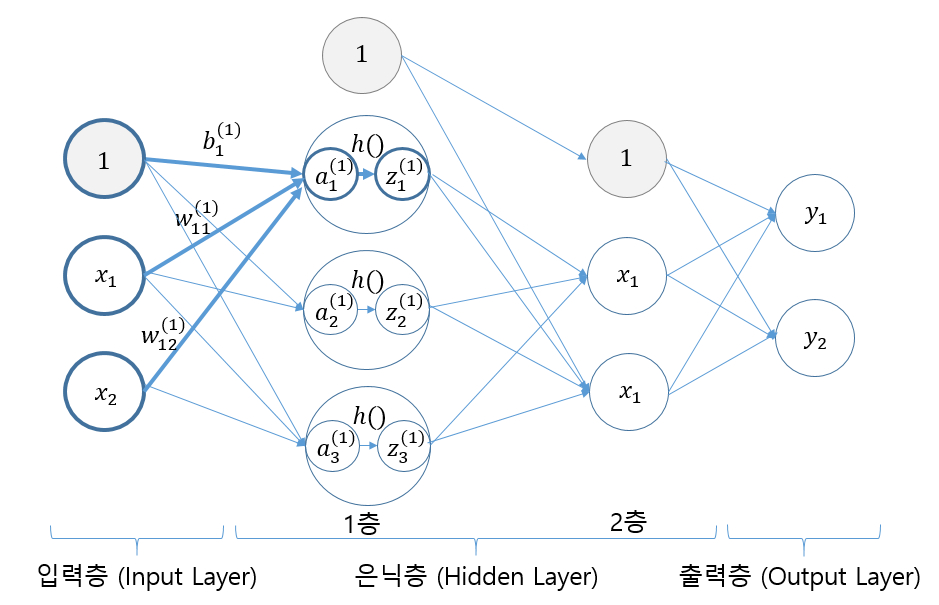
다층 퍼셉트론(Multi Layer Perceptron)



좌측의 입력층부터 0층으로 표시하며 우측으로 갈수록 높은 층으로 표시, 가장 우측 층을 출력층이라고 부르고, 입력층과 출력 층 사이를 은닉층이라고 부르는데, 은닉층은 여러 층으로 구성할 수 있다. 은닉층이 무수히 많은 신경망을 심층 신경망이라고 부른다.

은 1층의 퍼셉트론 중 두 번째임을 나타낸다. 은 n층의 가중치 중에서 n-1층의 j번째 퍼셉트론으로부터 n층의 i번째 퍼셉트론으로 향하는 가중치를 나타낸다. 은 n층의 i번째 퍼셉트론으로 향하는 편향(bias)를 나타낸다. 편향은 한쪽에 치우침을 뜻한다. 가중치가 각각의 입력 신호에 대한 중요도를 조절하는 매개변수라면, 편향은 퍼셉트론이 얼마나 쉽게 활성화하는지를 조절하는 매개변수다. 그리고 h()는 은닉층의 활성화 함수를 의미한다.

퍼셉트론를 수식으로 나타내면 다음과 같다.

입력층으로부터 은닉층의 세 개의 퍼셉트론까지 신호가 전달되는 과정을 행렬의 내적을 이용하여 나타낸 수식은 다음과 같다.

