

• 开发与应用 •

# 基于抛物线插值的正弦波拟合算法

李耀彬<sup>1</sup>, 曾祥斌<sup>2</sup>, 沈铨武<sup>1</sup>

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国人民解放军九五九七二部队, 甘肃 酒泉 735018)

**摘要:**为了合理利用原始采集数据并提高正弦波形数据拟合重构的精度,在最小二乘法原理的基础上,提出了抛物线拟合细分最小二乘法拟合正弦波形的方法。根据所采集数据点的特点,将数据进行分段抛物线插值细分。按照数据间抛物线变化趋势增加拟合数据量,然后对细分后的数据进行最小二乘拟合得到正弦曲线的参数。通过软件算法仿真并进行了实验数据处理,结果表明,该算法合理利用了所采集的数据,可使正弦波形参数的相位估计重复精度从最小二乘的 $-4.555 \times 10^{-4}$ 提高到 $1.367 \times 10^{-4}$ 。

**关键词:**正弦; 最小二乘; 抛物线插值; 算法; 相位

**中图分类号:** TN911.72 **文献标识码:** A **文章编号:** 1000-7024 (2009) 11-2793-03

## Fitting algorithm of sine wave based on interpolation of parabola

LI Yao-bin<sup>1</sup>, ZENG Xiang-bin<sup>2</sup>, SHEN Cheng-wu<sup>1</sup>

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China; 2. 95972 Unit, PLA Army, Jiuquan 735018, China)

**Abstract:** In order to make rationally use of the original data and improve the data fitting precision of the sine waveform, the method of the parabola and least square sine waveform fitting is presented based on the least square method. According to the character of the collected data, the parabola interpolation subdivision is carried out based the continuous serial points divided by some demand. According to the change trend of data, the least square fitting is carried out after the parabola subdivision. The experiment data are processed based on the software program after emulating. The result shows that the algorithm reasonably makes use of the gathered data and enhances effectively the precision of the parameter of the sine waveform after the curve fitting. And comparing by the only least square method, the precision of phase is improved from  $-4.555 \times 10^{-4}$  to  $1.367 \times 10^{-4}$  for data in the actual experiment.

**Key words:** sine wave; least square; parabola interpolation; algorithm; phase

## 0 引言

给定一组有序数据点,这些点可以是实际生产中测出的,也可以是设计员给出,要求一条曲线通过这些数据点,称为对这些数据点进行插值,所构造的曲线称为插值曲线。在某些情况下,测量所得或设计员给出的数据本身就很粗糙,要求构造一条曲线通过这些数据点就没有什么意义。更合理的提法应是构造一条曲线使之在某种意义下最为接近给定的数据点,称之为对这些数据点进行逼近,所构造的曲线称为逼近曲线<sup>[1]</sup>。

在工程设计中常常需要用曲线按一定要求逼近一系列实测的有序数据点,而这些离散的有序数据点由于各种因素的影响存在误差,因此反映这些实验数据的曲线或曲面不必通过每一点,常采用拟合方法对这些离散点进行逼近,其中最小二乘拟合是较常用的方法<sup>[1-2]</sup>。

在实际应用中,有些场合需要尽可能利用原始数据点,而且对拟合计算精度要求也较高,这时单单使用最小二乘法就不能满足数据处理的要求了。本文就基于这种情况,提出了通过抛物线插值增加数据量后采用最小二乘正弦曲线拟合的方法。该方法先将序列数据按一定规则进行抛物线拟合,然后根据抛物线拟合特点,按照处理要求进行细化取值,最后对细化的取值进行最小二乘拟合,求出正弦曲线参数。

## 1 最小二乘法拟合原理

设有  $n+1$  个相异点  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$  求一个  $m(m < n)$  次多项式  $P(x)$ , 使它在  $x_i$  点上取值尽量接近  $y_i (i=0, 1, \dots, n)$ 。

设所求多项式为

$$P(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m = \sum_{j=0}^m a_j x^j \quad (m < n) \quad (1)$$

