

## 7강 연습문제

1. (i) 5개의 데이터를 묶은 행렬

$$\begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 3 & -3 & 2 \\ 4 & -4 & 2 \\ 5 & -5 & 4 \end{pmatrix}$$

에 대하여 확대는  $\gamma = [\sqrt{2}, 2\sqrt{2}, \sqrt{30}]$ , 이동은  $\beta = [2, 4, 5]$ 로써 배치 정규화 (batch normalization)를 적용한 결과를 구하시오.

- (ii) 코드로 계산하시오.

2. (i) 5개의 데이터를 묶은 행렬

$$\begin{pmatrix} -2 & 0 & 4 \\ -1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 3 & 1 \\ 2 & 4 & 6 \end{pmatrix}$$

에 대하여 확대는  $\gamma = [\sqrt{2}, 2\sqrt{2}, 4\sqrt{5}]$ , 이동은  $\beta = [2, 4, 10]$ 로써 배치 정규화 (batch normalization)를 적용한 결과를 구하시오.

- (ii) 코드로 계산하시오.

3. (i) (2번 문제 계속) batch normalization층으로 흘러 들어온 미분이

$$\begin{pmatrix} 1 & -1 & 4 \\ -1 & 1 & -4 \\ 1 & -1 & 4 \\ -1 & 1 & -4 \\ 1 & -1 & 4 \end{pmatrix}$$

일 때, 미분  $\frac{\partial L}{\partial \beta}$ 와  $\frac{\partial L}{\partial \gamma}$ 의 값을 구하시오.

- (ii) 코드로 계산하시오.

4. (i) 신경망이



와 같이 주어져 있다. 첫번째 Affine층의 가중치 행렬  $W_1$ , 두번째 Affine층의 가중치 행렬  $W_2$ 가 각각

$$W_1 = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad W_2 = \begin{pmatrix} \log 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \log 2 & 0 \end{pmatrix}$$

이고 편향은 없다. 배치정규화층(batch normalization)의 확대(scale)는  $\gamma = [1, 2, 3]$ , 이동(shift)은  $\beta = [0, -1, -2]$ 이다. 입력 데이터  $[1, 2, 3], [3, 0, 1]$ 이 배치 묶음으로 입력되고 라벨이 각각  $[1, 0, 0], [0, 1, 0]$ 일 때 손실함수값을 구하시오.

- (ii) 코드로 계산하시오.

5. 다음 코드를 실행했을 때 출력될 값들을 대략적으로 쓰시오.

```
gamma = np.array([1,2,3])
beta = np.array([4,5,6])
BN = BatchNormalization(gamma, beta)
x = 2*np.random.randn(1000,3)+3
y = BN.forward(x)
print(np.mean(y,axis=0))
print(np.mean((y - np.mean(y,axis=0))**2, axis=0))
```