

5강. 신경망(2)

◆ 담당교수 : 장필훈

■ 정리하기

1. 앞선 회귀문제에서는 음의 로그최대화를 통해 매개변수를 결정했다. 뉴럴넷은 보통 오류 함수를 최소화한다. 둘의 본질적인 차이는 없다.
2. 오류함수와 활성화 함수 사이에는 짝이 있다. 제곱합 오류함수의 경우에는 항등함수를 쓴다.
3. 이진분류문제에서 출력을 시그모이드 함수로 가정하면 결과값을 확률로 바로 해석할 수 있다.
4. 이진분류문제에서 조건부 분포에 음의 로그를 취한 오류함수는 교차 엔트로피 오류함수이다.
5. 교차 엔트로피오류함수는 제곱오류함수보다 훈련과정이 더 빠르고, 일반화도 개선된다고 알려져 있다.
6. 비선형함수의 경우 해석적으로 미분값이 0이 되는 지점을 찾기 어렵기 때문에, 보통 경사하강법을 이용한다. $w^{(\tau+1)} = w^{(\tau)} + \eta \nabla E(w^{(\tau)})$
7. 배치방식으로 ∇E 를 계산하려면 데이터 집합 전체를 이용해야 하는 것이 원칙이다.
8. 전체집합을 이용하기 힘든 경우가 대부분이므로 stochastic gradient descent을 이용한다.
9. 오차역전파는 오류함수의 미분을 계산하고, 결과를 바탕으로 가중치 w 를 조절한다.
10. 다층으로 이루어져 있을 때는, 출력단부터 차례로 계산해서 결국 모든 유닛의 w 를 얻어낸다.
11. 뉴럴넷에서 입력과 출력의 차원은 데이터의 형태에 의해 정해지고, 중간 은닉유닛은 자유롭게 디자인할 수 있다.
12. 뉴럴넷도 과적합을 막기 위해 정규화 한다. 조기종료, 네트워크 구조 단순화등을 한다.
13. 구조를 단순화 하면서 효과적으로 과적합을 제어하고, 불변성까지 획득하는 대표적인 구조가 CNN이다.
14. 다루려고 하는 조건부 분포가 가우시안이 아닐 때는 혼합밀도 네트워크(가우시언이나 다른 분포들의 혼합)을 사용하여 근사한다.