

## 深層学習 day 2 Section 3 係数

Q) 以下は  $L_2$  正則化を適用した場合:

パラメータの更新を行う PG である

あるパラメータ  $param$  と正則化がないと

に そのパラメータに伝播される誤差

の勾配  $grad$  が与えられたとする.

最終的な勾配を計算する (2) に

あてはまるのはどれか.  $rate$  は  $L_2$  正則化  
の係数

def ridge(param, grad, rate):

$$grad += rate \times \boxed{(2)}$$

A) (4) param

$\therefore L_2$  ノルムは  $\|param\|^2$  なのだから  $\checkmark$  <sup>正則化</sup>  
その勾配が誤差の勾配に加えられる. 2 は吸収する.

---

Q) L1正則化の場合で.

最終的な勾配を計算する (a) に  
おてはまるのはどれか.

def lasso(param, grad, rate)

$$X = \boxed{(a)}$$

$$\text{grad} += \text{rate} * X$$

A) 3)  $\text{sign}(\text{param})$  ✓

∵ L1 ノルムは  $|\text{param}|$  なのでその勾配が  
誤差の勾配に加えられる。

---