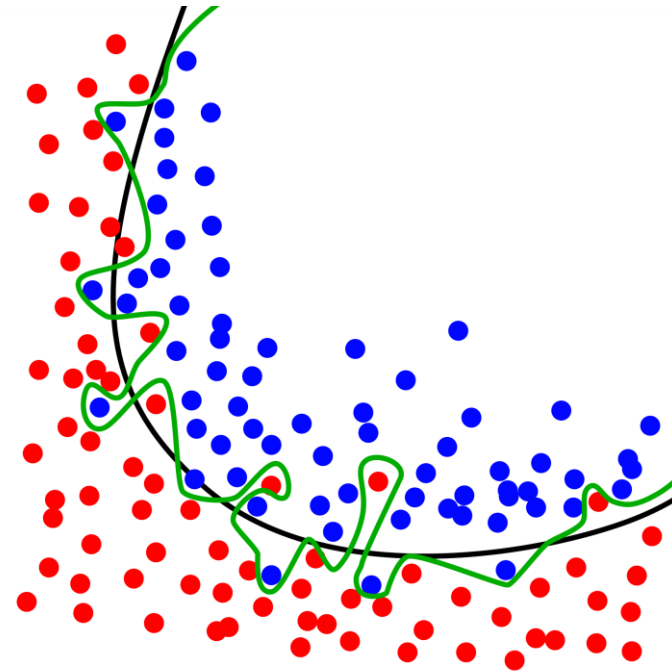
- 오버피팅
- 손실함수
- 오차제곱합, 교차 엔트로피 오차
- 왜 손실함수를 설정하는가
- 미분과 기울기
- 경사하강법



-오버피팅

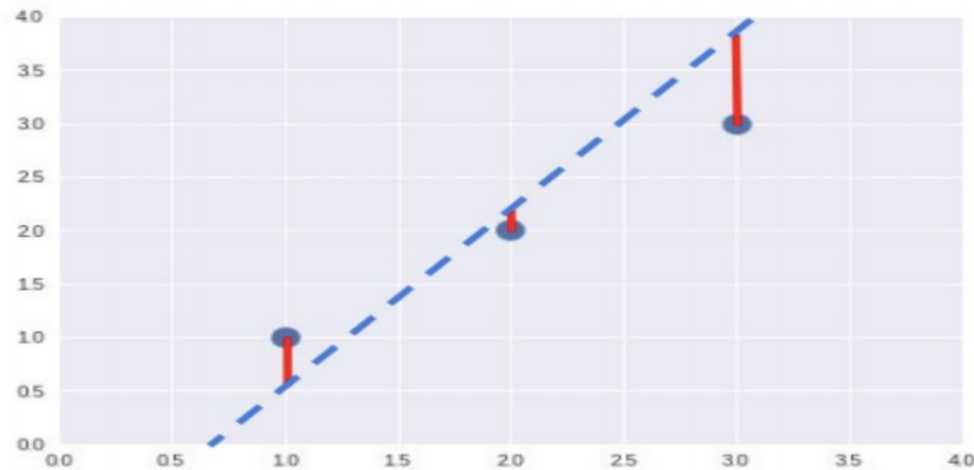
Overfitting은 학습 데이터(Training Set)에 대해 **과하게** 학습된 상황입니다. 따라서 학습 데이터 이외의 데이터에 대해선 모델이 잘 동작하지 못합니다. 학습 데이터가 부족하거나, 데이터의 특성에 비해 모델이 너무 복잡한 경우 발생합니다. Training Set에 대한 loss는 계속 떨어지는데, Test Set에 대한 loss는 감소하다가 다시 증가합니다.

1. 데이터 양 늘리기
2. 모델의 복잡도 줄이기
3. 가중치 규제(regularization) -> ??
4. dropout



- 손실 함수

- 정의 #
- 데이터를 토대로 산출한 모델의 예측 값과 실제 값과의 차이를 표현하는 지표
- 모델 성능의 '나쁨'을 나타내는 지표로 "현재의 모델이 데이터를 얼마나 잘 처리하지 못하느냐" 를 나타내는 지표



