

## 深層学習 day3 section 2

Q) 勾配爆発を防ぐために勾配クリッピングを行うという手法がある。具体的には勾配のノルムがしきい値を超えたら、勾配のノルムとしきい値に正規化するというものである。以下は、クリッピングを行う関数である。(±)に入るのは?

```
def gradient_clipping(grad, threshold):
```

コピペ

```
    grad: gradient
```

勾配 しきい値

```
    norm = np.linalg.norm(grad)
```

```
    rate = threshold / norm
```

```
    if rate < 1:
```

```
        return
```

```
    return grad
```

(±)

Ans) (1)  $\text{gradient} \times \text{rate}$

Q) 以下の PG は LSTM の順伝播を行う PG である。

ただし、 $\text{sigmoid}$  関数は要素ごとにシグモイド関数を用いる関数である。

(1) にあてはまるのはどれか。

```
def lstm(x, prev_h, prev_c, W, U, b):
```

```
    lstm_in = _activation(x.dot(W.T) + prev_h.dot(U.T) + b)
```

```
    a, i, f, o = np.hsplit(lstm_in, 4)
```

```
    a = np.tanh(a)
```

```
    input_gate = _sigmoid(i)
```

```
    forget_gate = _sigmoid(f)
```

```
    output_gate = _sigmoid(o)
```

```
    c = (1+)
```

```
    h = output_gate * np.tanh(c)
```

```
    return c, h
```

Ans) (3)  $\text{input\_gate} * a + \text{forget\_gate} * c$

∵ CEC の変えている情報のため。