STEP36. 고차 미분 이외의 용도

고차 미분 계산 정리

- 고차 미분을 하기 위해 역전파 시 수행되는 계산에 대해서도 연결을 만들도록 함
- 역전파의 계산 그래프를 만들수 있음
- 고차 미분 외에 어떻게 활용할 수 있는지를 살펴봄

> **Double Backpropagation**
> - 역전파로 수행한 계산에 대해 또 다시 역전파를 수행함
> - Double backprop 은 현대적인 딥러닝 프레임워크 대부분이 지원함

double backprop의 용도

double backprop 활용 용도

- 미분이 포함된 식에서 다시 한번 미분 수행
  - $y' = 2x, \quad z = (y')^3 + y = 8\boldsymbol{x}^3 + \boldsymbol{x}^2$
  - $z' = 24\boldsymbol{x}^2 + 2x$
  - a, b 가 정수일 때 $f(\boldsymbol{x_0}, \boldsymbol{x_1}) = \boldsymbol{b}(\boldsymbol{x_1} - \boldsymbol{x_0}^2)^2 + (a - \boldsymbol{x_0})^2$

DeZero를 사용하여 문제 계산

```
x = Variable(np.array(2.0))
y = x ** 2
y.backward(create_graph=True)
gx = x.grad
x.cleargrad()

z = gx ** 3 + y
z.backward()
print(x.grad)
```

- y.backward(create_graph=True) 는 미분을 하기 위한 역전파 코드
- 이 코드가 계산 그래프를 생성
- 역전파가 만들어낸 계산 그래프를 사용하여 새로운 계산을 하고 다시 역전파함
- 위와 같은 형태의 문제 해결
  - 미분식을 구하고, 그 식을 사용하여 계산 후 또 다시 미분하는 문제

딥러닝 연구에서의 사용 예

WGAN-GP 논문

- 수식을 최적화

$$L = \mathop{\mathbb{E}}_{\tilde{\boldsymbol{x}} \sim \mathbb{P}_g} [D(\tilde{\boldsymbol{x}})] - \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathbb{P}_r} [D(\boldsymbol{x})] + \lambda \mathop{\mathbb{E}}_{\hat{\boldsymbol{x}} \sim \mathbb{P}_{\hat{\boldsymbol{x}}}} \left[(\| \overset{\text{기울기}}{\underline{\nabla_{\hat{\boldsymbol{x}}} D(\hat{\boldsymbol{x}})}} \|_2 - 1)^2 \right]$$

- 최적화하는 식에 기울기(텐서의 각 원소에 대한 미분 결과)가 들어 있음
- 이 기울기는 첫 번째 역전파에서 구할 수 있음
- 이 기울기를 사용하여 함수 L을 계산하고, 함수 L을 최적화하기 위해 두 번째 역전파 수행

DeZero의 역전파를 수정하여 double backprop를 가능하게 됨