



1. 反向传播是训练神经网络的常用方法，通过链式求导法则和梯度下降法，从最后一层到第一层，逐渐实现神经网络的参数训练

$$2. (1) \begin{vmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{vmatrix} \quad (2) \begin{vmatrix} 3 & 2 & 4 \\ 2 & 3 & 2 \\ 3 & 2 & 2 \end{vmatrix} \quad (3) \begin{vmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 2 & 2 & 3 \end{vmatrix}$$

$$3. (1) 2 \times 2 \begin{vmatrix} 3 & 4 & 3 \\ 6 & 5 & 5 \\ 4 & 5 & 7 \end{vmatrix} \quad 3 \times 3 \begin{vmatrix} 3 & 5 \\ 6 & 7 \end{vmatrix}$$

$$(2) \begin{vmatrix} 3 & 3 & 4 & 4 & 3 \\ 3 & 3 & 6 & 4 & 5 \\ 6 & 3 & 5 & 5 & 5 \\ 6 & 5 & 5 & 5 & 3 \\ 4 & 5 & 5 & 6 & 7 \end{vmatrix}$$

$$4. (1) L = -\frac{1}{3} \log 0.7$$

$$L_1 = \frac{1}{N} \sum_{i=0}^{N-1} |y_i - \hat{y}_i|$$

$$= \frac{1}{3} (0.1 + 0.2 + 0.3) = 0.2$$

$$(2) L = -\frac{1}{3} \log 0.6$$

$$L_1 = \frac{1}{3} (0.4 + 0.2 + 0.2) = \frac{4}{15}$$

5. inception 模块可构造更稀疏的 CNN 结构，使用 1×1 的卷积核来进行降维从而极大地降低了卷积参数的数量，使用不同大小的卷积核来捕获不同大小的感受野

