**import modules**

In [ ]:

**import** **pandas** **as** **pd**

**import** **numpy** **as** **np**

**import** **matplotlib.pyplot** **as** **plt**

**from** **tensorflow.keras.models** **import** Sequential

**from** **tensorflow.keras.layers** **import** Dense

**data preparation and visualization**

In [ ]:

data=pd.read\_csv(‘regression.csv’)

하면 x, y 데이터 받아짐

x=data[‘X’]

Y=data[‘Y’]

plt.plot(X,Y,’.’)

* 받아진 데이터가 scatter plot됨. 그래프를 ‘.’ 으로 표시함

y=ax1+bx2+c ->3차원 상의 평면을 구할 수 있음 (입력2개)

**create model**

In [ ]:

model=sequential()

model.add (Dense(1, input\_shape=[1])

🡪y=ax+b와 같은 것을 만듦

입력 2개일 경우 input\_Shape=[2])

model.compile(optimizer=’sgd’,loss=’mse’)

**train model**

In [ ]:

model.fit((X,Y, epochs = 1000)

**model prediction**

In [ ]:

x=[0,1,2,3,4,5,6,7,8,9,10]

y=model.predict(X)

plt.plot(X,Y,’.’)

plt.plot(X,y,’x’)

plt.show()

* ‘.’과 ‘x’로 이뤄진 그래프가 나옴

**what are a and b in y = ax + b?**

In [ ]:

weights=model.layers(0)/get\_weights()

print(weights)

190507

함수의 형태

y=ax+b ->입력1, 출력1 의 가장 간단한 형태 (확장하려면 a1x1+a2x2+..)

y=ax^2+b 역시 입력이 2개가 아닌 1개

입력이 여러 개, 출력이 여러 개 인 경우.

x1 y1

x2 y2

1

각 화살표 값이 a1,a2,a3…가 됨.

Dense=fully connected

CNN=fully connected 아님. 이럴 때 더 잘 작동하는 경우도 있음. connection의 값(a)가 0인 경우

boston housing value 의 입력값 -> 집값에 영향을 미쳤는지…

x1 O y1

x2 O y2

1

3x2 – 2x2 행렬 이어진 경우. nod와 edge로 이어져 있음. nod 간의 관계를 행렬로 표현함.

연결되어 있는 3개의 layer. 그러나 x1과 y1처럼 바로 연결할 수 없는 nod들도 있음. 그래프 자체에 constraint가 있다고 이해하면 됨.

connection에 방향성이 있음. 위 그래프 같은 경우 무조건 왼쪽에서 오른쪽 방향으로.

중간에 layer를 하나 더 넣을 수 있음. (4x3, 3x4, 4x2 행렬 세개 이어진 형태)

이 경우 입력값 4개만 알고 출력값 2개만 있으면 중간 행렬들을 다 알 수 있다.

화살표에 얼마나 weight (가중치)를 줄지 표현 가능. 예를 들어 a 값이 0인 경우 가중치가 없다. weight matrix라고 함.

이때 데이터는 반드시 pair 가 되어야 함. 입력 값은 input, 출력 값은 target 이라고 부름. 계산된 출력이라기 보다는 데이터 속에 존재하는 input의 pair data에 해당하기 때문. (정답 값)

계산된 출력 값과 target값에는 차이가 있다. weight matrix를 훈련하다 보면 이 차이를 줄일 수 있다. 이 차이를 error라고 부름 (target과 output 간의 차이)

error를 줄이려면 predict를 잘해야 함 + 좋은 data가 많아야 함.

입력, 출력 사이에 있는 layer를 모두 hidden layer 라고 함.