若令

收稿日期: 2008-06-27; 修订日期: 2009-02-03。

基金项目: 中国科学院知识创新工程领域前沿基金项目。

作者简介: 李耀彬 (1969 -), 男, 辽宁锦州人, 研究员, 研究方向为光机总体设计; 曾祥斌 (1971 -), 男, 贵州贵定人, 工程师, 研究方向为靶场测量测控; 沈铨武 (1978 -), 男, 吉林松原人, 硕士, 助理研究员, 研究方向为光电测量。E-mail: liyaobin7025@163.com

$$\begin{cases} a_0 + a_1x_0 + a_2x_0^2 + \dots + a_mx_0^m - y_0 = R_0 \\ a_0 + a_1x_1 + a_2x_1^2 + \dots + a_mx_1^m - y_1 = R_1 \\ \dots \\ a_0 + a_1x_n + a_2x_n^2 + \dots + a_mx_n^m - y_n = R_n \end{cases} \quad (2)$$

亦即

$$\sum_{j=0}^m a_j x_i^j - y_i = R_i \quad (i=0,1,\dots,n) \quad (3)$$

由于曲线  $P(x)$  不一定会通过所有点  $P(x_i, y_i)$ , 故  $R_i (i=0,1,\dots,n)$  不会全为零值。为了使  $R_i (i=0,1,\dots,n)$  达到最小, 还能保证所求的偏差和不相互抵消, 便于分析, 使用偏差平方和的最小值作为讨论分析条件, 即

$$\sigma = \sum_{i=0}^n R_i^2 = \sum_{i=0}^n \left( \sum_{j=0}^m a_j x_i^j - y_i \right)^2 = \varphi(a_0, a_1, \dots, a_m) \quad (4)$$

当  $\sigma$  达到最小值时, 即保证了所拟合多项式的偏差的绝对值也最小。这种利用偏差的平方和最小为条件选择  $\varphi(a_0, a_1, \dots, a_m)$  的方法即为最小二乘法<sup>[2]</sup>。

由微分学知道, 使  $\varphi(a_0, a_1, \dots, a_m)$  达到最小的  $a_0, a_1, \dots, a_m$  满足必要条件

$$\frac{\partial M}{\partial a_k} = -2 \sum_{i=0}^n [y_i - P(x_i)] x_i^k = 0 \quad (k=0,1,\dots,m) \quad (5)$$

或写成如下正规方程

$$\begin{cases} a_0(n+1) + a_1 \sum_{i=0}^n x_i + a_2 \sum_{i=0}^n x_i^2 + \dots + a_m \sum_{i=0}^n x_i^m = \sum_{i=0}^n y_i \\ a_0 \sum_{i=0}^n x_i + a_1 \sum_{i=0}^n x_i^2 + a_2 \sum_{i=0}^n x_i^3 + \dots + a_m \sum_{i=0}^n x_i^{m+1} = \sum_{i=0}^n x_i y_i \\ \dots \\ a_0 \sum_{i=0}^n x_i^m + a_1 \sum_{i=0}^n x_i^{m+1} + a_2 \sum_{i=0}^n x_i^{m+2} + \dots + a_m \sum_{i=0}^n x_i^{m+m} = \sum_{i=0}^n x_i^m y_i \end{cases} \quad (6)$$

方程组 (6) 的解是惟一存在的; 方程组 (6) 的解  $a_j$  可使得  $\varphi(a_0, a_1, \dots, a_m)$  达到最小值。由方程组 (6) 解得  $a_j$ , 可以得到拟合多项式  $P(x)$ , 把  $x_i$  代回  $P(x)$  即可得到拟合值。

## 2 正弦曲线拟合

本文涉及的正弦曲线拟合步骤主要分为两部分: 先将数据点依次按照相邻 3 点进行抛物线拟合, 拟合后对数据点间进行插值细分, 再将细分后的数据进行最小二乘拟合求得正弦曲线的参数。

