5장

신재현

- 1. 신경망을 통과해서 도달한 결과값은 오차를 포함한다
- 2. 최종적으로 나온 오차는 여러 계산 과정들을 거치며 합산된 값
- 3. 각 계산 과정마다 어느 정도의 오차가 발생했는지를 관찰해서 계산과정을 다시 수정해준다
- 4.오차가 많이 발생하게 한 요인들은 큰 폭으로 값을 변화시킨다
- 5. 수치 미분 대신 오차역전파를 사용하면 효율이 올라간다

- Forward로 계산된 예측 값과 실제 값을 비교해 계산된 cost(error)를 backward로 미분값 혹은 예측에 필요한 값들을 계산한다. 정확한 이해를 위해
- 계산 그래프의 역전파로는 각 노드의 미분을 구할 수 있다.

계산 그래프

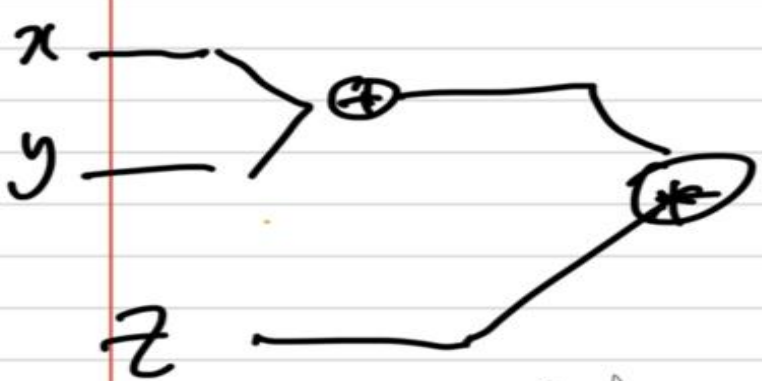
backpropagation

$f(x, y, z) = (x+y)z$

① $q = x+y$ $\quad \frac{\partial q}{\partial x} = 1, \ \frac{\partial q}{\partial y} = 1$ ② $\frac{\partial f}{\partial q} = z, \ \frac{\partial f}{\partial z} = q$

$f = q \cdot z$

$\left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right]$



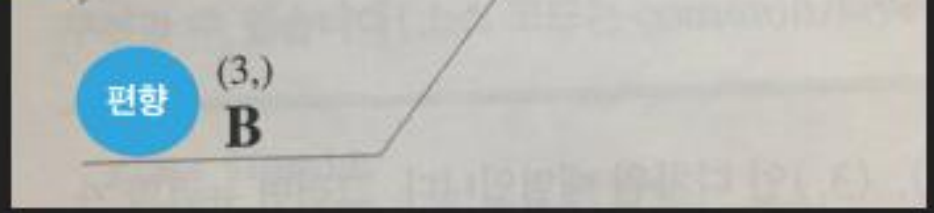
$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \cdot \frac{\partial q}{\partial x}$$

layer_naive.py

곱셈 계층과 덧셈 계층의 구현입니다.

