

彭钊、陈志遥、吕品姬, 2015, 多面函数法在钻孔应变时序资料处理中的应用, 中国地震, 31(2), 382 ~ 389.

多面函数法在钻孔应变时序 资料处理中的应用

彭钊 陈志遥 吕品姬

中国地震局地震研究所(地震大地测量重点实验室)武汉市武昌区洪山侧路 40 号 430071

摘要 基于现有的钻孔应变连续观测资料, 将多面函数拟合法应用于时序资料的处理, 讨论了核函数选择不同参数对于拟合效果的影响, 得到了最优的拟合结果并分析了其反映的原始数据的物理内涵及应用方向, 同时与多项式拟合法的拟合结果进行了简要对比。

关键词: 多面函数拟合 核函数 钻孔应变时间序列

[文章编号] 1001-4683(2015)02-0382-08 [中图分类号] P315 [文献标识码] A

0 引言

钻孔连续应变观测兼具良好的高频特性与高灵敏度, 资料的连续性强, 其中的信息十分丰富, 涵盖了从高频到零频的各种地壳运动信息, 同时包含大量噪声及干扰信息, 噪声中还包含着仪器随机噪声和各种环境影响周期成分, 这些特性要求在进行时序资料的处理时要做到: 理论完备、计算稳定简便, 以便快速处理大量的高频时序数据且便于分析其地球物理内涵; 拟合及插值精度高, 以免许多微弱的地壳运动信息被当做噪声处理; 能够有效进行信息分离和特征提取, 以分辨地壳信息及各种噪声的特征和变化规律。目前观测资料的处理方法有多项式拟合、切比雪夫曲线拟合等, 随着台站数量的增多和观测技术的进步, 这些方法逐渐显得较为复杂, 且分析得到的信息不够充分。因此, 很多学者在探讨一种功能较完备且简单易行的方法。本文介绍的多面函数拟合法是比较适合钻孔连续应变观测资料处理的较优秀的方法。

1 多核函数法

1.1 多核函数法的基本思想

多面函数拟合法由 Hardy(1978) 提出, 基本思想是任何数学表面和任何不规则的圆滑表面, 总可以通过一系列有规则的数学表面的合成, 以任意精度逼近。其二维数学表达式为

$$f(x, y) = \sum_{j=1}^m c_j q_j(x, y, x_j, y_j) \quad (1)$$

[收稿日期] 2014-12-26; [修定日期] 2015-04-29

[作者简介] 彭钊, 生于 1991 年, 男, 在读硕士研究生, 主要从事地壳形变资料分析研究。E-mail: pengzhao0505@163.com

式中 $q_j(x, y, x_j, y_j)$ 为核函数, (x_j, y_j) 为核点的坐标, c_j 为待定系数。核函数在理论上可以任意选择,但在地壳形变分析中,一般是一个非常简单的曲面函数,通常是“钟”形或“钵”形曲面函数,常用的“钟”形核函数有倒双曲函数、倒数型函数和高斯函数等,常用的“钵”形核函数有双曲函数等。多面函数法最初用于地壳垂直形变分析,后广泛应用于大地测量各方面,国内在 20 世纪 90 年代初就有学者对该方法进行了应用和研究,陶本藻等(2003)研究了该方法的数学模型,刘经南等(2001、2002)把该方法应用于地壳水平运动和应变场的描述。

1.2 核函数参数的选取

根据核函数的数学思想,当将其用于一维自变量的情况时(杨国华等,1990;杨博等,2010),其数学表达式为

$$f(t) = \sum_{j=1}^m c_j q_j(t, t_j) \quad (2)$$

多面函数的拟合效果与效用不仅取决于核函数的个数,还取决于核函数的选取与核函数点及其间隔的选取,需要根据处理数据的具体情况合理选择。假如处理的是时序资料,其最优模型为

$$f(t) = \sum_{j=1}^m c_j q_j(t, t_j) = q^T C \quad (3)$$

式中 $q^T = (q_1, \dots, q_m)$ 为核函数系数, $C^T = (c_1, \dots, c_m)$ 为待定参数, m 为核函数个数。用最小二乘法求解待定参数。

