



셋째마당

딥러닝의 시작, 신경망

9장 오차 역전파에서 딥러닝으로

- 1 딥러닝의 태동, 오차 역전파**
- 2 활성화 함수와 고급 경사 하강법**
- 3 속도와 정확도 문제를 해결하는 고급 경사 하강법**



오차 역전파에서 딥러닝으로

● 오차 역전파에서 딥러닝으로

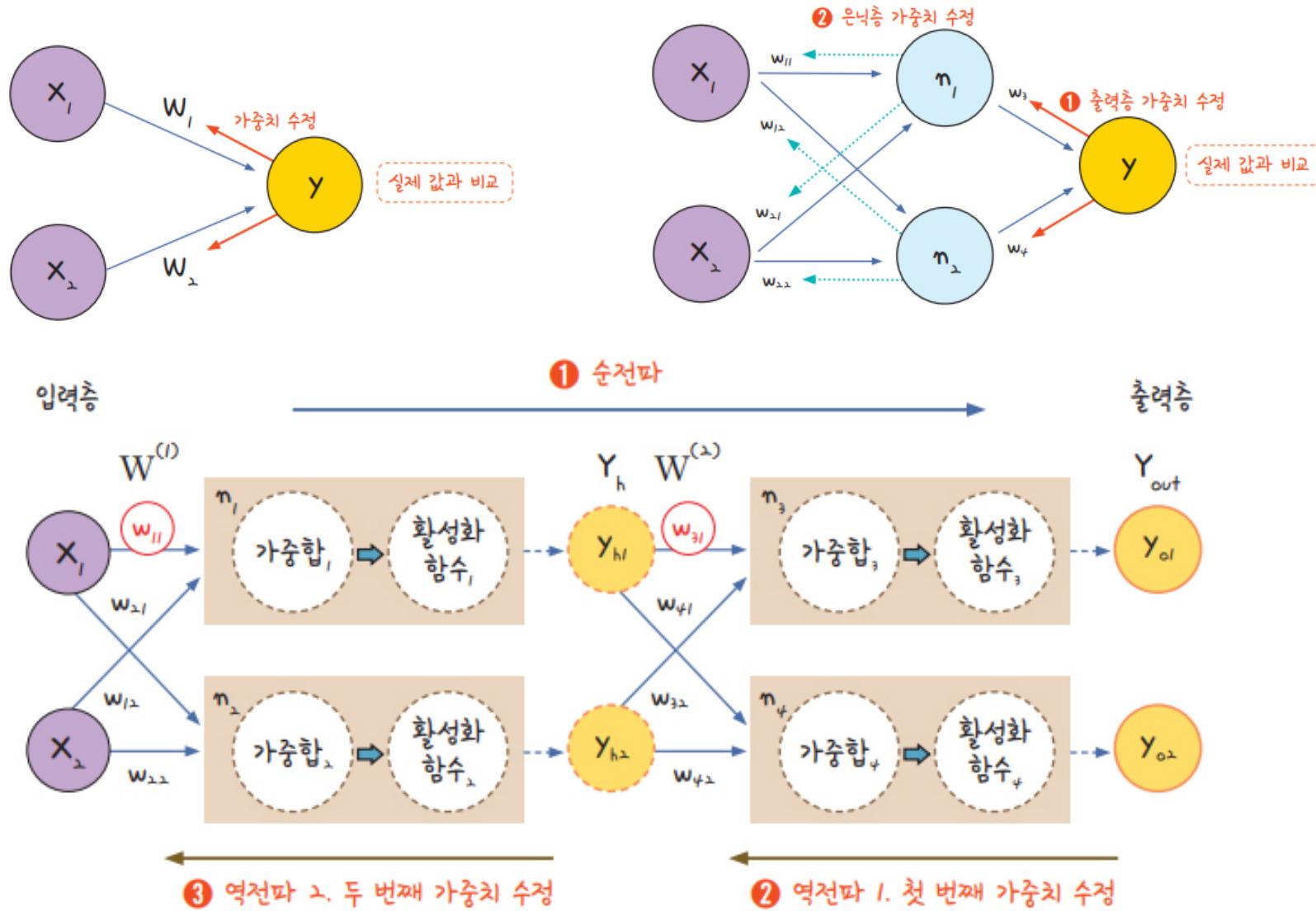
- 해결되지 않은 XOR 문제를 다층 퍼셉트론으로 해결
- 한 가지 문제를 만남
- 은닉층에 포함된 가중치를 업데이트할 방법이 없었던 것
- 이는 기나긴 인공지능의 겨울을 지나 오차 역전파라는 방법을 만나고 나서야 해결
- 오차 역전파는 후에 지금 우리가 아는 딥러닝의 탄생으로 이어짐
- 오차 역전파는 어떻게 해서 은닉층의 오차를 업데이트할 수 있었을까?



