

과제명 : 지식 그래프 기반 음식점 광고 콘텐츠  
(Title) 추천 알고리즘 개발

과제번호 :  
(Project No.)

### 3) 행렬 곱으로 이해하는 신경망

순전파는 행렬의 곱셈으로 이해할 수 있다.

#### -1. 순전파 (Forward Propagation)

입력층 → 출력층 연산 진행

#### -2. 행렬 곱으로 순전파 진행

$$h_1 = x_1w_1 + x_2w_2 + x_3w_3 + b_1$$

$$h_2 = x_1w_4 + x_2w_5 + x_3w_6 + b_2$$

$$[y_1, y_2] = \text{softmax}([h_1, h_2])$$

↓

$$\begin{matrix} X & W & B & Y \\ \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} & \begin{bmatrix} w_1 & w_4 \\ w_2 & w_5 \\ w_3 & w_6 \end{bmatrix} & \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} & \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ \text{Sample-1} & & & \end{matrix}$$

인공신경망이 다수의 샘플을 동시에 처리하는 것은 '배치연산'

### 4) 딥러닝의 학습방법

#### -1. 손실함수 (Loss function)

회귀에서는 MSE, 분류에서는 크로스 엔트로피를 주로 쓴다.

① MSE: 연속형 변수 예측할 때 사용.

② Binary Cross-Entropy: 출력층에서 시그모이드 사용할 때 사용.

③ Categorical Cross-Entropy: 출력층에서 소프트맥스 사용할 때 사용.

#### -2. 배치 크기 (Batch Size)에 따른 경사 하강법

배치: 가중치 등의 매개변수의 값을 조정하기 위해 사용하는 데이터의 양.

##### ① 배치 경사하강법 (Batch Gradient Descent)

가장 기본적인 경사 하강법.

옵티마이저 중 하나로 손실(loss)을 구할 때 전체 데이터 고려.  
1 Epoch 당 시간이 오래걸리고 메모리를 크게 뺏긴다.

##### ② 배치 크기가 1인 확률적 경사하강법 (Stochastic Gradient Descent, SGD)

매개변수 값 조정 시 전체 데이터가 아니라 랜덤으로 선택한 하나의 데이터에 대해서만 계산하는 방법.

매개변수의 변경 폭이 불안정하고 때로는 배치 경사하강법보다 낮은 정확도를 보인다.

##### ③ 미니 배치 경사하강법 (Mini-Batch Gradient Descent)

배치 크기를 지정하여 매개변수 값을 조정  
배치 경사하강법보다 빠르며 SGD보다 안정적이다.  
배치 크기는 2^n 이 보편적이다.

### -3. 옵티마이저 (Optimizer)

#### ① 모멘텀 (Momentum)

경사하강법에서 계산된 접선의 기울기에 한 시점 전의 접선의 기울기 값을 일정 비율만큼 반영.

→ 로컬 미니멈에 빠져든 관성의 힘으로 빠져나올 수 있다.

#### ② 아다그라드 (Adagrad)

매개변수들은 각자 의미하는 바가 달라 동일한 학습률을 적용하는 것이 비효율적이다.

그래서 각 매개변수에 서로 다른 학습률을 적용시킨다.

#### ③ 알엠에스프로프 (RMSprop)

아다그라드의 보완책

나중에 학습률이 지나치게 떨어지는 단점 보완.

#### ④ 아담 (Adam)

RMSprop + Momentum

### -4. 역전파 (Back Propagation)

역전파 1단계



$w_5$ 를 업데이트 하기 위해서

$$\frac{dE_{total}}{dw_5} = \frac{dE_{total}}{dO_1} \times \frac{dO_1}{dz_3} \times \frac{dz_3}{dw_5}$$

$$w_5^+ = w_5 - \alpha \frac{dE_{total}}{dw_5}$$

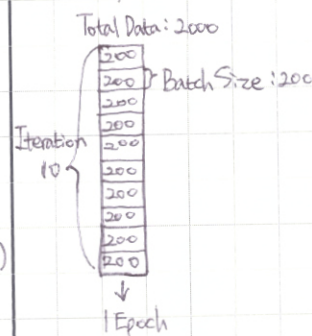
역전파 2단계

1단계와 같은 원리로  $w_1, w_2, w_3, w_4$ 를 업데이트 해준다.

업데이트 된 가중치를 통하여 순전파 진행.

\* 인공신경망의 학습은 오차를 최소화하는 가중치를 찾는 목적으로 순전파와 역전파 반복하는 것을 말한다.

### -5. Epoch 와 Batch Size 와 Iteration



Epoch: 인공신경망에서 전체 데이터에 대해 순전파와 역전파가 끝난 상태 문제지를 푼 횟수.

횟수에 따라 과적합, 과소적합 발생  
Batch Size: 몇개의 데이터 단위로 매개변수를 업데이트 하는지

Iteration: 배치의 수.

$$2000 / 200 = 10 \text{ Iteration.}$$

기록자  
(Recorded by)

이상준

서명  
(Signature)

이상준

일자  
(Date)

2022.4.23

점검자  
(Reviewed by)

정선태

서명  
(Signature)

정선태

일자  
(Date)

2022.4.26

이 노트의 지적재산권은 숭실대학교 산학협력단에 있음  
(Intellectual property right for this notebook belongs to Foundation of Soongsil)