



Topic 6 深度学习

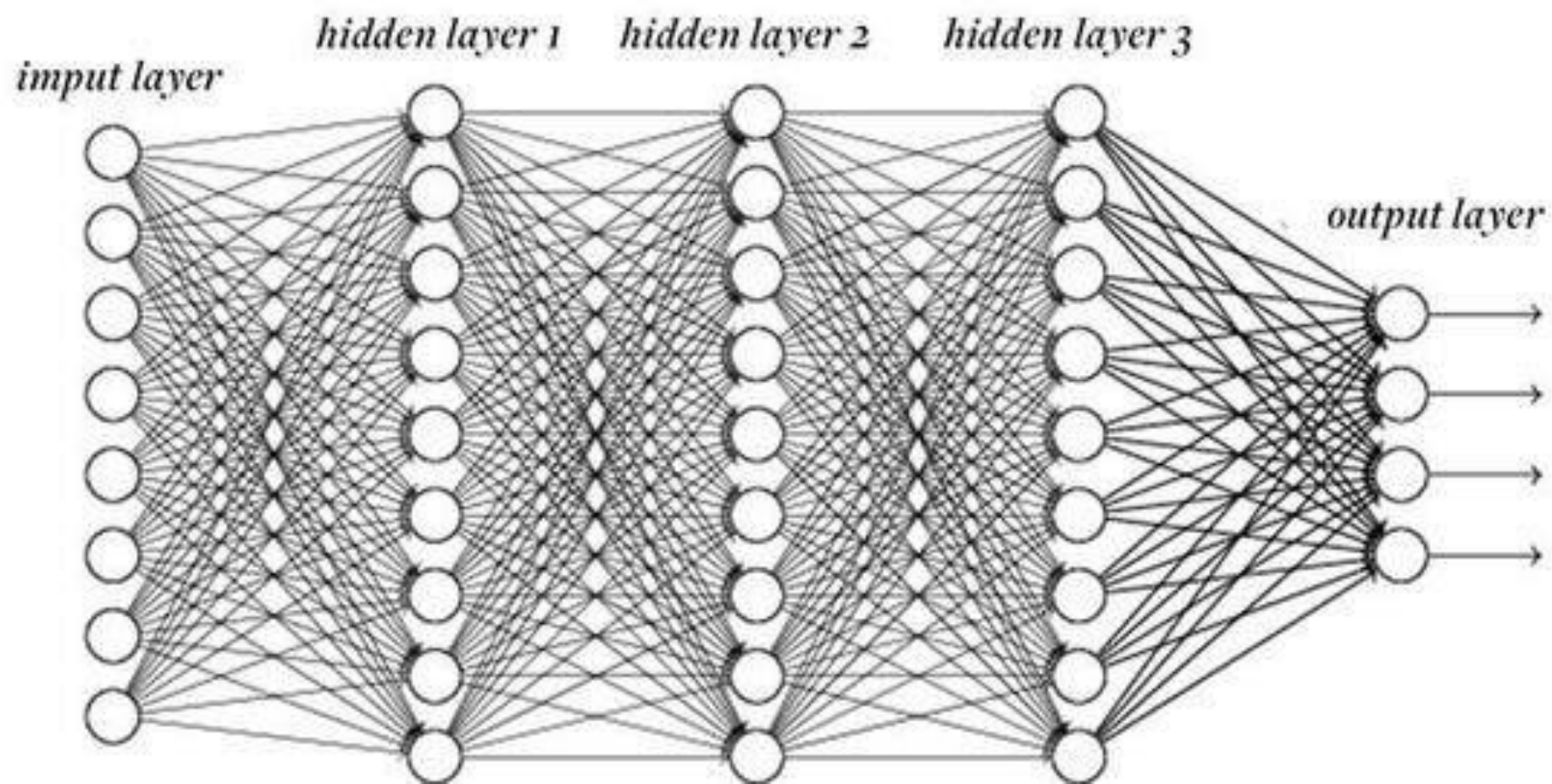
Part 1: 深度神经网络

魏准
浙江大学，信息与电子工程学院
2022

总体内容

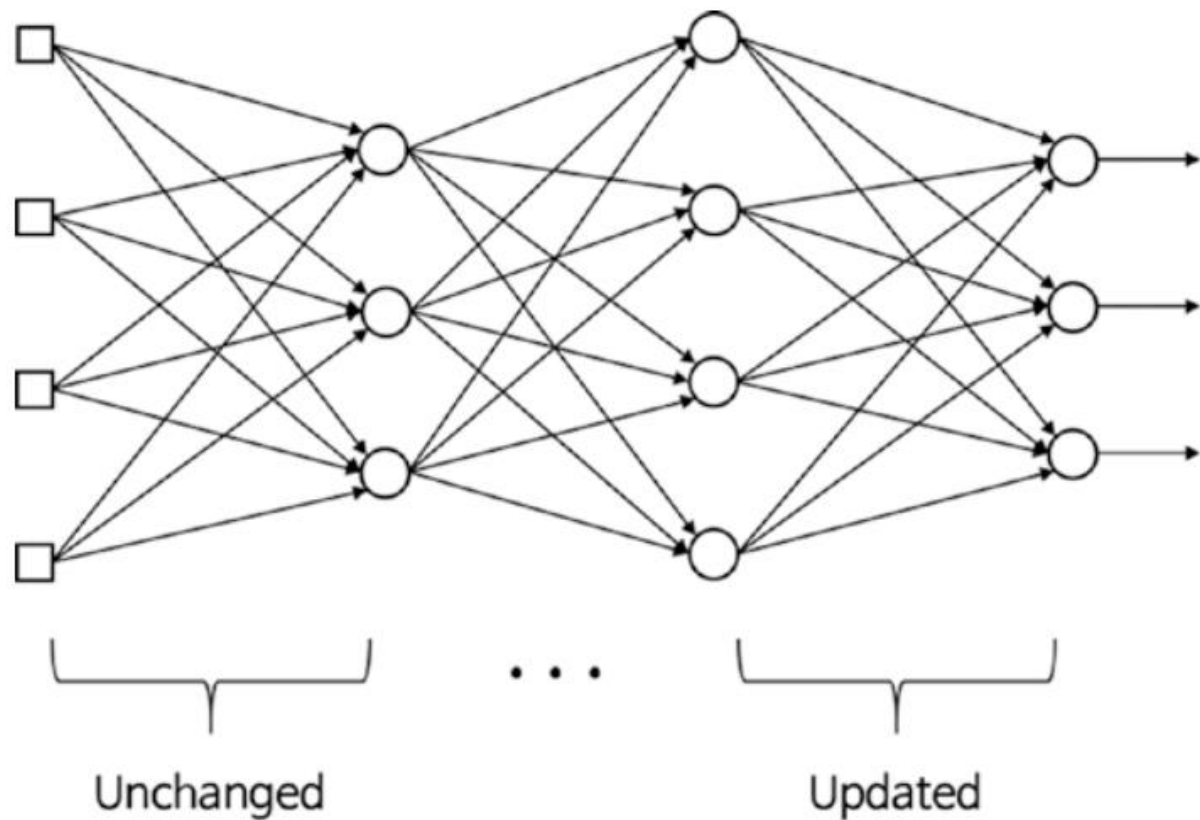
主要知识点

