Deep Learning Basic

Jaewon Kim, Dankook Univ.

Chapter 3-3

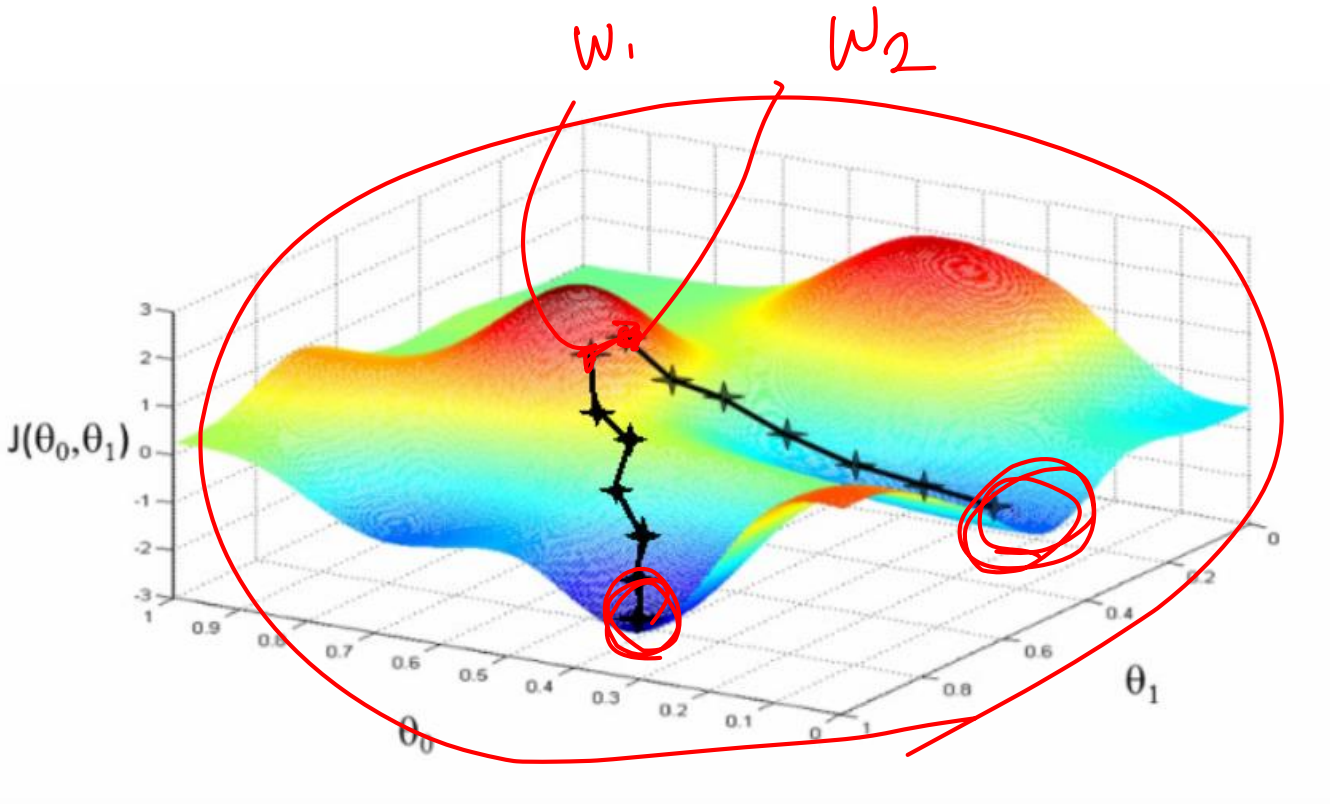
Contents

Part 1. Weight Initialization

- Why we use ?
- Zero Initialization
- Random Initialization
- Xavier Initialization
- He Initialization

Weight Initialization

- Why we use this ?