텐서(Tensor) – 동적 크기를 지니는 다차원 데이터 배열

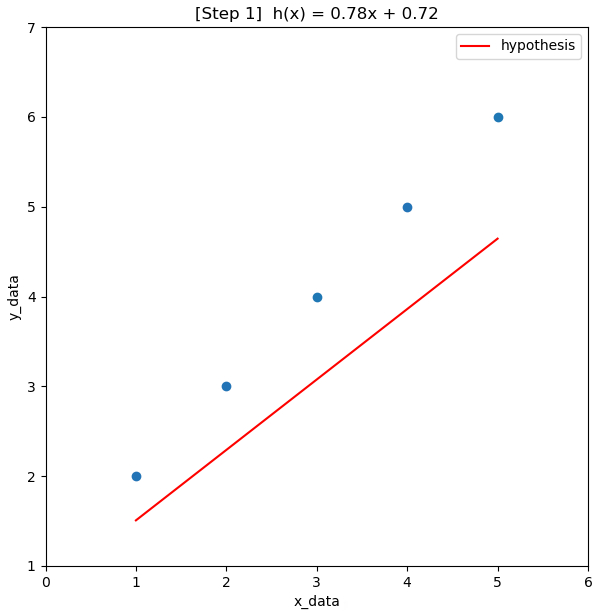
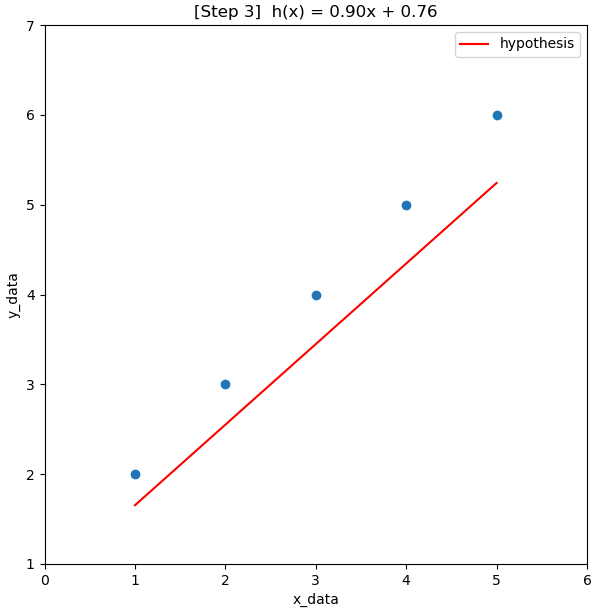
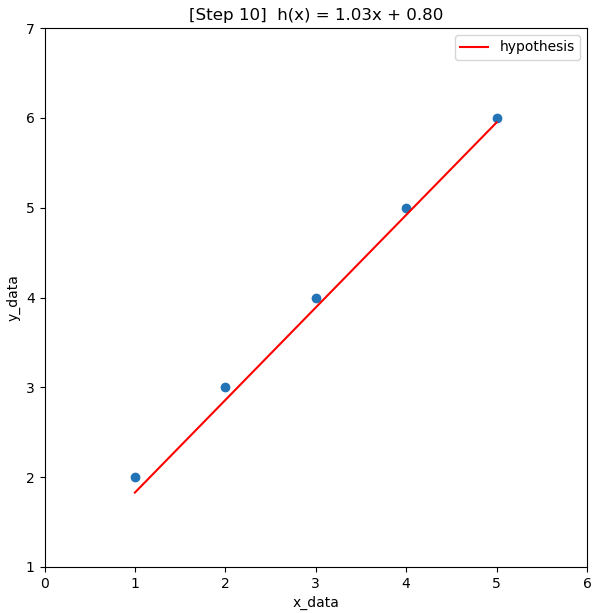
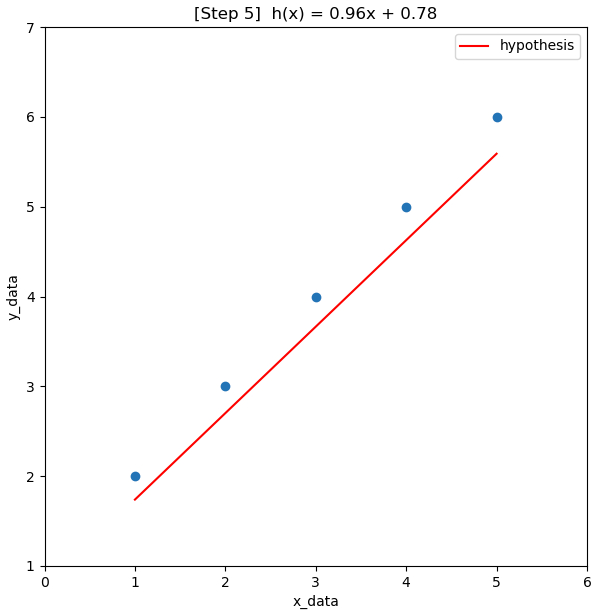
텐서는 n차원의 배열이나 리스트로도 표현할 수 있으며, 프로그램 내에서 모든 데이터는 텐서의 형태로 이동한다. 각각의 텐서는 차원(rank), 형태(shape), 자료형(data type)을 지니는데, 차원은 동적으로 변할 수 있다.