- 梯度消失(Vanishing Gradient);
- 实例：训练一个深度学习网络;
- 实验6-1;
- 过拟合, Dropout;
- 实验6-2;

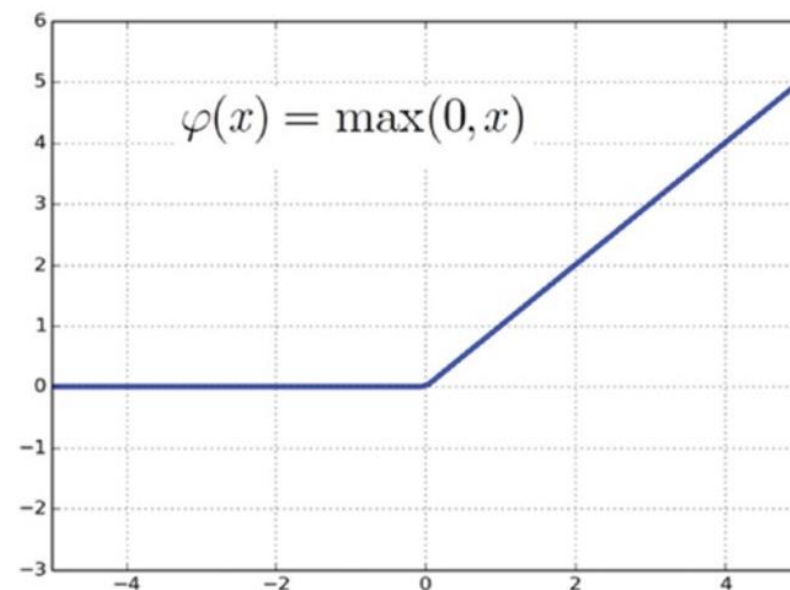


1.1: 梯度消失(Vanishing Gradient)

