

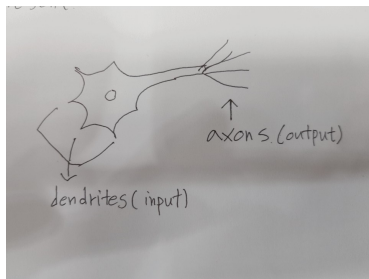
4주차 - Neural Networks : Representation

- 인공 신경망 background

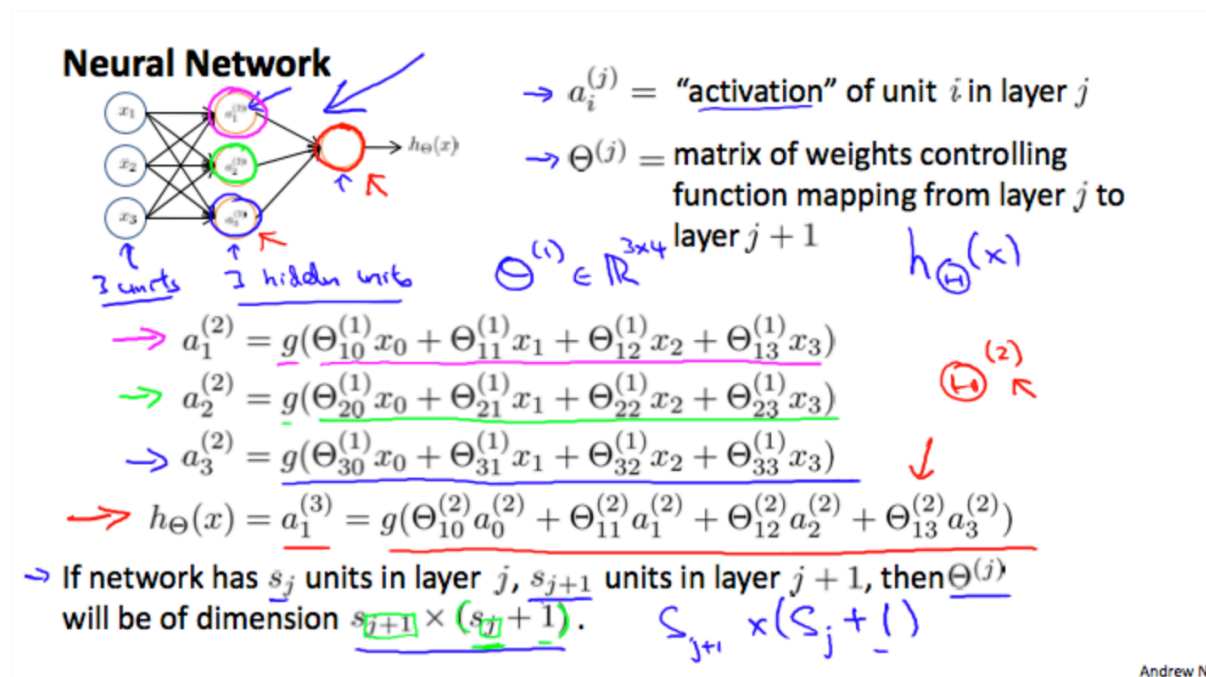
인공 신경망은 1980년대 부터 있었던 개념으로 인간의 뇌의 패턴을 모방한 분류기로 기계학습 최신 기술이다.

비선형 회귀는 특징이 많아 질 수록 $O(n^2)$ 꼴로 복잡해지는데, 생각보다 실생활에 비선형 회귀를 해야 할 경우가 많다. (주로 그림이나 사운드 같은 비정형 데이터들..) 이런 경우 인공 신경망 기술이 효과적 대안으로 이용된다고 한다. 다만 해석하기가 힘들다고 함.

- 인공 신경망 개념들



dendrites(input) : 수상돌기, axons(output) : 축삭돌기



input layer : 첫번째 신경망 층, input이 들어오는 곳

output layer : 마지막 신경망 층, output이 산출되는 곳

hidden layer : input, output을 제외한 층.

sigmoid activation function(활성화 함수) : 현재 신경망 층에서 다음 층을 활성화(즉 뉴런의 값을 1로

만드는 것)하는 방법을 나타내는 함수. 사실 그냥 로지스틱 함수이다.

weight : 신경망에서 각 연결선의 가중치를 의미하는데 기계학습 관점으로 보았을 때 parameter로 theta와 같은 의미이다(따라서 여기에서도 big theta로 표기한다). 인공 신경망에서 한 층의 weight는 다음 층 각 뉴런과 관련이 있다. 위 그림 하단부에서 볼 수 있듯이, s를 각 층에 존재하는 뉴런(unit)의 개수라고 하면 각 층의 big theta는 $s(j+1) \times s(j) + 1$ 크기의 matrix가 된다. +1은 x_0 , a_0 등의 bias를 추가한 것 때문에 발생했다.

$a_i^{(j)}$ = "activation" of unit i in layer j

$\Theta^{(j)}$ = matrix of weights controlling function mapping from layer j to layer $j + 1$

- feed forward propagation

$x(\text{input})$ 부터 시작하여 $h(x)$ 까지 단계적으로 a 를 구하면서 hypothesis를 구하는 방법.

- vectorized implementation

다음과 같이 z 라는 새로운 variable을 상정하고 각 단계의 뉴런의 활성정도를 간단하게 표현할 수 있다.

Handwritten mathematical equations showing the relationship between activation, weighted sum, and the previous layer's activation:

$$a_i^{(j)} = g(z_i^{(j)})$$
$$z_i^{(j)} = \sum_{k=1}^n \Theta_{i,k}^{(j-1)} x_k$$
$$z^{(j)} = \Theta^{(j-1)} a^{(j-1)}$$

- 그 외

인공 신경망은 마지막에서 두번째 층과 마지막 층만 두고 보면 마지막에서 두번째 층을 input으로 하는 로지스틱 회귀와 과정이 똑같다. 그렇게 보았을 때 처음 층위에서 마지막에서 두번째 층까지의 과정은 마치 raw한 특성을 다듬고 추상화시켜서 새로운 특성값으로 도출시키는 것과 같다고 볼 수 있다.