

机器学习 第5章 作业

5.1

使用线性函数作为激活函数时，无论是在隐藏层还是在输出层（无论传递几层），其单元值（在使用激活函数之前）都还是输入 x 的线性组合，这个时候的神经网络其实等价于逻辑回归的，若输出层也使用线性函数作为激活函数，那么就等价于线性回归。

5.2

使用Sigmoid激活函数，每个神经元几乎和对率回归相同，只不过对率回归在 $\text{sigmoid}(x) > 0.5$ 时输出为1，而神经元直接输出 $\text{sigmoid}(x)$

5.3

$\Delta v_i h = -\eta \frac{\partial E_k}{\partial v_i h}$ ，因 $v_i h$ 只在计算 b_h 时用上，所以 $\frac{\partial E_k}{\partial v_i h} = \frac{\partial E_k}{\partial b_h} \frac{\partial b_h}{\partial v_i h}$ ，其中
 $\frac{\partial b_h}{\partial v_i h} = \frac{\partial b_h}{\partial a_h} \frac{\partial a_h}{\partial v_i h} = \frac{\partial b_h}{\partial a_h} x_i$ ，所以 $\frac{\partial E_k}{\partial v_i h} = \frac{\partial E_k}{\partial b_h} \frac{\partial b_h}{\partial a_h} x_i = -e_h x_i$ ，即得原书中5.13式

5.4

学习率太高会导致误差函数来回震荡，无法收敛。

学习率太低则会收敛太慢，影响训练效率。