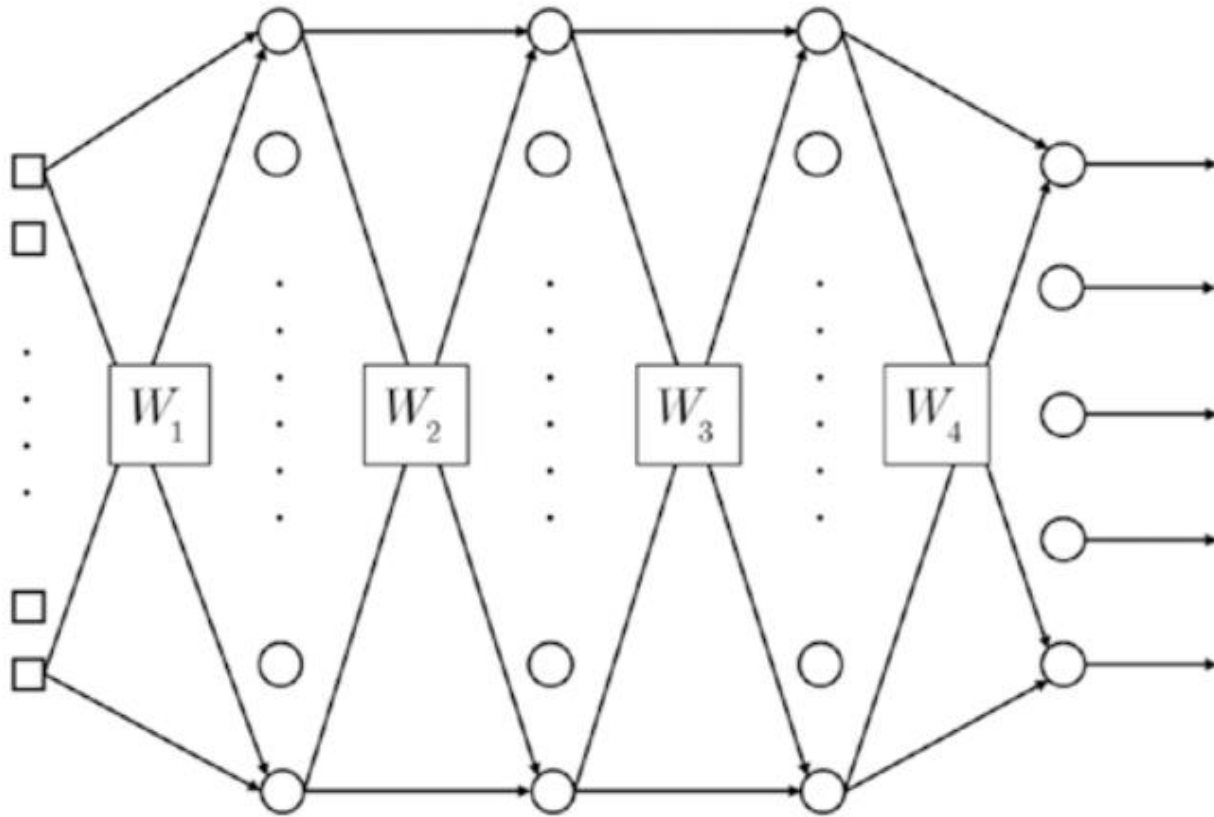
解决办法：ReLU激活函数



$$\varphi(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad \varphi'(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$
$$= \max(0, x)$$



