

딥러닝 세미나 2주차
김동협

CEE&MSE
DSP Lab, Inha University
August , 2021



-오차제곱합, 교차 엔트로피 오차

- 오차제곱합

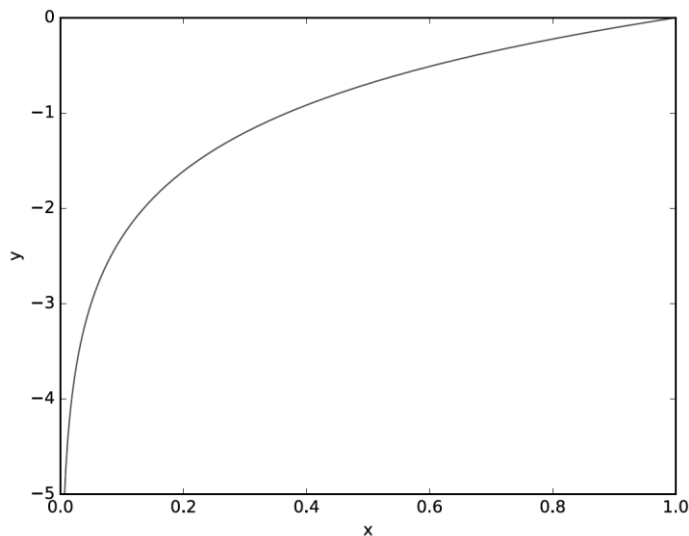
$$E = \frac{1}{2} \sum_k (y_k - t_k)^2$$

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



교차 엔트로피 오차

- T = 정답 레이블 (one hot vector)
- Y = 신경망 출력 (0~1)

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

- 식의 의미 : 정답일 때의 확률에만 로그를 취한 값

식만 봤을 땐 모든 정답에서 오차를 구해서 더하는 MSE와의 차이가 이 부분에서 있음.



교차 엔트로피 오차

정보 엔트로피 : 확률이 낮을수록, 어떤 정보일지는 불확실하게 되고, 우리는 이때 '정보가 많다', '엔트로피가 높다'고 표현한다.

Ex) 이영호(0.99) vs 김동협(0.01) < 이영호(0.5) vs 김택용(0.5)

즉 교차 엔트로피는 한 정답 레이블에 대한 불확실성을 표현한다.



교차 엔트로피 오차

- 교차 엔트로피는 확률이 input. 즉 정답의 오차에 대한 조정이 이루어지면 다른 다른 Label에 대한 확률에도 변화가 생긴다.
- 연속적인 데이터에 대해 각각 모두 Labeling을 하기에 제한된다.
- 각각의 출력값이 정답과 얼마나 차이가 나는지는 반영하지 못함



오차제곱합 vs 교차 엔트로피

MSE 는 correct answer과 incorrect answer에 대해 동등하게 고려하기 때문에 정답의 불확실성을 표현하려는 과정에 방해가 된다.

output				label	A	B	C	correct?
	0.3	0.3	0.4		0	0	1	Y
	0.3	0.4	0.3		0	1	0	Y
	0.1	0.2	0.7		1	0	0	N

첫 번째 sample: $-(\ln(0.3) * 0) + (\ln(0.3) * 0) + (\ln(0.4) * 1) = -\ln(0.4)$

3개 sample 모두에 대한 계산 및 ACE (Average cross-entropy): $-(\ln(0.4) + \ln(0.4) + \ln(0.1))/3 = 1.38$

첫 번째 sample : $(0.3 - 0)^2 + (0.3 - 0)^2 + (0.4 - 1)^2 = 0.09 + 0.09 + 0.36 = 0.54$

3개 sample 모두에 대한 계산 및 MSE(Mean squared error): $(0.54 + 0.54 + 1.34)/3 = 0.81$



Thanks!