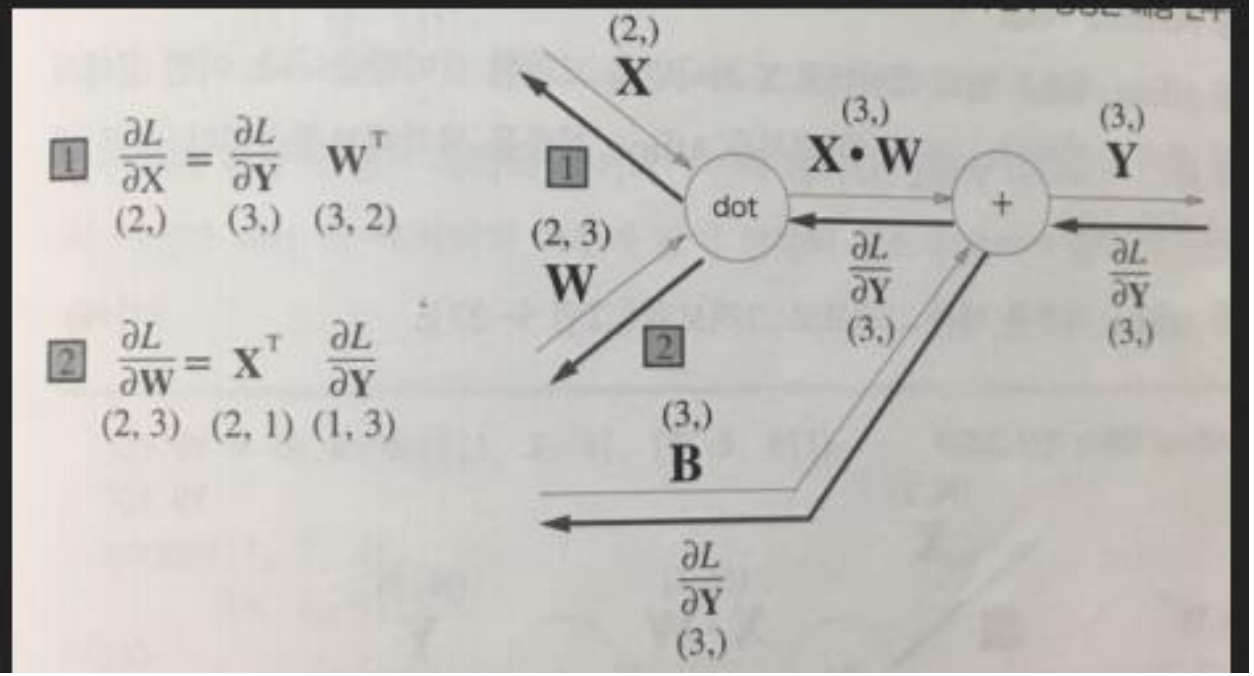
```python
class MulLayer:
    def __init__(self):
        self.x = None
        self.y = None

    def forward(self, x, y):
        self.x = x
        self.y = y
        out = x * y

        return out

    def backward(self, dout):
        dx = dout * self.y  # x와 y를 바꾼다.
        dy = dout * self.x

        return dx, dy
```

```python
class AddLayer:
    def __init__(self):
        pass

    def forward(self, x, y):
        out = x + y

        return out

    def backward(self, dout):
        dx = dout * 1
        dy = dout * 1

        return dx, dy
```



AFFINE/SOFTMAX 계층 구현하기

Affine 계층의 역전파 계산 그래프



정리

- Neural net학습시킬 때 forward propagation-backward propagation은 서로 의존하는 관계
- Chain rule을 적용하기 위해서는 모든 중간 변수를 저장하고 있어야 GRADIENT인 tensor를 계산할 수 있다.
- 수치 미분과 오차역전파법의 결과를 비교하면 오차역전파법의 구현에 잘못이 없는지 확인할 수 있다(기울기 확인).
- 훈련을 위해서 대량의 연산이 필요 매 시도마다 이전 검사와는 다른 ㅡㅌㄱ성을 고려하여

- 신경망을 훈련할 때, 훈련 데이터를 네트워크의 첫 번째 레이어에 넣으면, 수행 작업을 기준으로 입력의 정확도를 나타내는 가중치가 할당되며, 그 후 가중치를 모두 합산해 최종 출력이 결정됩니다.
- 이미지 인식 네트워크에 고양이의 이미지 데이터를 입력했다고 가정해볼게요. 첫 번째 레이어에서는 경계선을 찾습니다. 다음 단계는 직사각형, 원과 같은 좀더 세부적인 경계의 형태를 가려냅니다. 그리고 세 번째 레이어에서는 눈의 크기나 코의 모양과 같은 형태의 특징을 구분해냅니다. 각 레이어는 생성된 가중치의 총합을 바탕으로 최종 출력을 낼 때까지 다음 레이어로 이미지를 전달합니다.