

曲线拟合过程中的欠定过定问题

emacsun

目录

1 问题	1
2 抽象	1
3 总结	2

1 问题

曲线拟合过程中，我们有目标函数

$$y(x, \mathbf{w}) = w_0 + w_1x + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j \quad (1.1)$$

误差函数：

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 \quad (1.2)$$

对于这个问题，当训练数据集合小于 w_j 的个数时，我们可以认为对(1.2)求导，求得 w_j 的过程为欠定方程的求解过程；当训练数据集合大于 w_j 的个数时，对(1.2)求导，求得 w_j 的过程是过定方程的求解过程。

2 抽象

不管是欠定还是过定，最小化 (1.2)的解 $\mathbf{w} = \{w_j\}$ 都是一般线性方程：

$$\sum_{j=0}^M A_{ij}w_j = T_i$$

的解。其中 $A_{ij} = \sum_{n=1}^N (x_n)^{i+j}$ ， $T_i = \sum_{n=1}^N (x_n)^i t_n$



我们可以把(1.1)带入(1.2), 得:

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \left\{ \sum_{j=0}^M w_j x_n^j - t_n \right\}^2 \quad (2.1)$$

(2.2)

对于(2.1), 我们求 w_i , 则有:

$$\frac{dx}{dw_i} = \sum_{n=1}^N \left\{ \sum_{j=0}^M w_j x_n^j - t_n \right\} x_n^i \quad (2.3)$$

$$= \sum_{n=1}^N \sum_{j=0}^M w_j x_n^j x_n^i - \sum_{n=1}^N t_n x_n^i \quad (2.4)$$

令(2.3)等于零, 我们得到关于 $w_j, j = 0, \dots, M$ 的方程:

$$\sum_{n=1}^N \sum_{j=0}^M w_j x_n^j x_n^i = \sum_{n=1}^N t_n x_n^i \quad (2.5)$$

整理得到,

$$\sum_{j=0}^M w_j \sum_{n=1}^N x_n^{i+j} = \sum_{n=1}^N t_n x_n^i \quad (2.6)$$

令 $A_{ij} = \sum_{n=1}^N x_n^{i+j}$, $T_i = \sum_{n=1}^N t_n x_n^i$, 所以我们得到了:

$$\sum_{j=0}^M A_{ij} w_j = T_i$$

3 总结

我们看到, 对于(2.6), 可以扩展为 $M + 1$ 个方程, 即:

$$\begin{aligned} \sum_{j=0}^M A_{0j} w_j &= T_0 \\ \sum_{j=0}^M A_{1j} w_j &= T_1 \\ &\vdots \\ \sum_{j=0}^M A_{Mj} w_j &= T_M \end{aligned}$$

所以我们有 $M + 1$ 个方程 $M + 1$ 个未知数 $w_j, j = 0, \dots, M$