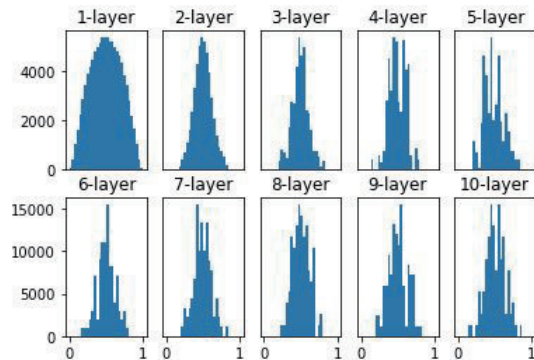
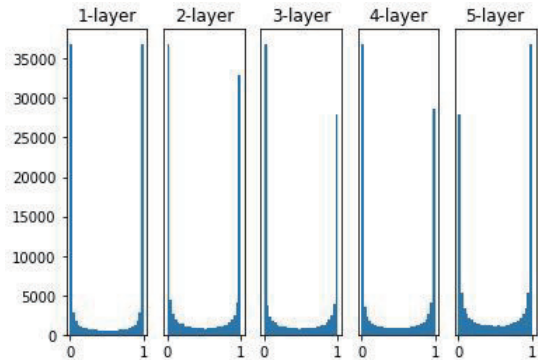


6장 연습문제

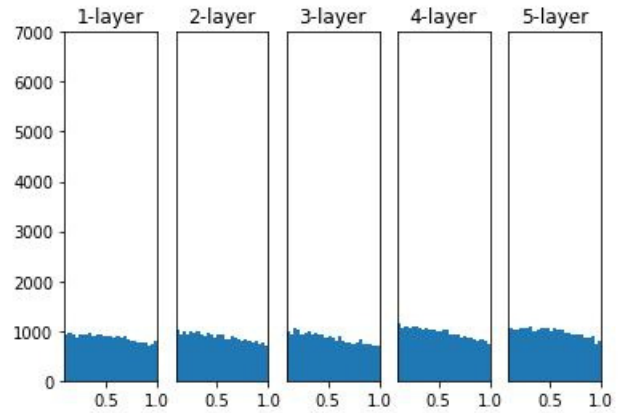
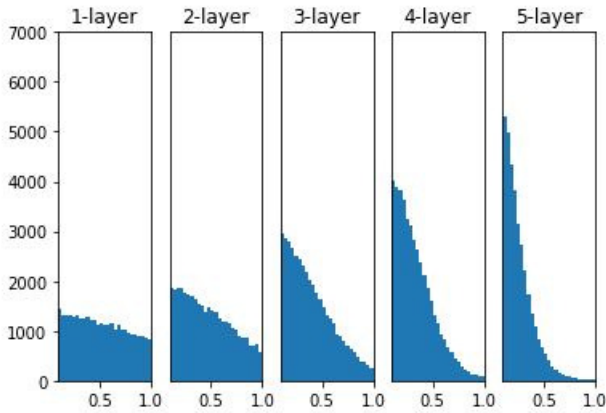
1. 다음 코드를 실행하면 좌측의 히스토그램을 얻는다. 오른쪽 히스토그램이 나오도록 코드를 수정하시오.



```
input_data = np.random.randn(1000, 100)
node_num = 100
hidden_layer_size = 5
activations = {}
x = input_data
for i in range(hidden_layer_size):
    if i != 0:
        x = activations[i-1]
    w = np.random.randn(node_num, node_num) * 1
    a = np.dot(x, w)
    z = sigmoid(a)
    activations[i] = z
for i, a in activations.items():
    plt.subplot(1, len(activations), i+1)
    plt.title(str(i+1) + "-layer")
    if i != 0: plt.yticks([], [])
    plt.hist(a.flatten(), 30, range=(0,1))
plt.show()
```

2. 뉴런의 개수가 각각 784, 400, 300, 200, 100, 10인 5층 신경망을 생각하자. 각 $k = 1, 2, 3, 4, 5$ 에 대하여 $k - 1$ 층에서 k 층으로 가는 Affine 변환의 가중치 행렬을 W_k 라 하자.
 - (i) 활성화 함수(activation function)를 sigmoid함수로 잡고 gradient vanishing이나 표현력 제한이 일어나지 않도록 Xavier 초기값을 따라 초기값을 설정하려 한다. 초기 가중치 행렬 W_1, W_2, W_3, W_4, W_5 를 각각 어떤 분포를 따라 랜덤하게 선택해야 하겠는가?
 - (ii) 활성화 함수(activation function)를 ReLU함수로 잡고 gradient vanishing이나 표현력 제한이 일어나지 않도록 He 초기값을 따라 초기값을 설정하려 한다. 초기 가중치 행렬 W_1, W_2, W_3, W_4, W_5 를 각각 어떤 분포를 따라 랜덤하게 선택해야 하겠는가?

3. 다음 코드를 실행하면 좌측의 히스토그램을 얻는다. 오른쪽 히스토그램이 나오도록 코드를 수정하시오.



```
input_data = np.random.randn(1000, 100)
node_num = 100
hidden_layer_size = 5
activations = {}
x = input_data
for i in range(hidden_layer_size):
    if i != 0:
        x = activations[i-1]
    w = np.random.randn(node_num, node_num) * np.sqrt(1.0 / node_num)
    a = np.dot(x, w)
    z = ReLU(a)
    activations[i] = z
for i, a in activations.items():
    plt.subplot(1, len(activations), i+1)
    plt.title(str(i+1) + "-layer")
    if i != 0: plt.yticks([], [])
    plt.hist(a.flatten(), 30, range=(0,1))
plt.show()
```

4. 활성화 함수를 시그모이드 함수로 사용하는 10층 신경망에서 가중치를 표준정규분포를 따라 초기화하였다. 다음 그림은 신경망에 데이터를 입력했을때 활성화층 출력값들의 히스토그램이다. 이후 다량의 데이터를 입력하여 여러번 학습을 시도하여도 낮은 층으로 내려갈수록 가중치와 편향이 제대로 학습이 되지 않았다. 그 이유를 수학적으로 설명하시오.