1.2: 实例：训练一个深度学习网络



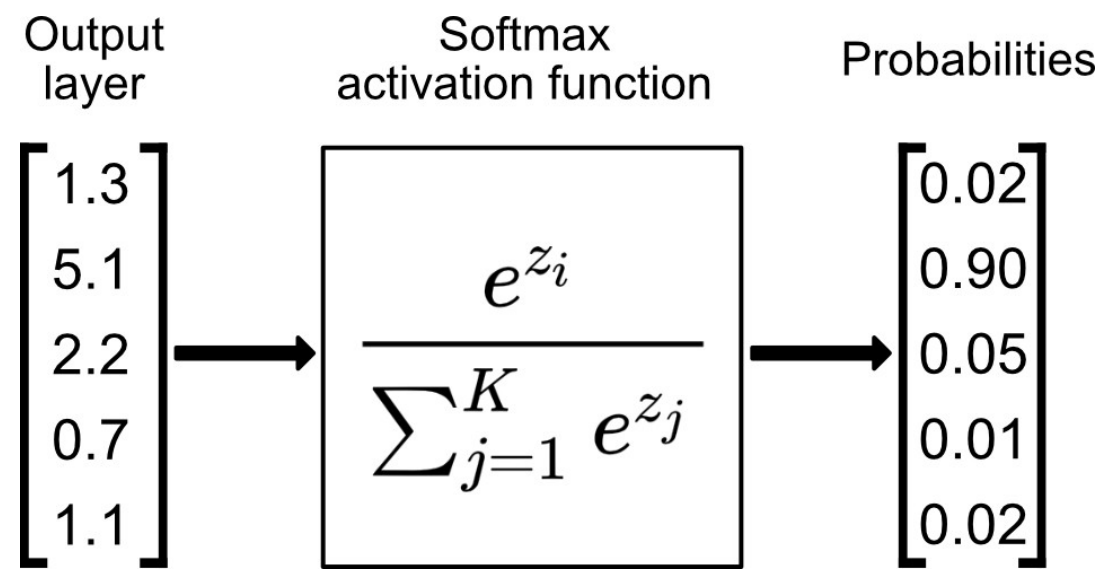
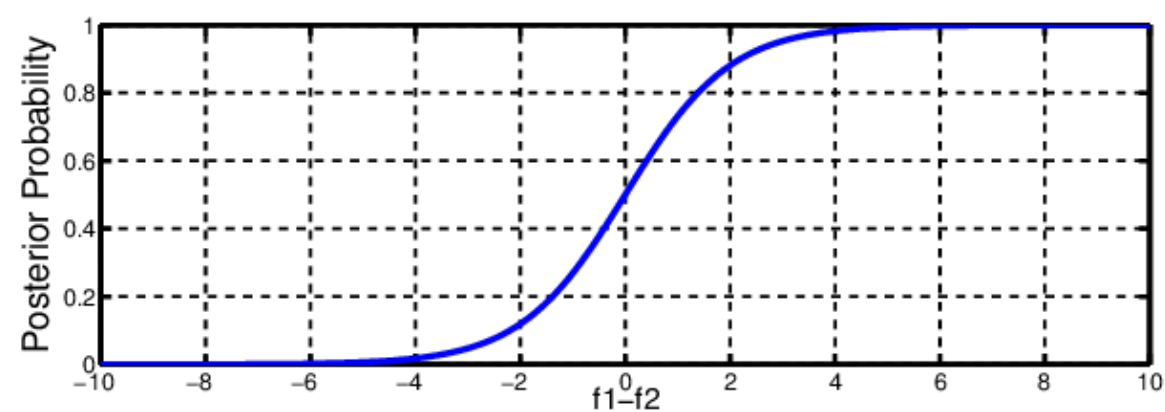
W_1 : 20×25

W_2 : 20×20

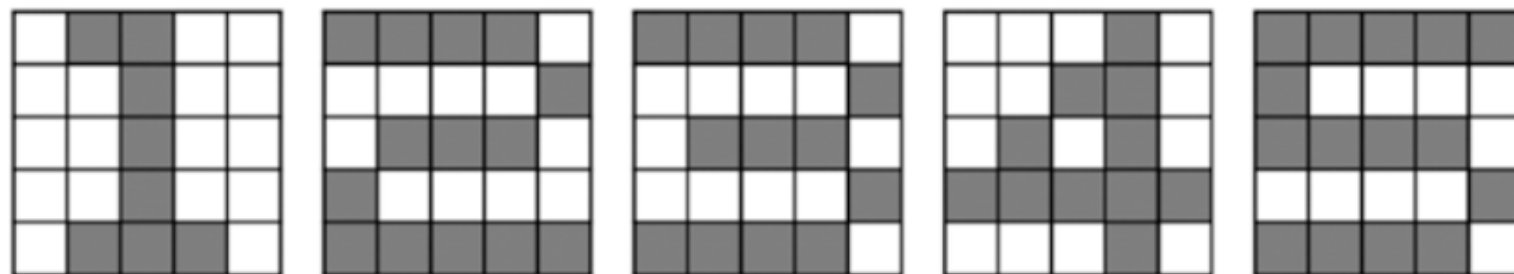
W_3 : 20×20

W_4 : 5×20

激活函数：除最后一层采用softmax外，其他层采用ReLU



训练数据:



