

1 Introduction

2017年12月8日 星期五 下午12:54

1.1 例子

- 一个简单的应用广泛的误差函数：

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$

是每个数据点的预测值与目标值差的平方和

- 实现泛化性的选择（生成数据情况）
 - 考虑一个额外的测试集，这个测试集的数据点的生成与训练集的生成方式完全相同，但是在目标值中包含的随机噪声的值不同。对于不同的训练结果，计算训练集和测试集的误差经常使用根均方误差（RMS）

$$E_{RMS} = \sqrt{2E(\mathbf{w}^*)/N}$$

平方根确保了与目标变量 t 使用相同的规模 and 单位进行度量

- 正则化。给误差函数增加一个惩罚项，使得系数不会达到很大的值
eg：

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

使得系数不会达到很大的值。在多项式函数拟合的情况下，通常系数 ω_0 会从正则化项中省略，因为包含它会使结果依赖于目标变量原点的选择。 ω_0 也可以被包含在正则化项中，但是必须有自己的正则化系数。

【注意观察数据的特点，对于结果要有深刻直观的理解，了解取得该结果的原因】

1.2 概率论

1.3 模型选择

1.4 维度灾难

1.5 决策理论

1.6 信息论