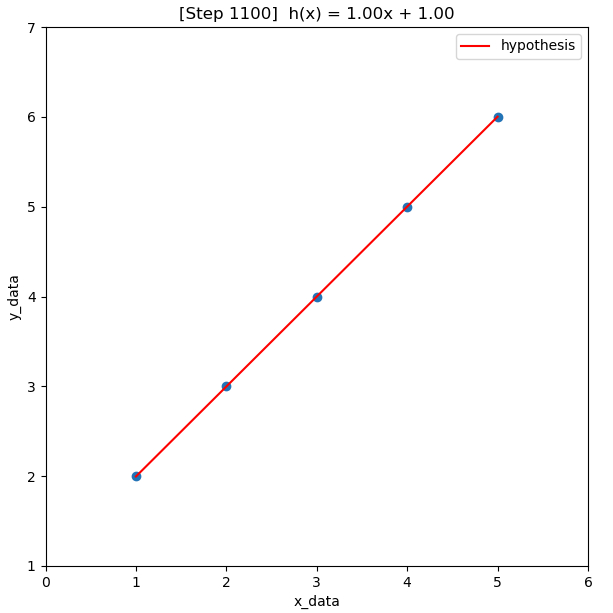
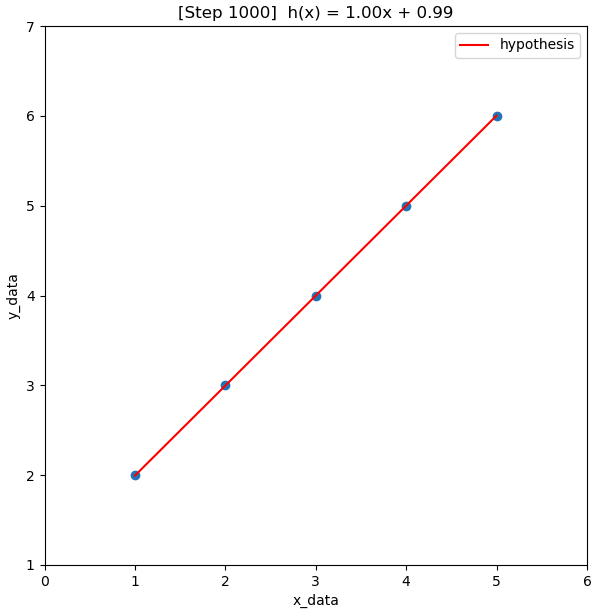
|  |  |  |
| --- | --- | --- |
| 차원 | 수학 엔티티 | 파이썬 코드 예 |
| 0 | 스칼라(크기만 지님) | s = 123 |
| 1 | 벡터 (크기와 방향) | v = [1.1, 2, 2, 3, 3] |
| 2 | 매트릭스 (숫자 테이블) | m = [[1, 2, 3], [4, 5, 6], [7, 8, 9]] |
| 3 | 3-텐서 (숫자 큐브) | t = [[[2], [4], [6]], [[8], [10], [12]], [[14], [16], [18]]] |
| n | n-텐서 | … |

선형 회귀 문제

선형 회귀 문제를 머신러닝으로 푼다는 것은 주어진 데이터를 가장 잘 나타내는 형태 즉 a와 b 값을 찾아내는 것이다. 머신러닝에서는 라고 표현하는데, w(가중치)와 b(편향)의 표현만 다른 것이다. 컴퓨터가 주어진 데이터를 이용하여 가설 h(x)에서 최적의 w와 b값을 찾는 과정을 학습이라 하며, w값과 b값을 알아내면 결국 임의의 x값이 주어졌을 때 그에 맞는 y값을 추론할 수 있게 된다.

파란 점: 실제 데이터, 빨간 선: 가설 주어진 데이터를 이용해 선형 회귀 모델을 학습하는 과정



1100회 수행하여 직선을 찾았다.