-오차제곱합, 교차 엔트로피 오차

- 오차제곱합

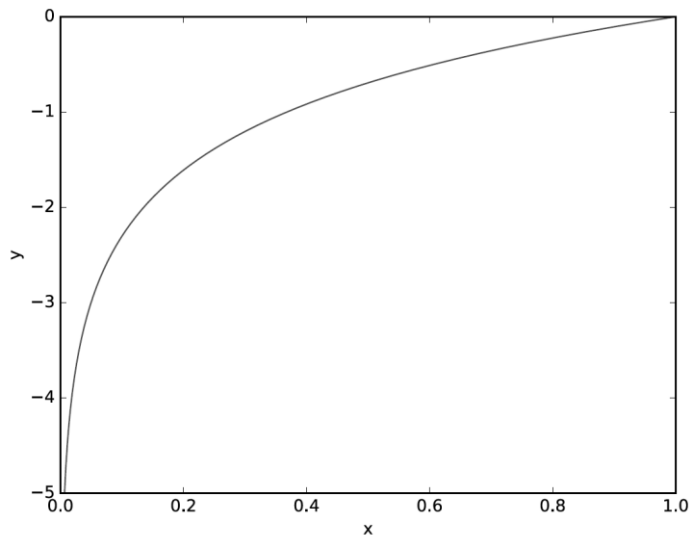
$$E = \frac{1}{2} \sum_k (y_k - t_k)^2$$

1. 제곱을 하는 이유 : 양수로 통일하기 위함
2. 1/2은 고정값은 아님

- 교차 엔트로피 오차

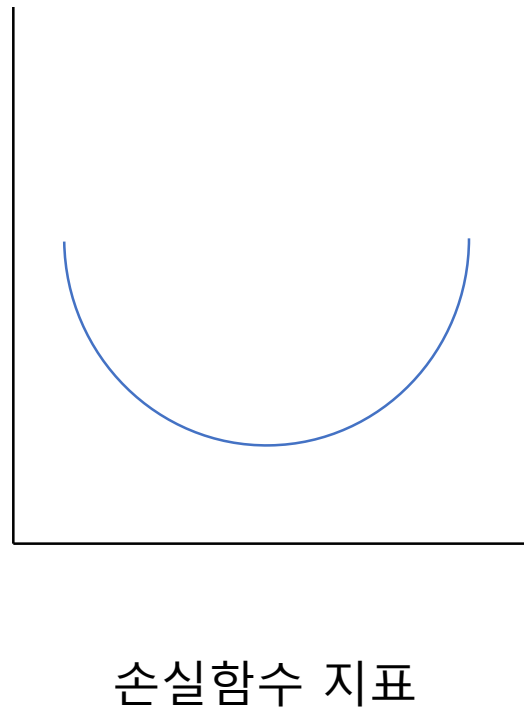
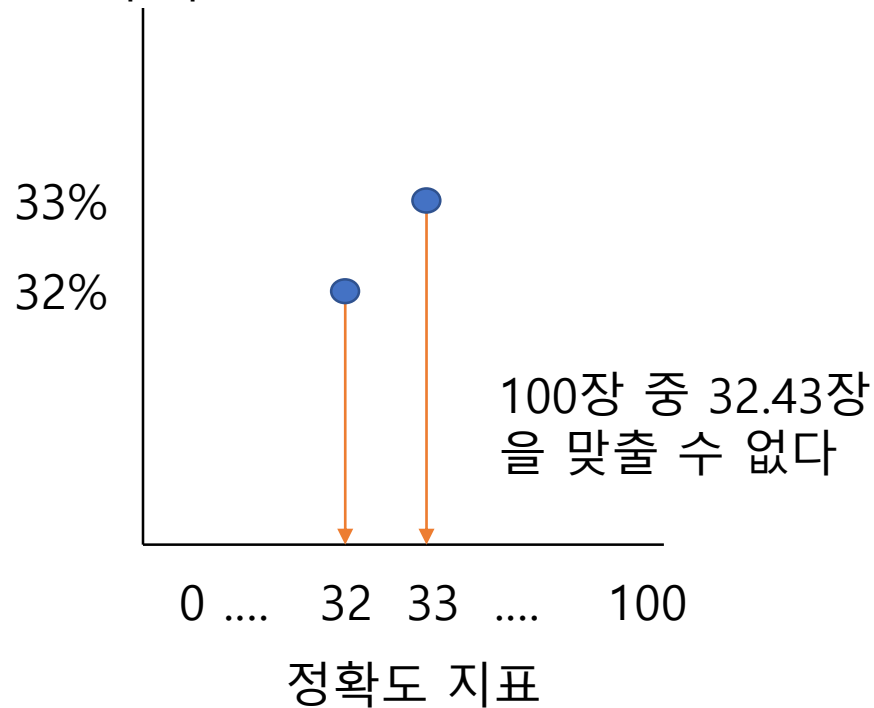
$$E = -\sum_k t_k \log y_k$$