Reference : https://yngie-c.github.io/deep%20learning/2020/03/17/parameter_init/

Weight Initialization

- Zero_Initialization

Weight parameter를 zero 초기화

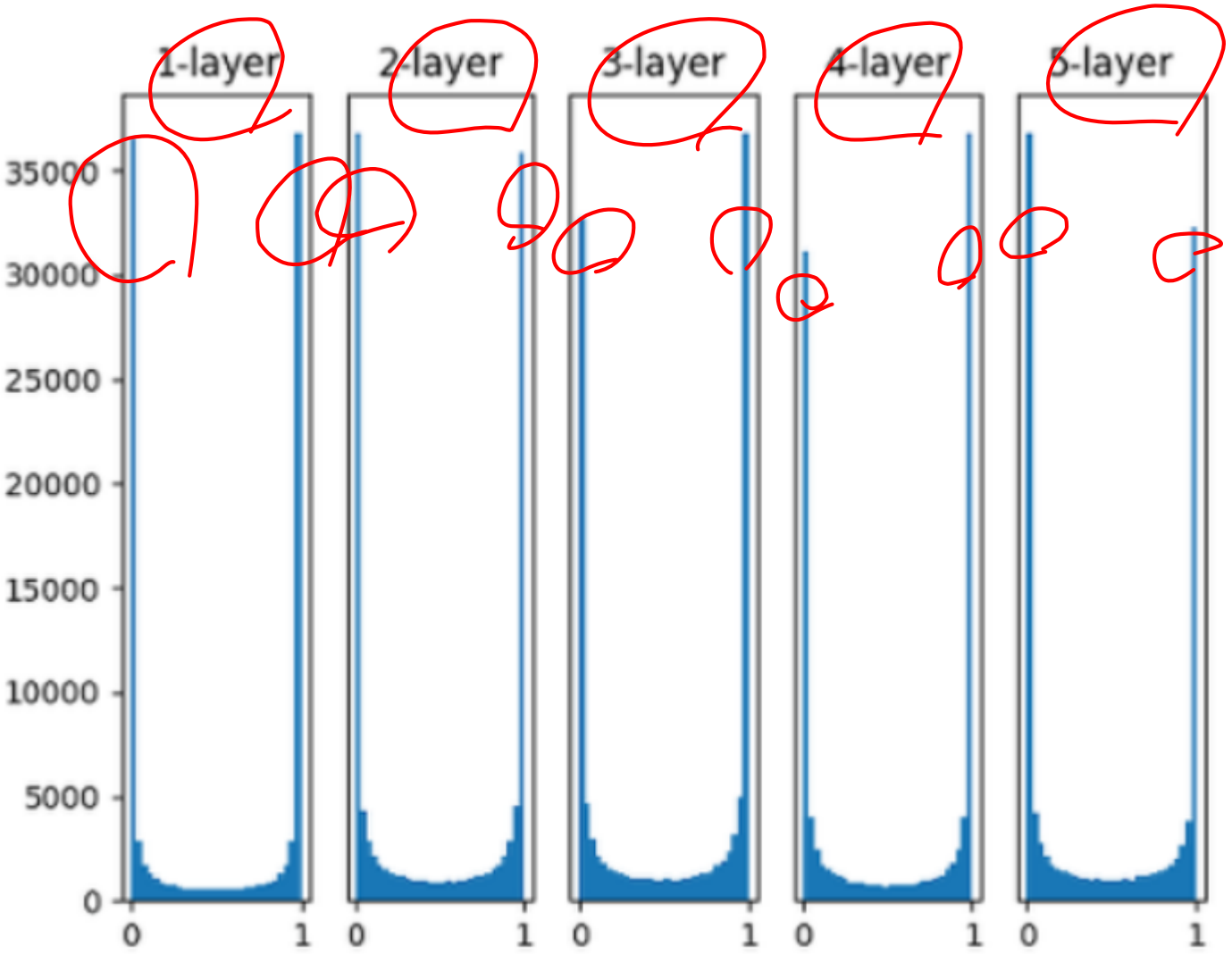
But, Gradient Vanishing