### 2.1 抛物线插值

抛物线拟合就是利用抛物线作为基本曲线, 通过一定的数学方法, 把一组离散的数据点用一条复合的二次曲线光滑地连接起来。

抛物线拟合的基本原理是通过 3 点可以确定惟一的一条抛物线<sup>[6]</sup>。若设抛物线方程为

$$y = ax^2 + bx + c \quad (8)$$

表达式中的 3 系数  $a$ 、 $b$  和  $c$  是未知的, 所以只要确定该 3 系数的值, 则抛物线就能惟一确定。假定给定的 3 样本点的坐标为  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$  和  $P_3(x_3, y_3)$ , 则由 3 点  $P_1$ 、 $P_2$ 、 $P_3$  求出抛物线方程  $y = ax^2 + bx + c$  中  $a$ 、 $b$ 、 $c$  的值, 分别为

$$a = \frac{(x_1 - x_3)(y_1 - y_2) - (x_1 - x_2)(y_1 - y_3)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)} \quad (9)$$

$$b = \frac{(y_1 - y_2) - a(x_1^2 - x_2^2)}{(x_1 - x_2)} \quad (10)$$

$$c = y_1 - ax_1^2 - bx_1 \quad (11)$$

即可确定过  $P_1$ 、 $P_2$ 、 $P_3$  这 3 点的抛物线。

给定一组样本点  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$ ,  $\dots$ ,  $P_n(x_n, y_n)$ , 按顺序每次取 3 点进行最小二乘抛物线拟合, 每相邻 3 点可以确定一段抛物线。保留  $P_1$  到  $P_2$  这一段抛物线, 然后再向后移动一个点, 再以  $P_2$ 、 $P_3$ 、 $P_4$  这 3 点形成一条新的抛物线, 保留  $P_2$  到  $P_3$  这一段, 以此类推, 我们可以得到后续的各段曲线。但全部  $n$  个数据点, 总共只能产生  $n-1$  个曲线段。若每次取前两点画所拟合的抛物线, 则还有最后两个点之间的曲线无法确定; 若每次取最后两点画出所拟合的抛物线, 则最初的 3 点所拟合的抛物线中前两点之间的线段没有确定。

两端曲线无法确定的原因在于缺乏所要求的连续相邻 3 点来拟合出所需的抛物线。解决问题的方法就是想办法补足 3 点, 或者在数据条件及精度要求允许的情况下, 直接利用开始及其最后所拟合的抛物线确定所缺的曲线段<sup>[7]</sup>。所谓补足 3 点, 即在  $P_1$  前补一点  $P_0$ , 或在  $P_n$  后补一点  $P_{n+1}$ , 以使首、尾两端具备有相邻点的条件  $P_0$ 、 $P_1$ 、 $P_2$  和  $P_{n-1}$ 、 $P_n$ 、 $P_{n+1}$ , 从而可产生  $P_1$ 、 $P_2$  曲线段和  $P_{n-1}$ 、 $P_n$  曲线段。所谓直接画出, 就是根据相邻 3 点  $P_1$ 、 $P_2$ 、 $P_3$  和  $P_{n-2}$ 、 $P_{n-1}$ 、 $P_n$  拟合出的每个抛物线直接画出曲线; 当每次取前两点画所拟合的抛物线时, 需要直接画出最后 3 点所拟合抛物线在点  $P_{n-2}$  与点  $P_n$  区间之间的所有曲线即可; 当每次取后两点画所拟合的抛物线时, 需要在开始拟合的抛物线中画出  $P_1$  到  $P_3$  区间之间的所有曲线即可<sup>[3-5]</sup>。

### 2.2 改进最小二乘正弦波拟合

当实验中可以获得每个采样值的位置量时, 就相当于已知采样的角频率。通过研究知道<sup>[7-9]</sup>, 对已知频率, 关于幅度、相位和直流分量的三参数正弦波曲线拟合是一种闭合的线性过程, 绝对收敛<sup>[6-10]</sup>。

最小二乘曲线拟合方法是拟合正弦曲线的一种常用方法, 它是一种可以获得正弦波波形参数和最小残差平方和的方法。

假设理想的正弦信号为

$$\begin{aligned} y(t) &= A_0 \cos(2\pi ft) + B_0 \sin(2\pi ft) + D_0 \\ &= C_0 \cos(2\pi ft + \theta_0) + D_0 \end{aligned} \quad (12)$$