신경망 출력(확률) 0~1



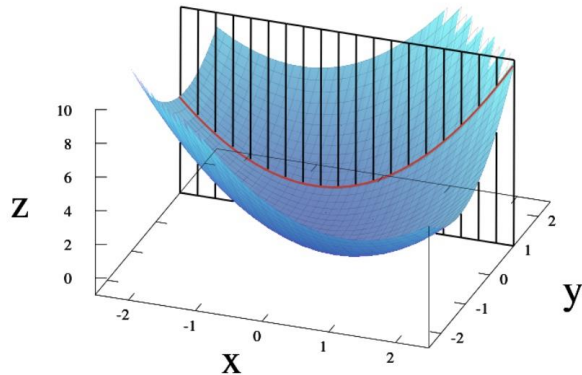
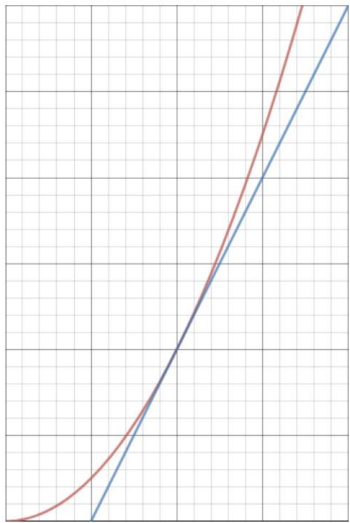
-왜 손실함수를 설정하는가

- 함수의 연속성과 관계가 있음
- 즉, 매개변수 변화량에 따른 지표의 변화량을 (ex. dL/dW) 구해야하는데 정확도는 변화량을 보기에 부적합



-미분과 기울기

- 전미분과 편미분의 기하학적 의미



$$y = f(x_1, x_2) = x_1^3 + 7x_1^2 + 8x_2^3 - 2x_2 + 7 \quad \dots (1)$$

$$y = (6x_1x_2 - x_2^3)dx_1 + (3x_1^2 - 3x_1x_2^2)dx_2$$

$$f_1 = 6x_1x_2 - x_2^3$$

$$f_2 = 3x_1^2 - 3x_1x_2^2$$

편미분 -> 선택한 축 이외의 변화량은 무시

전미분 -> 모든 축에 대한 변화량을 포함

```
grads['W1'] = numerical_gradient(loss_W, self.params['W1'])
```

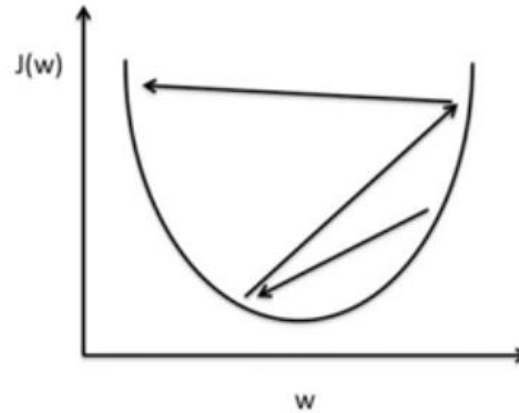


-경사하강법

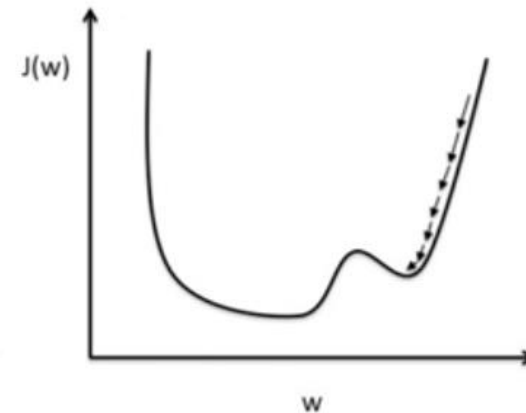
$$x_0 = x_0 - \eta \frac{\partial f}{\partial x_0}$$

$$x_1 = x_1 - \eta \frac{\partial f}{\partial x_1}$$

Eta는 학습률을 의미.
한 번 학습 시 학습되는 양



Large learning rate: Overshooting.



Small learning rate: Many iterations until convergence and trapping in local minima.