由此可以列出误差方程为

$$V = QC - L \quad (4)$$

法方程为

$$NC + W = 0 \quad (5)$$

因此待定参数 c 的求解表达式为

$$C = -N^{-1}W \quad (6)$$

式中 Q 为根据核函数

$$q^T = \begin{bmatrix} q_{1,1} & \cdots & q_{m,1} \\ q_{1,2} & \cdots & q_{m,2} \\ \vdots & \ddots & \vdots \\ q_{1,i} & \cdots & q_{m,i} \end{bmatrix} \quad (i = 1, 2, \dots; n, n \text{ 为观测值个数}) \quad (7)$$

所组成的误差方程系数矩阵,(4)式中的 L 为由观测数据所组成的观测量值矩阵, $N = Q^T Q$, $W = Q^T L$,只要 $Q^T Q$ 满足非奇,根据以上各式可以计算得到待定参数 c 以及拟合值中误差

$$M = \pm \sqrt{\frac{\sum V^T V}{n - m}} \quad (8)$$

由中误差 M 的大小来衡量拟合的效果。

就核函数的选择而言,选择形态较简单“钟”形或“钵”形函数较为可靠,当数据点较为密集时,选择“钟”形函数较为实用(杨国华等,1990),由于钻孔应变观测时序资料采样率

高,时间连续性好,因此本文采用“钟”形高斯型核函数

$$q_j(t, t_j) = e^{-\left(\frac{t-t_j}{k}\right)^a} \quad (9)$$

式中 a 为幂次值, k 为平滑因子。影响拟合效果的主要是核函数点和幂次 a 的选择,若核函数的形式一旦选定,则平滑因子 k 的改变也将影响拟合效果,不过一般来讲影响较小。

现利用实际观测资料讨论核函数参数的选取,选择云南省昭通台分量式钻孔应变仪 2014 年 1 月 1 ~ 3 日的观测数据进行了验算,且年初台站仪器经过调零,数据可靠性提高。通过对原始观测数据进行标定和换算(邱泽华等,2009),得到 EW 和 NS 分量应变变化分钟值时间序列(图 1)。

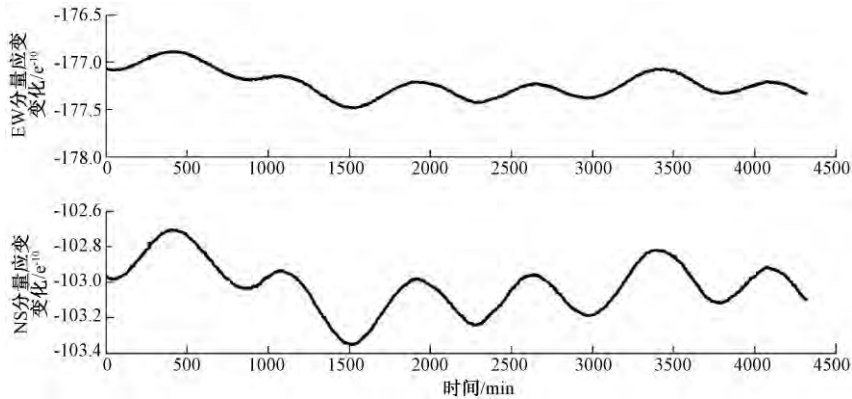


图 1 昭通台 EW 分量、NS 分量应变的时序变化曲线

平滑因子 k 越大,核函数越平缓,单个数据点对拟合的影响越小(齐娜等,2010),由图 1 可知时序数据的数据点密集,曲线较平滑,故 k 取较大值为宜,本文取 10000;核函数点的选择从观测时间开始按等间隔确定,分别取 60、90、120、180、240、300min 作为时间间隔;幂次 a 的初始值设定为 0.8,增量为 0.05,逐渐累加并逐一计算相应的中误差,得到的结果如图 2 和图 3,并以中误差的大小作为衡量与选择的标准。

图 2、3 表明滤波结果与核函数点的间隔、幂次值大小密切相关。当幂次值为 0.5 ~ 3.5、核点间隔小于 240min (4h) 时,同一幂次值下,中误差随着核点间隔的增大而增大,同一核点间隔下,大部分区间内中误差随着幂次的增大而减小,但在幂次值约等于整数 2、3 的小范围区间内结果不稳定,中误差增大,这表明当幂次 a 取近似整数时结果常不稳定,应该尽量避免,这与杨国华等(1990)的研究结果一致;当幂次超过 3.5 并逐渐增大时,拟合精度逐渐降低,计算结果趋于不稳定,核点间隔越小,趋于不稳定的临界幂次值越小,核点间隔大于 180min (3h) 的情况下,幂次小于 3.5 时拟合结果都是稳定的;但当核点间隔小于 3h 的情况下,临界幂次都小于 3.5。核点间隔 90 ~ 120min 时,临界幂次大约为 3;而当核点间隔减小到 60min 时,临界幂次降到约 2.5。表明虽然拟合中误差大致随着核点间隔的较小而降低,其计算稳定性也随之降低,因此核点间隔并不是越小越好,应根据具体情况加以选择。综合来看,核点间隔取 90min,幂次取 2.6 是一个较优的选择,此时中误差为 0.0039 (EW 分量) 和 0.0036 (NS 分量)。

