

# 5-8단계

## 5단계 역전파 이론

전파되는 데이터는 모두  $y$ 의 미분값( $dy/dy, dy/db, dy/da, dy/dx$ )

손실함수의 각 매개변수에 대한 미분을 계산해야함

미분값을 출력에서 입력방향으로 전파하면 **한 번의 전파만으로 모든 매개변수에 대한 미분을 계산** 할 수 있음 → 미분을 역전파하는 방식 사용

변수는 '통상값', '미분값' 존재

함수는 '통상 계산(순전파)', '미분값을 구하기 위한 계산(역전파)' 존재

역전파 시에는 순전파 시 이용한 데이터가 필요

## 6단계 수동 역전파

기울기(gradient): 다변수(e.g. 벡터, 행렬)에 대한 미분

## 7단계 역전파 자동화

**Define-by-Run** (동적 계산 그래프): 딥러닝에서 수행하는 계산들을 계산 시점에 '연결'하는 방식

### 변수와 함수의 관계

함수에게 변수는 '입력'과 '출력'

변수에게 함수는 '창조자' 혹은 '부모'

## 8단계 재귀에서 반복문으로

반복문으로 바꾸는 이유!

1. 효율성: 재귀는 함수를 재귀적으로 호출할 때마다 중간 결과를 메모리에 유지하면서(스택에 쌓이면서) 처리를 이어가기 때문에 일반적으로 반복문 방식이 효율이 더 좋다. &

‘꼬리재귀(tail recursion)기법을 이용하여 재귀를 반복문처럼 실행.

생각해본 사항)

1. Define-by-Run의 개념과 장점은 무엇이기에 딥러닝에서 수행하는 계산들을 계산 시점에 ‘연결’하는 방식이 되는것인가?
2. 역전파 자동화에서 창조자가 없을때 Variable인스턴스는 함수 바깥에 생성되었음을 뜻한다고할때 왜 높은 확률로 사용자가 만들어 건넨 변수일까?