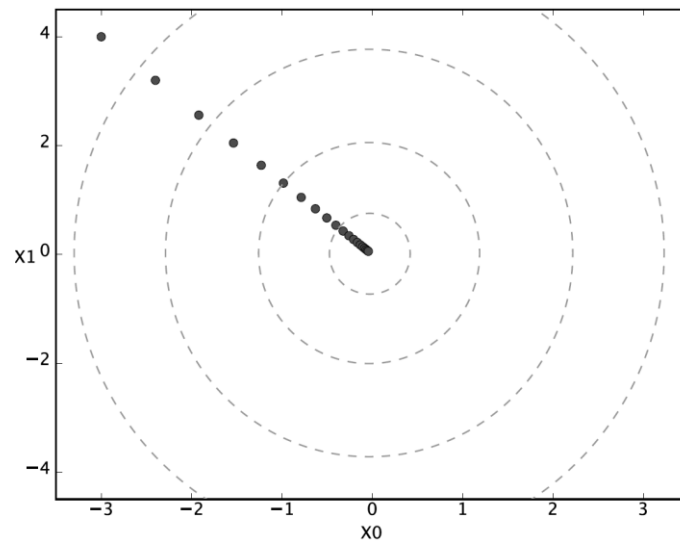
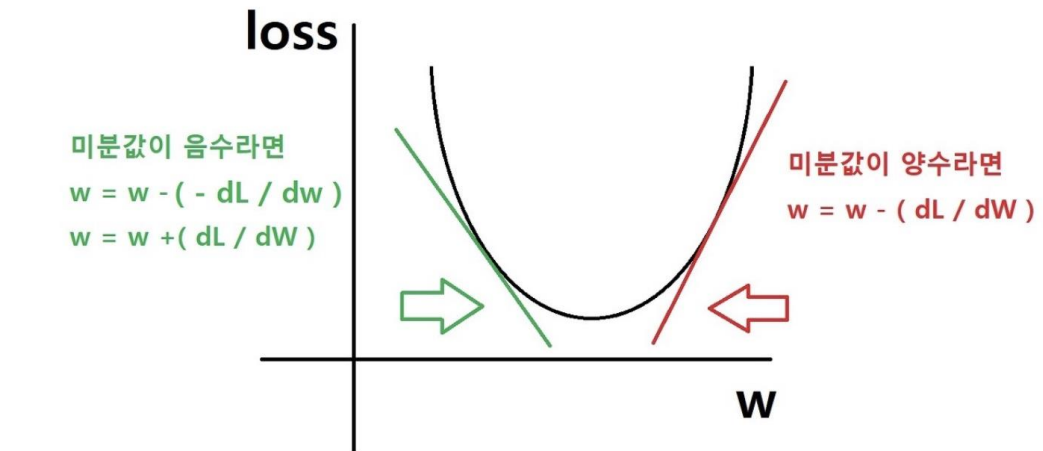
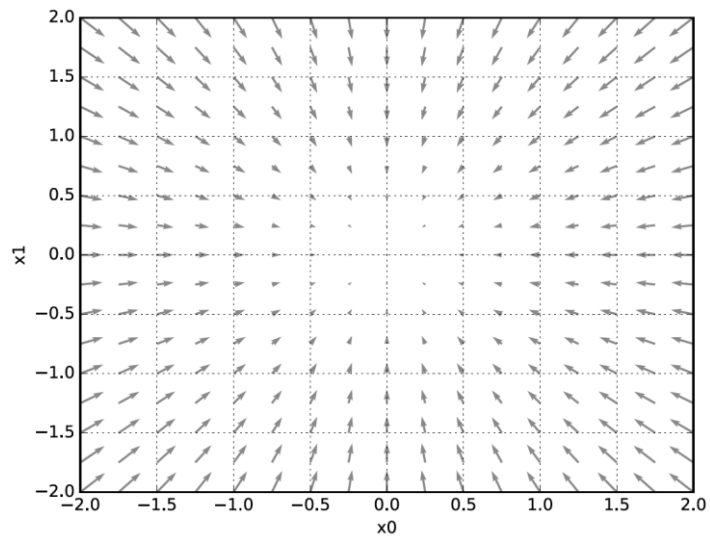


-경사하강법

$$f(x_0, x_1) = x_0^2 + x_1^2$$

$$x_0 = x_0 - \eta \frac{\partial f}{\partial x_0}$$

$$x_1 = x_1 - \eta \frac{\partial f}{\partial x_1}$$



Thanks!