↓

Random

Weight Initialization

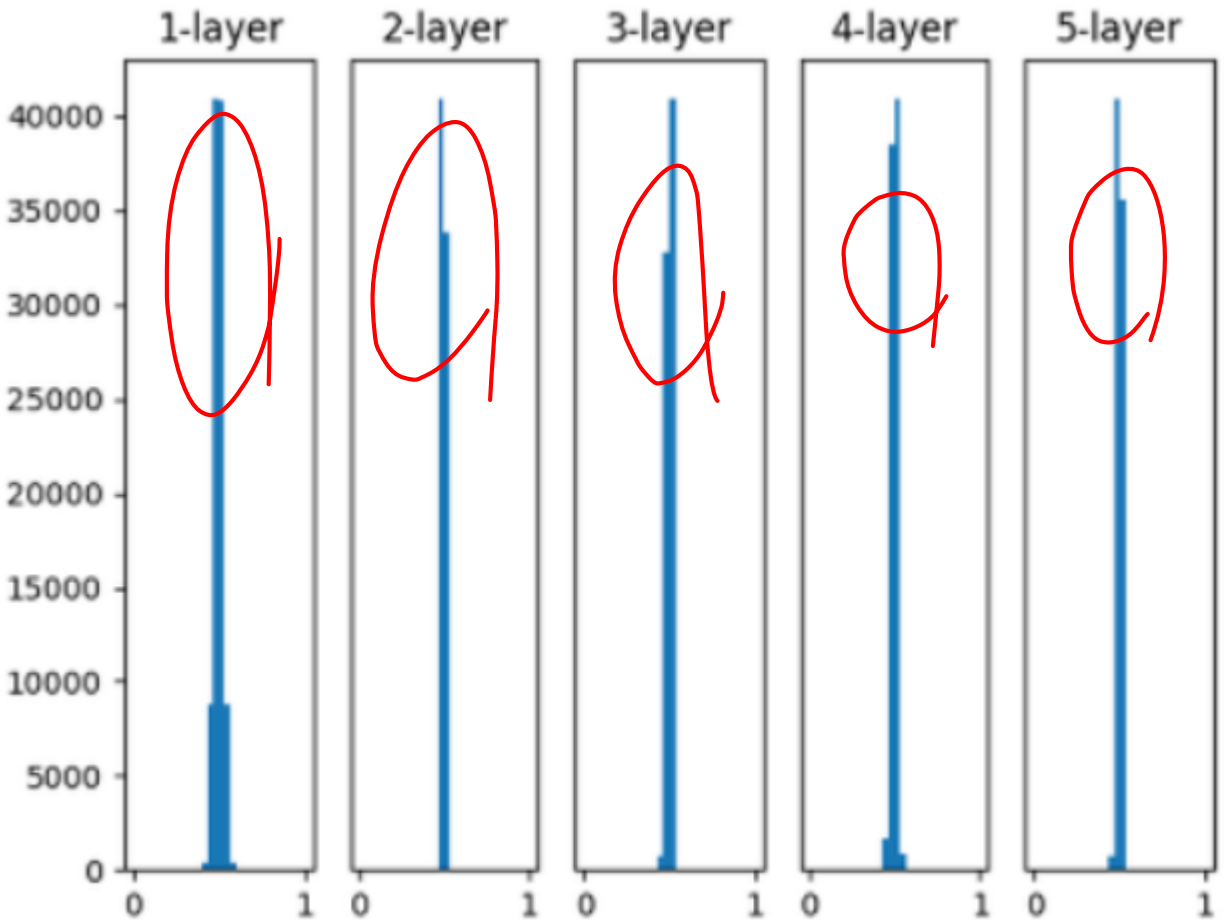
- Using Sigmoid Function



출처 : https://reniew.github.io/13/

Weight Initialization

- Using Sigmoid Function (표준편차=0.01)