1 딥러닝의 태동, 오차 역전파

1 딥러닝의 태동, 오차 역전파

▼ 퍼셉트론에서 오차 수정





1 딥러닝의 태동, 오차 역전파

첫 번째 가중치 업데이트 공식 = $(y_{o1} - \underline{y_{\text{실제 값}}}) \cdot \underline{y_{o1}(1-y_{o1})} \cdot \underline{y_{h1}}$

두 번째 가중치 업데이트 공식 = $(\underline{\delta y_{o1} \cdot w_{31}} + \underline{\delta y_{o2} \cdot w_{41}}) \underline{y_{h1}(1-y_{h1})} \cdot \underline{x_1}$

- 은닉층이 많아지면 연산과정이 복잡해 보이지만 편미분한 결과 값은 다음과 같다.
- 공식 중 밑줄 친 부분을 잘 보면 두 식 모두 'out(1-out)' 형태를 취하고 있다는 것을 알 수 있음
- 이 식을 델타식이라고 함. 층이 많이져도 델타식을 사용하면 은닉층의 오차를 수정할 수 있는 방법이된다.

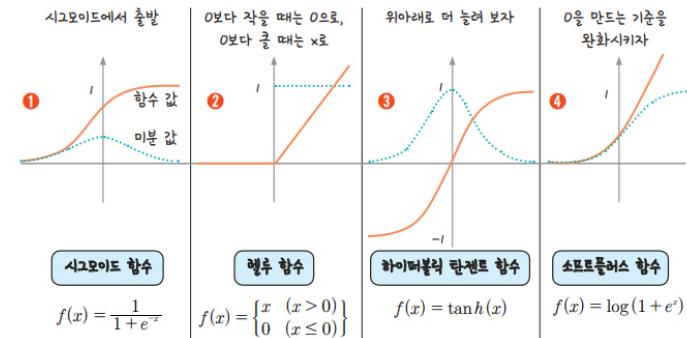
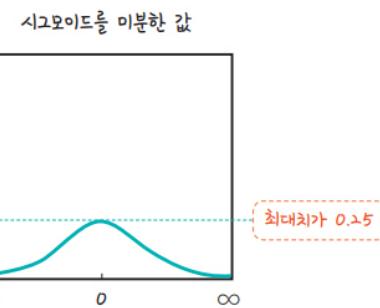
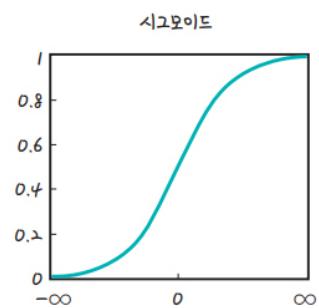
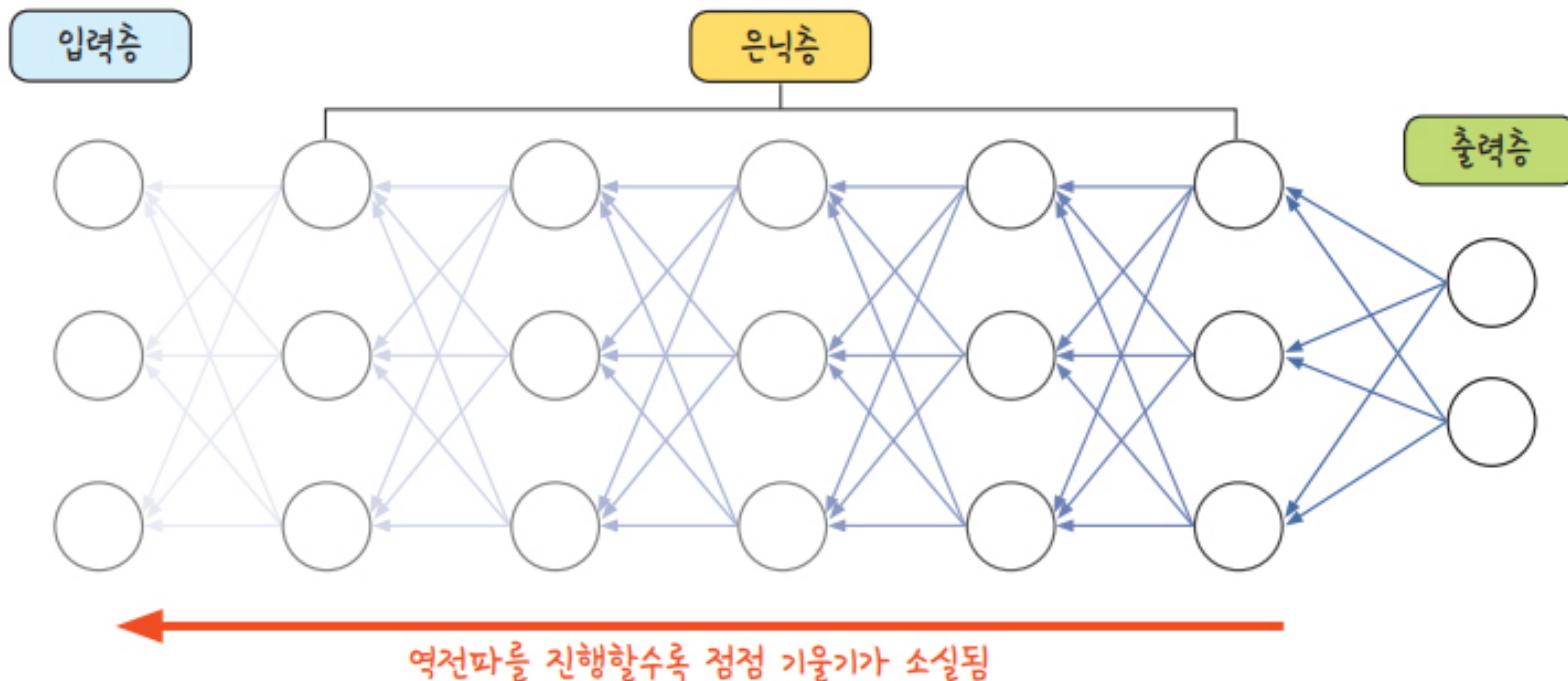