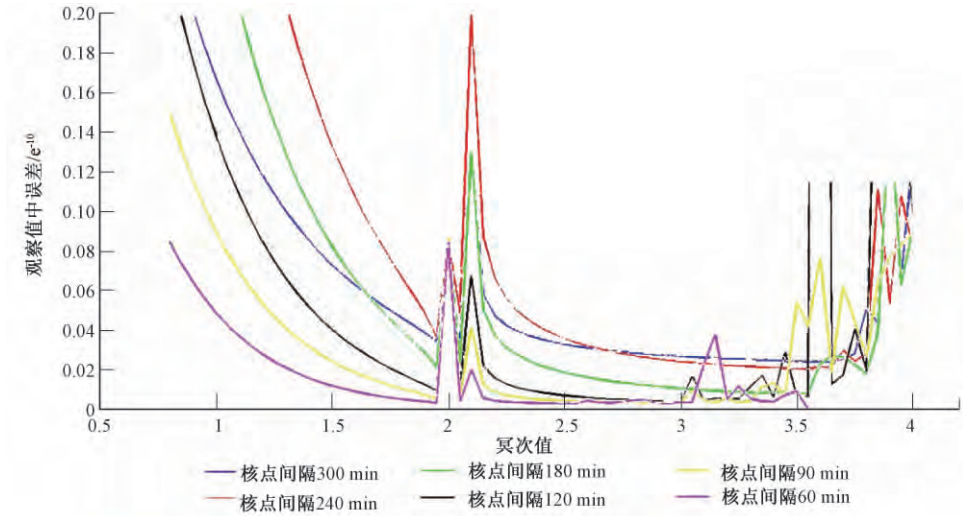


图2 核函数间隔与中误差变化关系(EW分量)

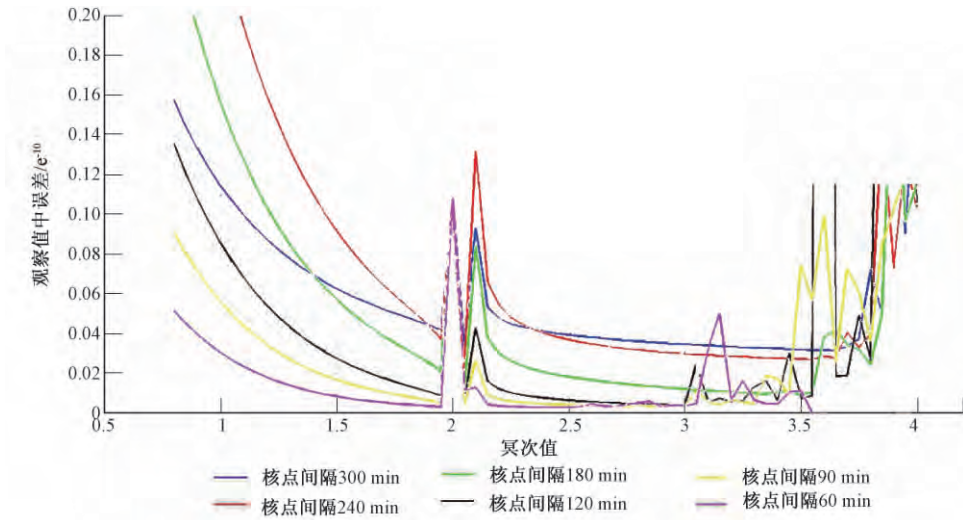


图3 核函数间隔与中误差变化关系(NS分量)