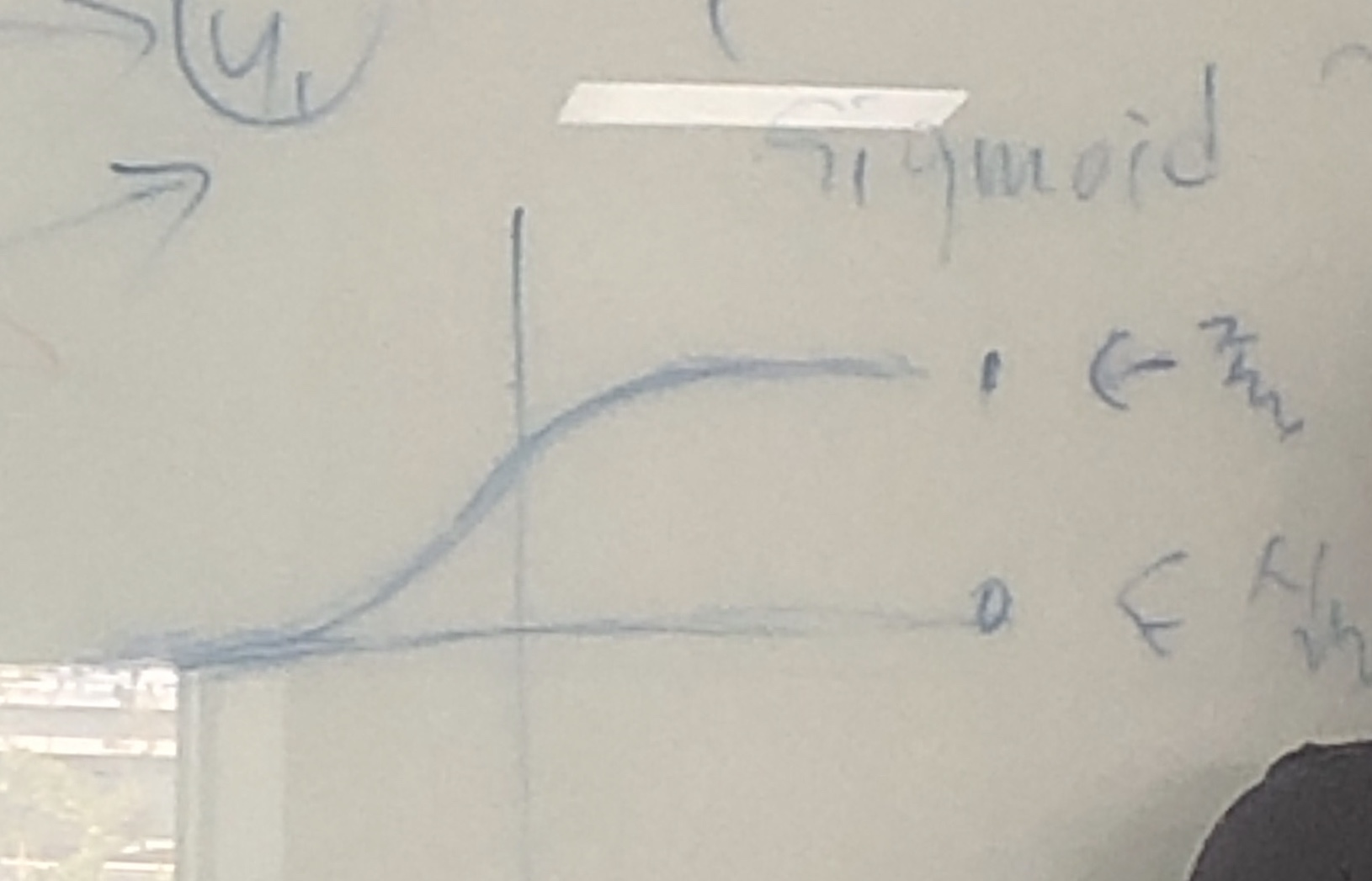
layer의 개수 + 각 layer의 nod의 개수 많아지면 deep

linear

sigmoid- 모든 linear data를 0과 1 사이에 나타내라 (??)

입력은 숫자인데 출력이 숫자가 아닌 경우 (예: 음성인식)

continuous한 입력이 들어오더라도 출력 부분은 0과 1 사이로 나타나도록 (예: 영화리뷰에 대한 입력 값-영화가 좋은지 싫은지를 0과 1 사이로 나타내는 출력 값)



190509

다양한 형태의 non-linearity를 붙여주는 형태

집값에 영향을 주는 요소들

(x\_train, y\_train) , (x\_test, y\_test)=

tensorflow.keras.datasets.boston\_housing.load\_data(path=’boston\_housing.npz)

boston housing price에 영향을 미치는 요인 13개

식으로 쓰면 y=a1x1+a2x2+….+a13x13

x1

x2

x3 y 화살표가 각각 a1,a2,a3…a13이 됨(데이터가 여기 들어감)

.

.

x13

예측을 잘 해야 함