基于最小二乘法的曲线拟合

一、前言

曲线拟合就是求一条曲线,使数据点均在离此曲线的上方或下方不远处,它既能反映数据的总体分布,又不至于出现局部较大的波动,能反映被逼近函数的特性,使求得的逼近函数与已知函数从总体上来说其偏差按某种方法度量达到最小。如图 1 所示。

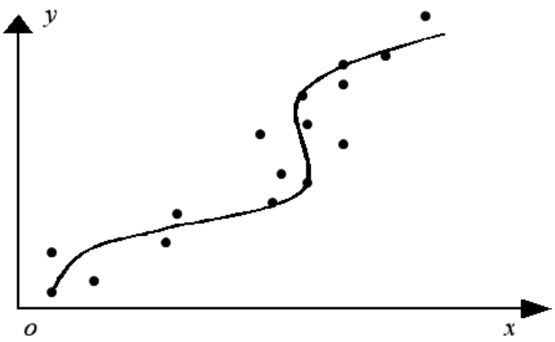

图 1 曲线拟合示意图

二、最小二乘法原理

假设有 m 组数据且 m>n，设近似方程为：

$$y = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_n\varphi_n(x) \tag{1}$$

$$\delta(a_0, a_1, \cdots, a_n) = \sum_{i=1}^{m}(f(x_i) - y_i)^2 = \min \tag{2}$$

对函数δ求偏导，并使其为零，如公式（3）所示。

$$\frac{\partial\delta}{\partial a_j} = 0(j = 0,1,\cdots,n) \tag{3}$$

上式即可得：

$$\sum_{i=1}^{m} 2(\sum_{k=0}^{n} a_k\varphi_k(x_i) - y_i))\varphi_j(x_i) = 0 \tag{4}$$

引入记号：

$$(\varphi_j, \varphi_k) = \sum_{i=1}^{m}\varphi_j(x_i)\varphi_k(x_i) \tag{5}$$

$$(\varphi_j, f) = \sum_{i=1}^{m}\varphi_j(x_i)f(x_j) \tag{6}$$

综上则有：

$$\sum_{k=0}^{n} a_k(\varphi_j,\varphi_k)=(\varphi_j,f) \qquad （7）$$

在$\{\varphi_0,\varphi_1,\cdots,\varphi_n\}$满足一定条件时，可以证明方程（7）有唯一解，且这个解就是我们需要找的最小二乘拟合公式。

三、编程原理及结果展示

基于 C 语言，设计了一个 50 组数据内的曲线拟合程序，程序流场图如图 2 所示。

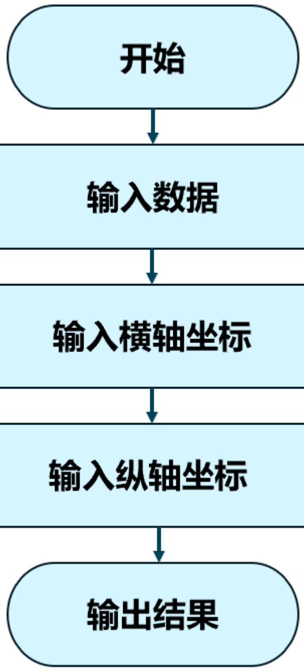

图 2 程序流程图

以网上找的一组数据为例，在化学反应中，生成物浓度 y（%）与时间 t（min）的数据见表 1，建立这组数据的拟合公式。

表 1 生成物浓度与时间数据对应表

t	1	2	3	4	5	6	7	8
y	4.00	6.40	8.00	8.80	9.22	9.50	9.70	9.86
t	9	10	11	12	13	14	15	16
y	10.00	10.20	10.32	10.42	10.50	10.55	10.58	10.60

通过程序运算最终求得如公式（8）所示。

$$y = -0.044t^2 + 1.066t + 4.387 \quad (8)$$

将函数与表 1 数据画图，如图 3 所示。

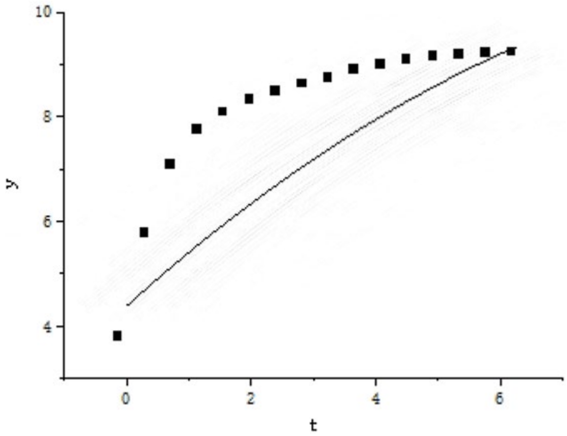

图 3 拟合效果图