2 结果分析

2.1 拟合结果分析

图2和图3的拟合结果清楚地说明拟合效果受控于核函数点的间隔大小,实际上核函数点间隔的长短主要体现了信息的不同频域,当核函数点的间隔越长,所显现的信息越趋于低频,反之亦然。幂次则影响同一信息频域的拟合效果,幂次越大,单个数据点对拟合结果的影响就越大,钻孔应变仪置于钻孔中进行观测,隔绝了地表诸多自然与人类的干扰,但是由于钻孔应变仪观测精度高,而且时间连续性好,故观测数据除了应变固体潮和各频段的地壳应变信息外,还夹杂有大量的大小不一、频段各异的干扰和噪声信息,因此拟合函数的幂次不宜过大。

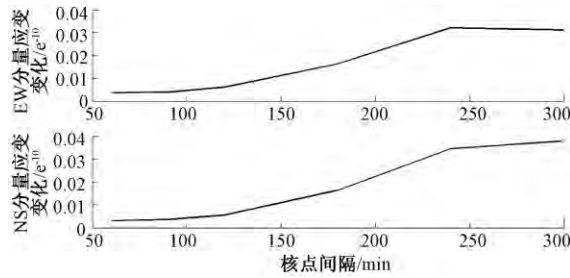


图4 核点间隔与拟合中误差的关系

图4给出的核点间隔与拟合中误差的关系表明误差在不同频谱范围所占的比重:EW分量和NS分量总体趋势上中误差随着核点间隔的增大而增大,体现出钻孔应变时序观测数据中信息丰富,频域分布广,符合钻孔应变观测本身的特点;核点间隔小于90min(1.5h)时,中误差变化小,在90~240min内中误差明显增大,它表明时序观测数据中的干扰和噪声信息主要集中在90~240min(1.5~4.0h)的周期内,由台站观测日志的记载可知该时期基本没有大的地震或人为干扰,因此该时期的干扰和噪声信息应该是主要来自周围短期环境变化,核点间隔在240min以上时,EW分量中误差较稳定,变化率较小,而NS分量240min以上区间的中误差变化率明显大于60~90min区间,这可能反映了NS方分量上某种低频长周期的地壳运动趋势;而NS分量中误差的平均值比EW分量略小,这可能表明NS分量的观测环境较EW方向的好,也可能反映了NS方向上某种低频长周期的地壳运动趋势代表了该方向的应变变化的主要影响因素,而EW方向没有整体的应变变化趋势,故而背景中误差略大。

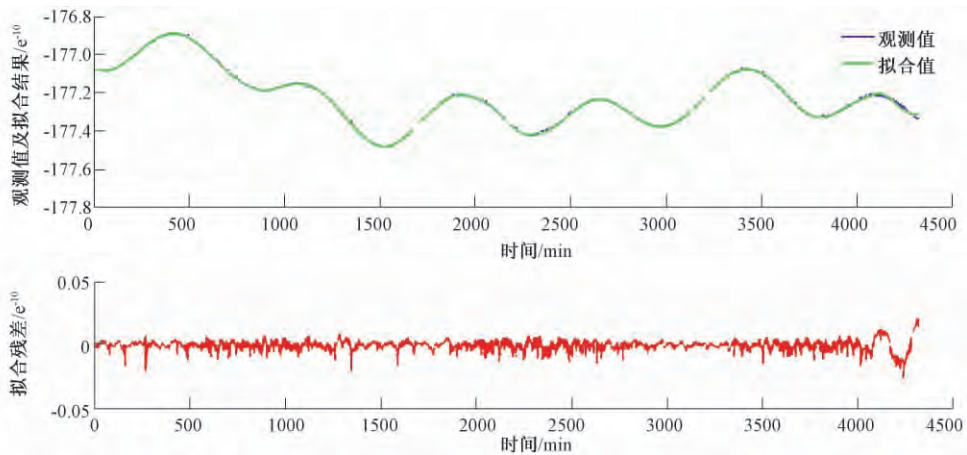


图5 昭通台EW分量时序数据拟合结果和残差

图5和图6选择的核点间隔为90min,核函数选择的是高斯核函数,幂次 c 和平滑系数 k 分别为2.6与10000时昭通台时序资料的拟合结果和相应的残差,由图可知EW向和NS向平均残差都在 ± 0.01 以内,拟合结果较好,仅两侧端点处残差较大,是端点效应所致,要消除端点效应,可多取边界外一段数据参与拟合计算。该方法可用于时序数据的插值、预处理