数据记录序列为时刻  $t_1, t_2, \dots, t_n$  的采集样本  $y_1, y_2, \dots, y_n$ , 采样速率  $\nu$  已知时, 采样间隔为  $\Delta t$ ,  $t_i = i \times \Delta t = i/\nu$ , ( $i=1, 2, \dots, n$ ), 数字角频率  $\omega = 2\pi f/\nu$ , 则上面的公式可变成下列离散形式

$$\begin{aligned} y(i) &= A_0 \cos(\omega i) + B_0 \sin(\omega i) + D_0 \\ &= C_0 \cos(\omega i + \theta_0) + D_0 \end{aligned} \quad (13)$$

正弦参数波形曲线拟合过程, 当已知数据信号的采样频率, 即为信号的数字角频率已知, 选取  $A$ 、 $B$ 、 $D$ , 使下式所述残差平方和最小

$$\varepsilon = \sum_{i=1}^n [y_i - A \cos(\omega i) - B \sin(\omega i) - D]^2 \quad (14)$$

则参数  $A$ 、 $B$ 、 $D$  即为  $A_0$ 、 $B_0$ 、 $D_0$  的最小二乘拟合值。为寻找出  $A$ 、 $B$ 、 $D$ , 构造矩阵

$$\Psi = \begin{bmatrix} \cos(\omega \cdot 1) & \sin(\omega \cdot 1) & 1 \\ \cos(\omega \cdot 2) & \sin(\omega \cdot 2) & 1 \\ \dots & \dots & \dots \\ \cos(\omega \cdot n) & \sin(\omega \cdot n) & 1 \end{bmatrix}; y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}; x_0 = \begin{bmatrix} A \\ B \\ D \end{bmatrix} \quad (15)$$

则残差可表示为

$$\varepsilon = \varepsilon(\omega) = (y - \Psi x_0)^T (y - \Psi x_0)$$
 (16)

当式(4)~(15)最小时可得  $x_0$  的最小二乘  $\hat{x}_0$

$$x_0 = (\Psi^T \Psi)^{-1} (\Psi^T y)$$
 (17)

拟合函数为

$$\hat{y}(i) = A \cos(\omega i) + B \sin(\omega i) + D$$
 (18)

其幅度和相位表达式为

$$\hat{y}(i) = C \cos(\omega i + \theta) + D$$
 (19)

即可得到幅值和相位的表达式为

$$C = \sqrt{A^2 + B^2}$$
 (20)

$$\theta = \arctan \left( \frac{-B}{A} \right)$$
 (21)

3 拟合算法

3.1 算法流程

根据抛物线拟合理论和最小二乘法，提出了该算法原理，其算法流程如图 1 所示。

数据读入程序后，就根据数据的数量进行抛物线拟合，然后将拟合点间的曲线细分取值，细分后根据系统要求判断是否满足细分条件，不满足重新细分，直至满足条件，然后再继续程序进行细化的最小二乘拟合，得到拟合后所需要的正弦参数并输出，即可得到正弦波形及其数据。

3.2 Matlab 软件仿真

据拟合思想及算法流程利用 Matlab 软件编写了相关算法，并按照算法对一组给定数据进行了仿真实验<sup>[11]</sup>。根据实验方案，正弦信号经直流滤波后，只留下交流分量，在频率已知的情况下，只对位相和幅值两个未知正弦参数进行拟合就可以了。在仿真实验中所采用的理论输入正弦函数表达式为  $f=2\sin(2x+\pi/3)+1$ 。给定一组 101 个对应  $(x,f)$  的采样数据序列数进行拟合计算。经过最小二乘拟合之后的正弦曲线，曲线整体平滑，与理论曲线极其相近。计算其误差如表 1 所示。

表 1 仿真拟合误差比较

		直接拟合值	抛物线二乘拟合值
幅值 A	计算值	1.999 947 97	2.000 061 70
	误差	0.000 061 70	-0.000 052 03
相位 $\phi$	计算值	1.047 331 174	1.047 211 3
	误差	-0.000 133 622	-0.000 013 81

