曲线拟合数值计算方法论文

一． 简介

1.1 最小二乘原理

最小二乘法是一种数学优化技术。它通过最小化误差的平方和寻找数据的最佳函数匹配。利用最小二乘法可以简便地求得未知的数据，并使得这些求得的数据与实际数据之间误差的平方和为最小。最小二乘法还可用于曲线拟合。其他一些优化问题也可通过最小化能量或最大化熵用最小二乘法来表达。

给定合适的拟合函数 $\varphi(x)$ ，常用的拟合函数为多项式，则多项式拟合函数的形式为 $\varphi(x)=a_k x^k + a_{k-1}x^{k-1}+\cdots+a_1 x+a_0$ ，测量 n 组 $\left(x_i, y_i\right)$ ，令

$$R(a_k)=\sum_{i=1}^{n}\left[\varphi(x_i)-y_i\right]^2$$

公式 1.1

求适当的 a_k 使得 $R(a_k)$ 最小的原理成为最小二乘原理。

1.2 多项式拟合函数参数的确定

给定数据 $\left(x_i, y_i\right), j=1,2,\cdots,n$ ，拟合函数的形式为

$$p(x)=a_0\varphi_0(x)+a_1\varphi_1(x)+\cdots+a_m\varphi_m(x)$$

公式 1.2

其中 $\left\{\varphi_k(x)\right\}_{k=0}^{m}$ 已知，且线性无关。求系数 $a_0^*, a_1^*, \cdots, a_m^*$ 使得

$$\varphi(a_0,a_1,\cdots,a_m)=\sum_{j=1}^{n}\left[p(x_j)-y_j\right]^2=\sum_{j=1}^{n}\left[\sum_{k=0}^{m}a_k\varphi_k(x_j)-y_j\right]^2$$

公式 1.3

取最小值。称

$$p^*(x)=\sum_{k=0}^{m}a_k^*\varphi_k(x)$$

公式 1.4

为拟合函数。

若$\varphi_k(x)=x^k(k=0,1,\cdots,m)$ ，则称式(1.4)为 m 次最小二乘拟合多项式。

由公式 1.3 可以看出，$\varphi(a_0,a_1,\cdots,a_m)$ 为$a_0,a_1,\cdots,a_m$ 的（m+1）元二次多项式，可以用多元函数求极值的方法求最小值。φ 对a_k 求偏导，得

$$\frac{\partial\varphi}{\partial a_k}=2\sum_{j=1}^{n}(\sum_{i=0}^{m}a_i^*\varphi_i(x_j)-y_j)\varphi_k(x_j)(k=0,1,\cdots,m)$$

即

$$\sum_{i=0}^{m}(\sum_{j=1}^{n}\varphi_i(x_j)\varphi_k(x_j))a_i^*=\sum_{j=1}^{n}y_j\varphi_k(x_j)(k=0,1,\cdots,m)$$

公式 1.5

引入内积符号，公式 1.5 可写为

$$\begin{bmatrix}(\varphi_0,\varphi_0) & (\varphi_0,\varphi_1) & \cdots & (\varphi_0,\varphi_m)\\ (\varphi_1,\varphi_0) & (\varphi_1,\varphi_1) & \cdots & (\varphi_1,\varphi_m)\\ \vdots & \vdots & \ddots & \vdots\\ (\varphi_m,\varphi_0) & (\varphi_m,\varphi_1) & \cdots & (\varphi_m,\varphi_m)\end{bmatrix}\begin{bmatrix}a_0^*\\ a_1^*\\ \vdots\\ a_m^*\end{bmatrix}=\begin{bmatrix}(y,\varphi_0)\\ (y,\varphi_1)\\ \vdots\\ (y,\varphi_m)\end{bmatrix}$$

公式 1.6

其中

$$\varphi_k=\begin{bmatrix}\varphi_0(x_1)\\ \varphi_0(x_2)\\ \vdots\\ \varphi_0(x_n)\end{bmatrix}(k=0,1,\cdots,m),y=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}$$

