

8강 연습문제

1. 출력될 9개를 순서대로 쓰시오.

```
x=np.arange(18).reshape(2,3,3)

for i in range(2):
    print(x[i,:,:])

for j in range(3):
    print(x[:,j,:])

for k in range(3):
    print(x[:, :,k])
```

2. 7강 2번 문제를 TimeRNN 클래스로 검산을 하시오.

3. 데이터에 대한 가중치 행렬 W_x , hidden state에 대한 가중치 행렬 W_h , 편향 벡터 b 가 각각

$$W_x = 0.01 \times \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}, \quad W_h = 0.02 \times \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}, \quad b = [0, 0, 0]$$

인 RNN층을 생각하자. 배치크기가 3, 시간 블록의 길이가 5, 벡터표현의 차원이 3인 데이터

$$xs = \begin{pmatrix} [0, 1, 2] & [1, 2, 3] & [2, 3, 4] & [3, 4, 5] & [4, 5, 6] \\ [1, 2, 3] & [2, 3, 4] & [3, 4, 5] & [4, 5, 6] & [5, 6, 7] \\ [2, 3, 4] & [3, 4, 5] & [4, 5, 6] & [5, 6, 7] & [6, 7, 8] \end{pmatrix}$$

가 RNN층으로 들어오고 위층에서 흘러들어온 미분은

$$dhs = 0.1 \times \begin{pmatrix} [0, 1, 2] & [1, 2, 3] & [2, 3, 4] & [3, 4, 5] & [4, 5, 6] \\ [1, 2, 3] & [2, 3, 4] & [3, 4, 5] & [4, 5, 6] & [5, 6, 7] \\ [2, 3, 4] & [3, 4, 5] & [4, 5, 6] & [5, 6, 7] & [6, 7, 8] \end{pmatrix}$$

라 하자.

TimeRNN 클래스를 사용하여 hidden state의 묶음 hs 와 밑으로 흘러보내는 미분 dxs 와 파라미터들의 미분

$$\frac{\partial L}{\partial W_h}, \quad \frac{\partial L}{\partial W_x}, \quad \frac{\partial L}{\partial b}$$

을 수치적으로 구하시오.