

4강. 선형분류(2), 신경망(1)

◆ 담당교수 : 장필훈

■ 주요용어

용어	해설
퍼셉트론	$f(\mathbf{w}^T \phi(\mathbf{x}))$ 가 1 또는 -1의 계단함수로 정의되는 유닛. ϕ 는 비선형 변환으로 activation function이라고도 불린다.
경사하강법	함수의 계산이 해석적으로 매우 어렵거나 불가능할 때, 수치적인 방법으로 근사하는 최적화방법이 쓰이는데 그 중 하나가 경사하강법. 가장 흔하게 쓰이고, 퍼셉트론이나 피드포워드네트워크 모두 이 방법을 사용한다. 식으로 나타내면, $\mathbf{w}^{(\tau+1)} = \mathbf{w}^{(\tau)} - \eta \nabla E_p(\mathbf{w})$
로지스틱 시그모이드	시그모이드 함수와 같은 말. $\frac{1}{1+e^{-x}}$
피드포워드 네트워크 함수	뉴럴넷의 기본 유닛. 층 하나는 (모든 element들의 activation을 제외하면) 행렬의 곱 하나로 단순하게 표현될 수 있다. 자세한 것은 ppt참고. 이 함수를 다층으로 쌓으면 모든 공간에서 임의의 함수를 근사할 수 있다.

■ 정리하기

1. 퍼셉트론에서 하나의 unit은 비선형변환함수를 통과한 입력벡터의 출력이 특정값에서 step function을 이룬다. $y(\mathbf{x}) = f(\mathbf{w}^T \phi(\mathbf{x}))$
2. 퍼셉트론의 기댓값은 오분류된 패턴 전체에 대해서만 정의되므로 모든 공간에서 미분은 불가능하다.
3. 퍼셉트론기준의 계산은 경사하강법을 이용한다.
4. 오분류된 패턴의 오류함수에 대한 기여도는 점점 감소한다.
5. 훈련집합이 선형분리 가능하면 퍼셉트론 학습 알고리즘은 정확한 해를 유한한 단계로 구할 수 있다.
6. 선형분리 불가능하면 수렴하지 않는다.
7. 퍼셉트론알고리즘은 확률적 출력값을 내지 않고 다중클래스에 대해 일반화되지 않는다.

8. 확률적 생성모델의 경우 logistic sigmoid를 출력함수로 이용한다.
9. 로지스틱 회귀라고 불리지만 분류모델에 속한다.
10. 로지스틱 회귀는 수치적 방법으로 근사하여 계산한다.
11. 데이터의 양이 너무 크면 기저함수 자체를 매개변수 적응가능하게 만든다. 신경망이 그 대표적인 예.
12. 피드포워드 네트워크함수는 모든 공간에 대해 정의되므로 퍼셉트론과 다르다.
13. 피드포워드 네트워크함수는 차원감소역할을 하지만, 선형은 아니다.
14. 뉴럴넷은 universal approximator다.