由公式 1.6 即可求出 m 次二乘拟合多项式的 m 个系数。

二．问题背景

如何确定视频的拍摄地点和拍摄日期是视频数据分析的重要方面，太阳影子定位技术就是通过分析视频中物体的太阳影子变化，确定视频拍摄的地点和日期的一种方法。

对于确定视频拍摄地点的经度，可根据视频拍摄时间与影子相对物体的长度的相对大小的变化进行确定，而根据离散的数据拟合出影子-时间曲线是其中关键一步。

三．数据

时间/小时	影子长度/m
14.7	1.149626
14.75	1.182199
14.8	1.215297
14.85	1.249051
14.9	1.283195
14.95	1.317993
15.00	1.353364
15.05	1.389387
15.10	1.426153
15.15	1.463400
15.2	1.501482
15.25	1.540232

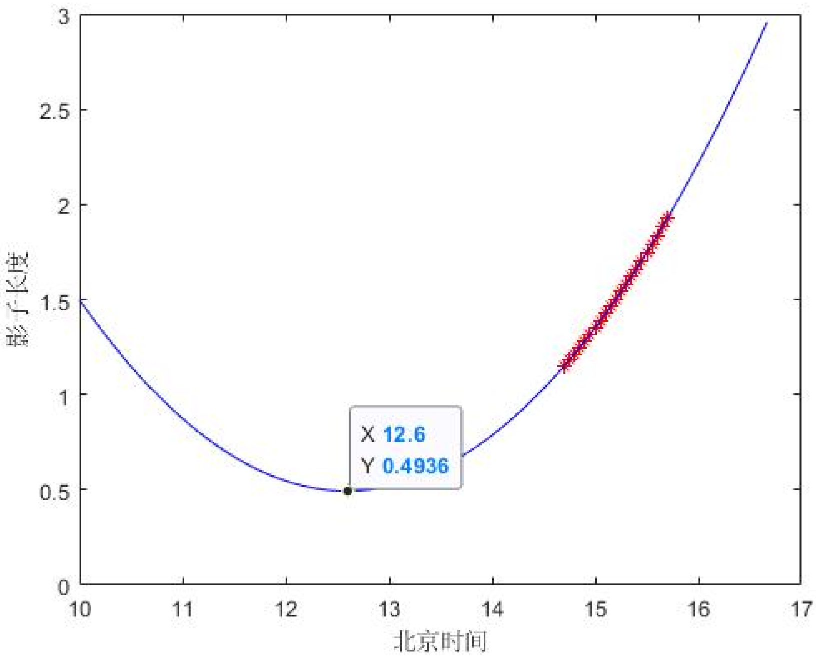
15.3	1.579853
15.35	1.620145
15.4	1.661271
15.45	1.703291
15.5	1.746206
15.55	1.790051
15.6	1.835014
15.65	1.880875
15.7	1.927918

表 1 时间、影长数据

表 1 为某地测得不同时刻测得某一物体的长度及其对应的时间。

四．曲线拟合

根据查阅相关资料，影子长度与时间为二次函数关系，所以用二次最小二乘拟合。代码见附录一。



图表 1 影子长度-北京时间拟合曲线

如图表一所示，红色点为表一内测量数据，用二次最小拟合，得到二次拟合函数

$$\varphi(x) = 0.1489x^2 - 3.7519x + 24.1275$$

然后求出 10 点到 17 点的影子拟合长度，如上图蓝线所示。

五．总结

最小二乘曲线拟合在很多问题中都有应用，是解决一些问题的关键一步，是很重要的一种方法。

附录

附录一:

```
clc,clear
time=[14.7000000000000,14.7500000000000,14.8000000000000,14.8500000000000,14.9000000000000,14.9500000000000,15,15.0500000000000,15.1000000000000,15.1500000000000,15.2000000000000,15.2500000000000,15.3000000000000,15.3500000000000,15.4000000000000,15.4500000000000,15.5000000000000,15.5500000000000,15.6000000000000,15.6500000000000,15.7000000000000]';
length=[1.14962582608430;1.18219897648408;1.21529695548043;1.24905105179892;1.28319533976710;1.31799314869236;1.35336404932302;1.38938709149035;1.42615285646385;1.46339985308186;1.50148162159915;1.54023181696782;1.57985331597589;1.62014451515906;1.66127061311515;1.70329063286334;1.74620590996595;1.79005091547699;1.83501427242406;1.88087500116302;1.92791844744533];
p=polyfit(time,length,2);
x=(600:0.01:1000)/60;
y=polyval(p,x);
time1=x(y==min(y));
plot(time,length,'r*');
hold on
plot(x,y,'b-');
hold on
plot(time1,min(y),'g*')
xlabel('北京时间')
ylabel('影子长度')
```