퍼셉트론 : 선형 분류기라고도 한다.

(선형회귀와는 다르다.)

선형회귀는 데이터를 가장 잘 표현하는 하나의 직선!!

선형 분류기는 데이터를 0과 1을 가장 잘 구분 지어줄 수 있는 직선

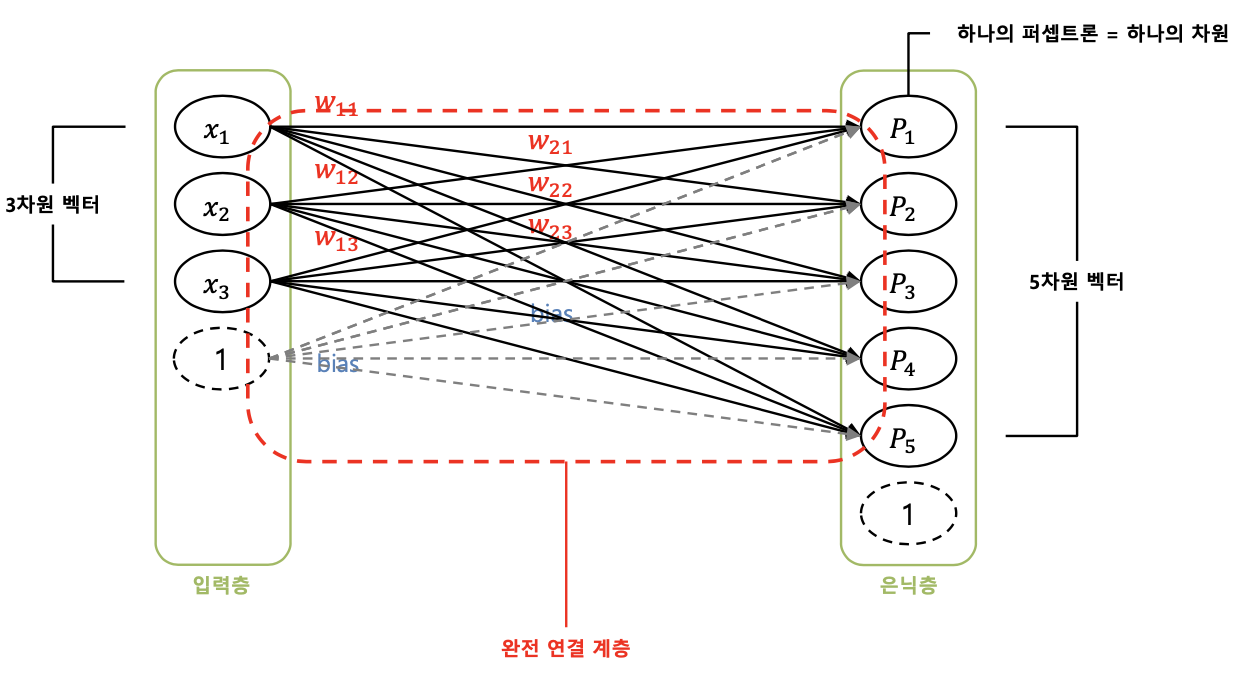
단층 퍼셉트론 : 입력과 출력만

다층 퍼셉트론(MLP) : 입력과 출력사이에 은닉층이 추가가 되었다.

은닉층이 2개 이상이면 심층 신경망

입력층에서 3차원 벡터라고 하면 퍼셉트론을 합쳐서 하나의 벡터이다.

그럼 3개를 합쳐서 하나의 단어라고 생각해야 한다.