2 활성화 함수와 고급 경사 하강법



2 활성화 함수와 고급 경사 하강법

● 활성화 함수와 고급 경사 하강법



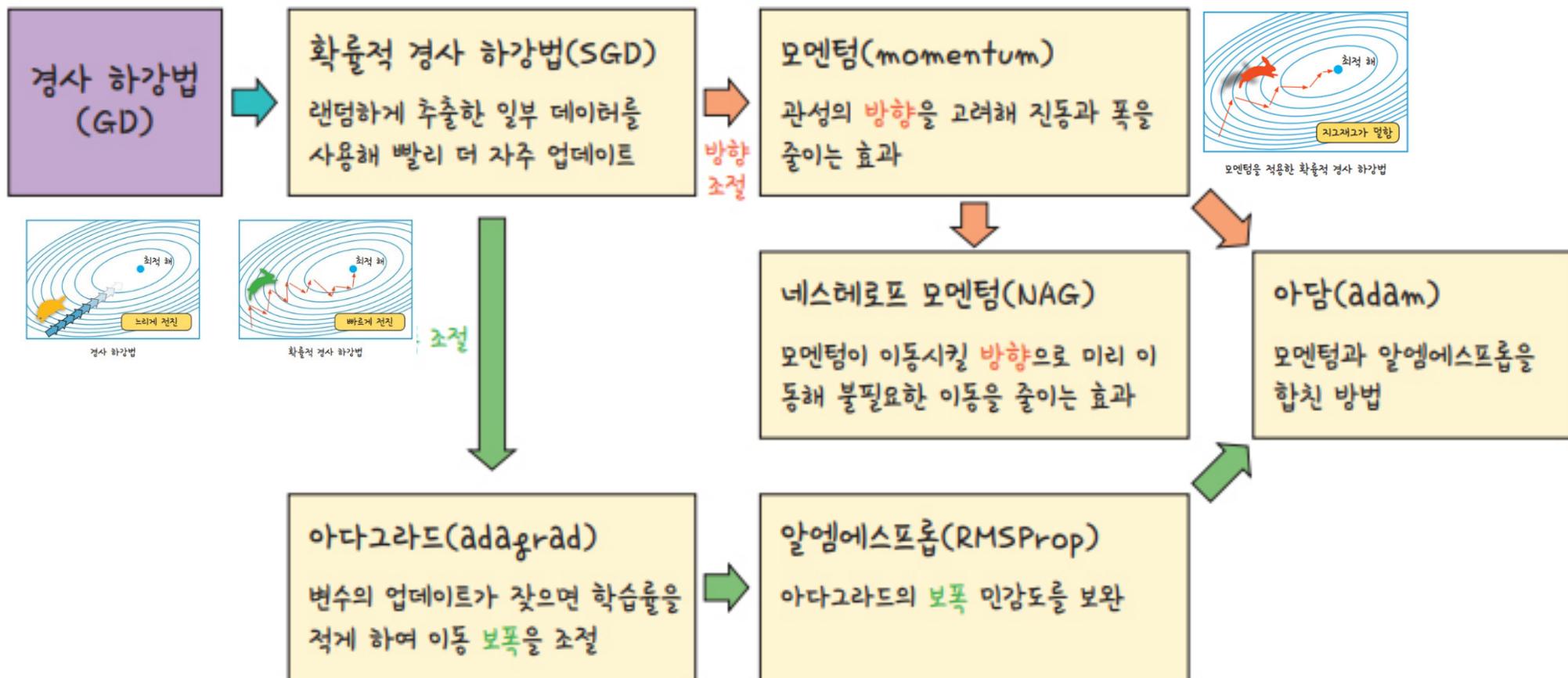


3 속도와 정확도 문제를 해결하는 고급 경사 하강법



3 속도와 정확도 문제를 해결하는 고급 경사 하강법

▼ 그림 9-9 | 딥러닝에 사용되는 고급 경사 하강법의 변천





3 속도와 정확도 문제를 해결하는 고급 경사 하강법

● 속도와 정확도 문제를 해결하는 고급 경사 하강법

- 이렇게 오차를 최소화하는 경사 하강법들을 딥러닝에서는 '옵티마이저'라고 한다고 했음
- 앞에 소개한 고급 경사 하강법들은 텐서플로에 포함되어 있는 optimizers라는 객체에 이름을 적어 주는 것만으로 손쉽게 실행할 수 있음
- 또 앞 절에서 소개된 시그모이드, 렐루 등 활성화 함수도 activation이라는 객체에 이름을 적어 주는 것으로 손쉽게 실행할 수 있음
- 이어지는 장에서 이러한 부분을 실습해 보자
- 이렇게 해서 4장부터 9장까지, 텐서플로에서 사용되는 대부분의 개념과 용어를 상세히 배웠습니다.
- 드디어 텐서플로를 이용한 모델링을 할 준비가 되었음