从表 1 中的仿真结果可以看出，经过最小二乘法拟合之后的正弦曲线参数值与直接数据拟合相比，幅值误差稍提高不多，但相位误差提高很多，误差达到了  $10^{-5}$  以下，对于正弦参数的数据处理来说，幅值足可以满足其精度需求，相位精度则得到了很大提高。

3.3 实验结果分析

用该算法对 6 组实验采集数据进行处理。该 6 组数据为

同一状态环境下不同时间实验采集到的 6 组数据，其理论曲线为一正弦曲线。传感器在旋转转台上旋转一周，将不同位置采集到的离散数据点与位置点对应进行拟合，并求出正弦曲线的参数，这里主要考虑相位参数。如表 2 所示。

由实验结果比较分析可以看出，经过抛物线细分后的最小二乘方法有效利用了原始数据的特点，在多次测量时通过数据变化规律增加数据点有效减小了测量过程中的随机干扰，可更为精确的求出正弦参数中的相位。

表 2 实验结果比较

	最小二乘	抛物线最小二乘
(1)	3.898 253	3.898 290
(2)	3.897 826	3.898 186
(3)	3.898 332	3.898 375
(4)	3.897 968	3.898 315
(5)	3.898 125	3.898 287
(6)	3.898 312	3.898 295
均值	3.898 137	3.898 362
方差	0.000 455 5	0.000 136 7

4 结束语

为了提高正弦曲线数据拟合与处理的精度，本文在抛物线拟合和最小二乘拟合法的基础上，提出了抛物线—最小二乘正弦曲线拟合法。该方法按照数据序列的先后，顺序以每一数据点为首连续取 3 个点进行局部抛物线拟合，依照抛物线拟合函数式对所采集数据（仿真为给定的数据）拟合公式所用的前两点进行细分插值，将细分后的数据序列利用最小二乘拟合法按照系统要求进行高精度拟合求出正弦曲线参数，适用于相位精度要求较高的场合。通过 Matlab 仿真试验和实验数据分析，该方法有效利用了原始数据，并提高了正弦曲线参数拟合的精度和局部数据处理的精度，在实验中验证可使相位估计重复精度从最小二乘的  $-4.555 \times 10^{-4}$  提高到  $1.367 \times 10^{-4}$ 。

参考文献：

[1] 芦文波.提取正弦信号参数的非线性寻优最小二乘算法[J].上海交通大学学报,2003,37(10):1613-1615.

[2] 梁志国,张大治,孙璟宇,等.四参数正弦波曲线拟合的快速算法[J].计测技术,2006,26(1):4-7.

[3] 张凤蛟.快速曲线拟合的方法[J].延边大学学报(自然科学版),2006,32(3):208-211.

[4] 司少玲,关水.三角函数曲线数据拟合最佳次数的确定[J].计算机工程与设计,2006,27(24):4660-4662.

[5] 邓振森,刘渝,王志忠.正弦波频率估计的修正 Rife 算法[J].数据采集与处理,2006,21(4):473-477.

[6] 梁志国,孙璟宇.波形测量中的基本问题讨论[J].计量技术,2004,(12):13-17.

[7] 李太福,熊隽迪.基于梯度下降法的自适应模糊控制系统研究[J].系统仿真学报,2007,19(6):1265-1268.

[8] 杨汉生,吕家云,李明,等.一种拟合型的提升小波预测方法[J].系统仿真学报,2006,18(9):2681-2683.

[9] 薛国新,孙玉强.正弦曲线三点拟合问题的一种新方法[J].计算机仿真,2006,23(2):107-109.

[10] 苏德伦,王仕成,张安京.基于遗传算法的正弦波四参数曲线拟合[J].计测技术,2005,25(6):18-20.

[11] 黄忠霖.控制系统 MATLAB 计算及仿真[M].北京:国防工业出版社,2006.

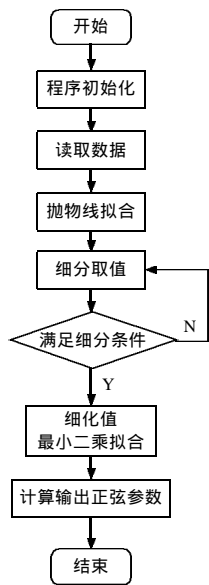


图 1 算法程序流程