NTD 여기서 D는 3

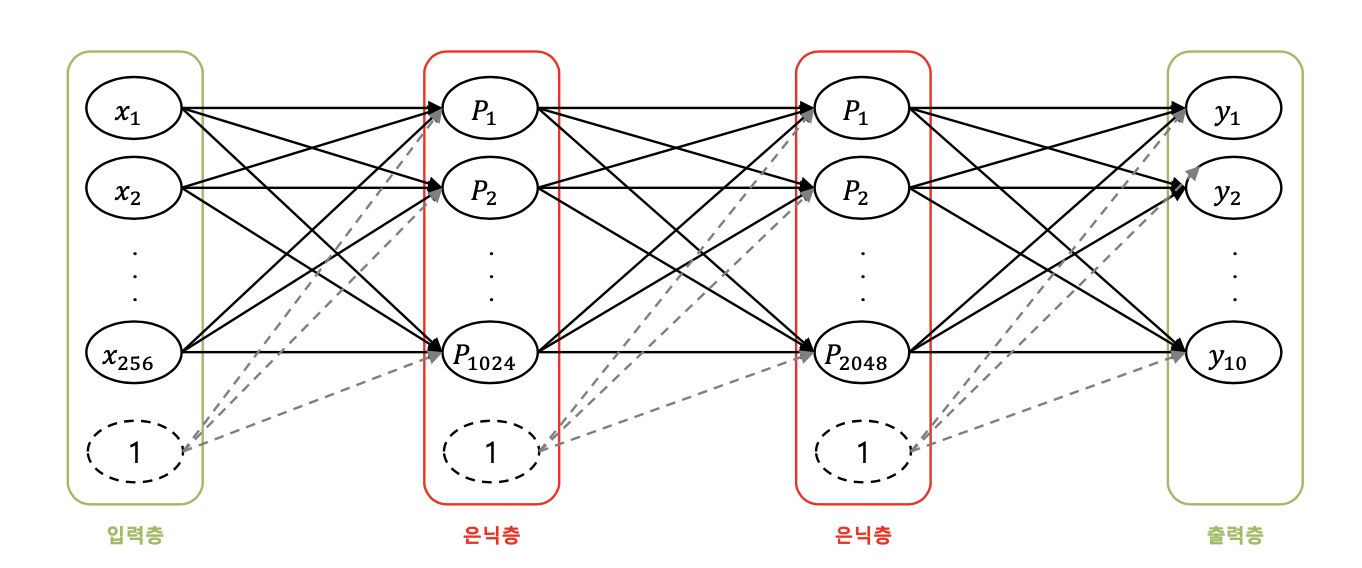
NTH 여기서 H는 5

V는 출력

D차원의 벡터 : 숫자가 D개인 벡터(숫자 배열)

숫자 3개가 들어가서 5개가 나오는 과정을 20번한다.(N:5, T:4 가정)

가중치 매트릭스는 3X5다.



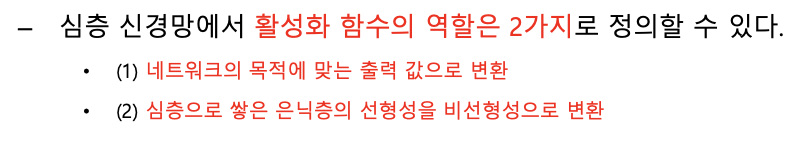
입력층이 256개면 한개의 단어를 256개의 숫자로 표현

은닉층이 1024개면 넘겨줘야 하는 값이 1024개

활성화 함수의 단순한 목적은 내가 원하는 출력의 형태로 변환

1은 출력층 부분

2는 은닉층 부분(단순 퍼셉트론 말고 은닉층이 쌓이면서 추가된 목적)



입력층(임베딩-선형레이어 / 차원을 맞춰줘서)은 D차원의 벡터로 변환

출력층은 은닉차원은 출력차원으로 변환

은닉층은 복잡도를 셋팅

ReLU활성화 함수는 시그모이드 함수의 기울기 손실문제가 없으므로 성능이 좋다(복잡도가 유지)

오차가 최소가 되는 방향으로 매개 변수(가중치 행렬)들을 업데이트 하고, 이 과정을 반복하는 것이 학습이다.

경사하강법 : 기울기가 0이 되는 곳을 찾는다(최솟값을 찾기 위해서)

Chain Rule :