

## 第八讲 课后作业

---

1. (50) 考虑两层网络函数

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma \left( \sum_{j=1}^M w_{kj}^{(2)} h \left( \sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right)$$

其中，隐含单元非线性激活函数是 logistic sigmoid 函数

$$\sigma(a) = \{1 + \exp(-a)\}^{-1}$$

证明：存在一个准确计算相同函数的等价网络，但该等价网络隐含单元激活函数是  $\tanh(a)$  函数

$$\tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}$$

2. (50) 考虑多目标变量回归问题，假设目标标量对输入矢量  $\mathbf{x}$  的条件分布满足高斯分布

$$p(\mathbf{t}|\mathbf{x}, \mathbf{w}) = \mathcal{N}(\mathbf{t}|\mathbf{y}(\mathbf{x}, \mathbf{w}), \Sigma)$$

其中  $\mathbf{y}(\mathbf{x}, \mathbf{w})$  是网络在输入矢量  $\mathbf{x}$  和权矢量  $\mathbf{w}$  时的输出， $\Sigma$  是目标变量上高斯噪声的协方差矩阵。已知独立观测量集合  $\{\mathbf{x}, \mathbf{t}\}$ ，如果假设  $\Sigma$  是不变和已知的，那么为了找到  $\mathbf{w}$  的最大似然解必须最小化的误差函数是什么？现在假设  $\Sigma$  也是由数据决定的，写出  $\Sigma$  最大似然解的表达式。注意参数  $\mathbf{w}$  和  $\Sigma$  的优化是耦合的。