$$X(:, :, 1) = \begin{bmatrix} 0 & 1 & 1 & 0 & 0; \\ 0 & 0 & 1 & 0 & 0; \\ 0 & 0 & 1 & 0 & 0; \\ 0 & 0 & 1 & 0 & 0; \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix};$$

$$X(:, :, 2) = \begin{bmatrix} 1 & 1 & 1 & 1 & 0; \\ 0 & 0 & 0 & 0 & 1; \\ 0 & 1 & 1 & 1 & 0; \\ 1 & 0 & 0 & 0 & 0; \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix};$$

$$X(:, :, 3) = \begin{bmatrix} 1 & 1 & 1 & 1 & 0; \\ 0 & 0 & 0 & 0 & 1; \\ 0 & 1 & 1 & 1 & 0; \\ 0 & 0 & 0 & 0 & 1; \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix};$$

$$X(:, :, 4) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0; \\ 0 & 0 & 1 & 1 & 0; \\ 0 & 1 & 0 & 1 & 0; \\ 1 & 1 & 1 & 1 & 1; \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix};$$

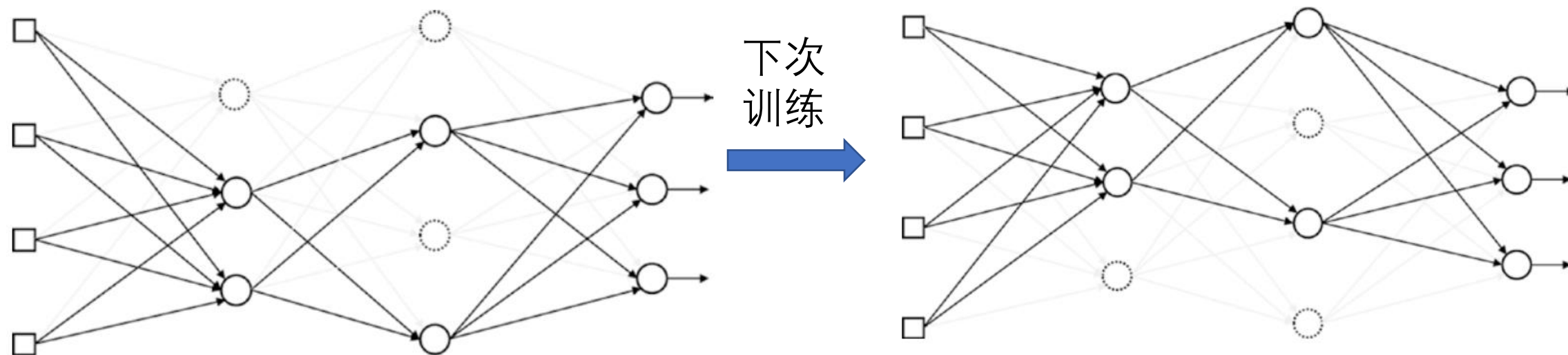
$$X(:, :, 5) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1; \\ 1 & 0 & 0 & 0 & 0; \\ 1 & 1 & 1 & 1 & 0; \\ 0 & 0 & 0 & 0 & 1; \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix};$$

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0; \\ 0 & 1 & 0 & 0 & 0; \\ 0 & 0 & 1 & 0 & 0; \\ 0 & 0 & 0 & 1 & 0; \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix};$$

实验6-1:

通过SGD训练方法、ReLU激活函数（**Softmax 最后一层**）及BP规则，对上述深度神经网络进行训练，并输出训练后的结果。（为简单起见softmax函数的导数直接取1）；

1.3: 过拟合(Overfitting)



解决办法: Dropout (Dropout_rate: p), 剩下的元素乘以 $1/(1-p)$, 仅在训练的时候用!

实验6-2:

加入**dropout**技巧，采用**Sigmoid**激活函数（Softmax 最后一层）及BP规则，对上述深度神经网络进行训练，并输出训练后的结果。

$$y_1 = \begin{bmatrix} 0.5356 \\ 0.9537 \\ 0.5442 \\ 0.0821 \\ 0.3663 \\ 0.8509 \end{bmatrix} \quad ym = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} \quad y_1 * ym = \begin{bmatrix} 1.0712 \\ 1.9075 \\ 0 \\ 0 \\ 0 \\ 1.7017 \end{bmatrix}$$