출처 : 위와 동일

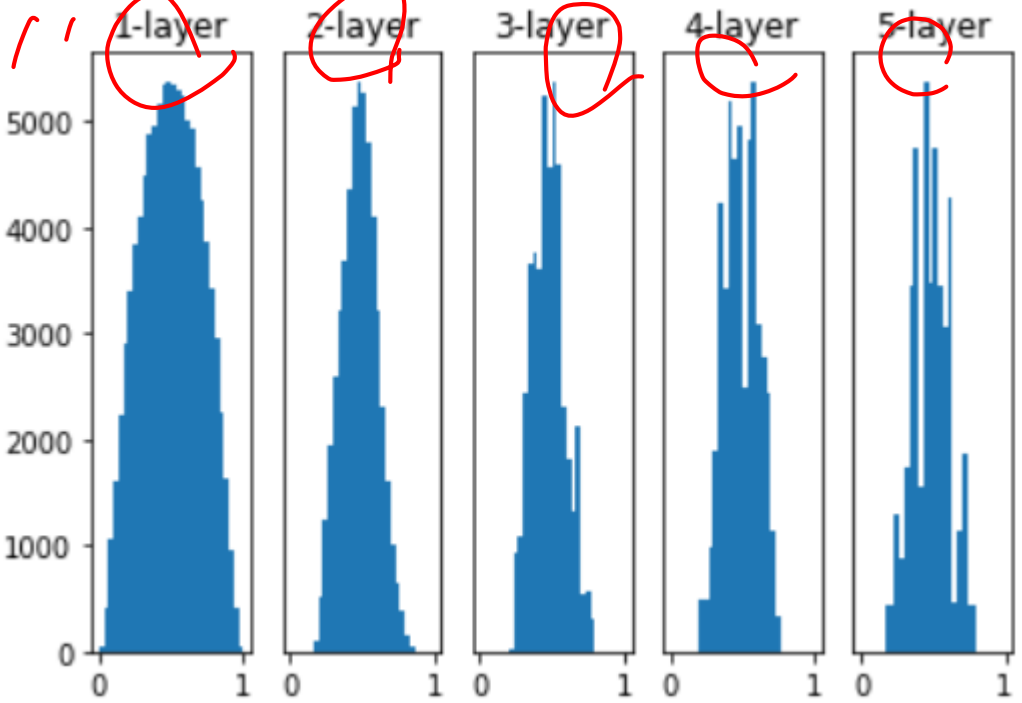
Weight Initialization

- Xavier Initialization

$$W \sim N(0, Var(W))$$

$$Var(W) = \sqrt{\frac{2}{n_{in} + n_{out}}}$$

n_{in} : 이전 hidden layer node
n_{out} : 현재 hidden layer node



- Sigmoid, tanh 사용 / ReLU 사용시 0으로 수렴

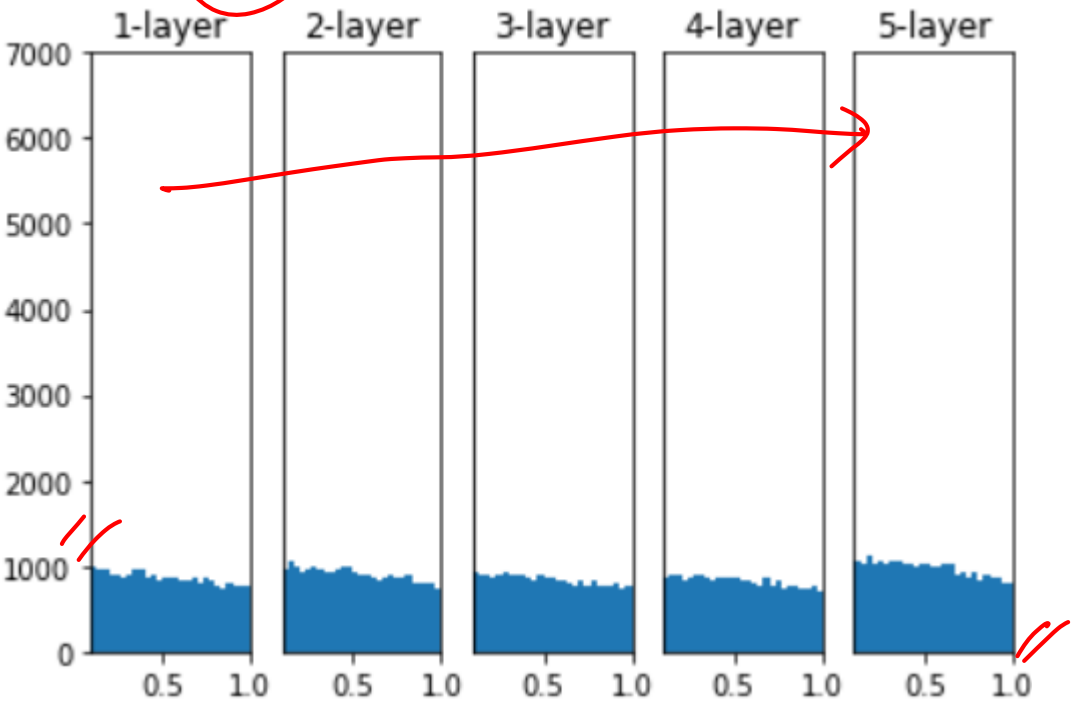
Weight Initialization

- He Initialization

$$W \sim N(0, Var(W))$$

$$Var(W) = \sqrt{\frac{2}{n_{in}}}$$

n_{in} : 이전 hidden layer node



- ReLU 계의 Activation Function 사용

Reference : https://yngie-c.github.io/deep%20learning/2020/03/17/parameter_init/

Thank you..!!!