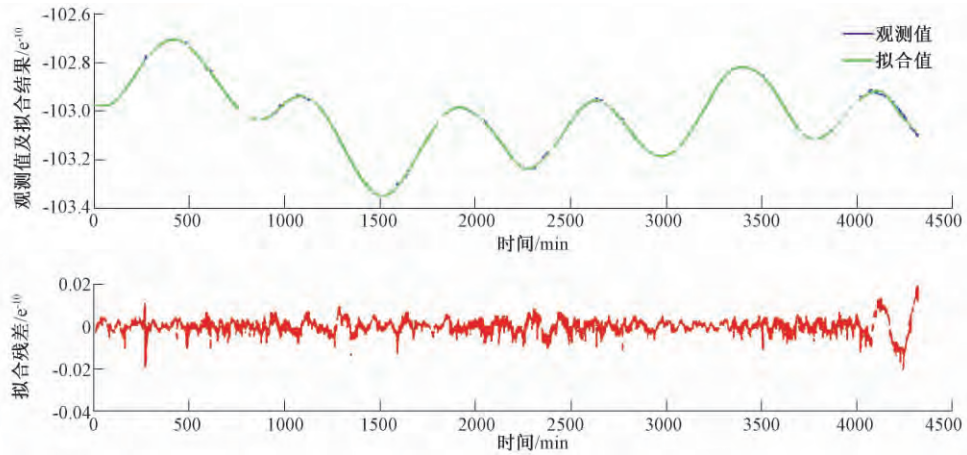


图 6 昭通台 NS 分量时序数据拟合结果和残差

理等,并可以拟合残差为判断因子,以数倍中误差或合适值为阈值,应用于异常识别及自动报警系统,另外,由于得到了时序数据的数学解析式,方便了进一步的信息分离、趋势分析等计算。

2.2 与多项式拟合结果的对比

使用多项式拟合法对昭通台该段时序数据进行处理,采用 21 阶多项式拟合结果如图 7,拟合中误差分别为 0.0062(EW 分量)和 0.0076(NS 分量),将该结果与多面函数拟合法进行比较,可知:

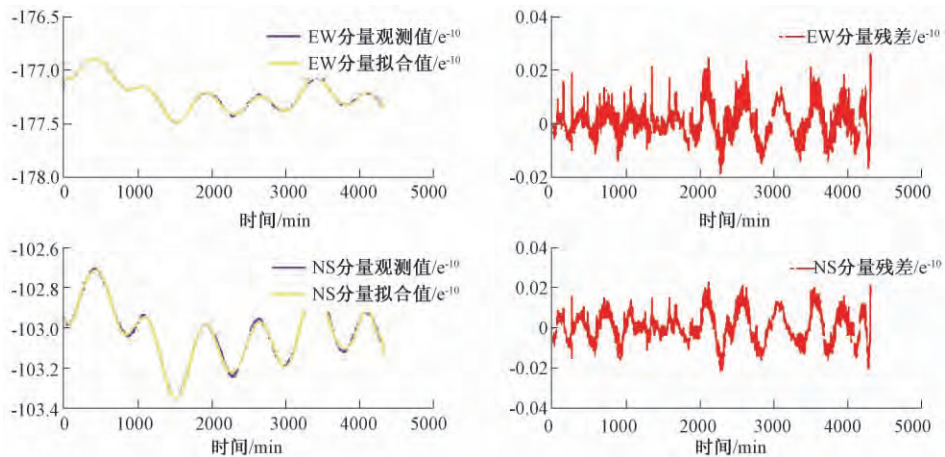


图 7 昭通台时序资料多项式拟合结果及残差

(1) 多面函数法拟合结果精度较高,中误差大约是多项式拟合法的 50%,且多项式拟合法的拟合精度随着阶数的下降而快速下降,当阶数为 10 时,拟合中误差增大到 0.08 左右,而多面函数法拟合精度受核点间隔与幂次的改变影响较小,核点间隔 60 ~ 180min,幂次在 2 ~ 3.5 变化时,拟合中误差都在 0.02 以下,计算稳定性和可靠性也高于多项式拟合法。

(2) 多项式拟合残差曲线与原始数据曲线相似,原始数据变化大的时段拟合残差也大,

说明多项式拟合精度受图形形状影响较大,而多面函数拟合残差在整个时段上分布比较均匀,受图形形状影响较小,可靠性比多项式拟合要高。

(3) 两种方法都有较严谨的理论基础,同时计算也比较简单,适于计算机自动处理,但都存在端点效应,需注意避免。

3 结语

本文利用多面函数法拟合昭通台分量式钻孔应变仪 2014 年 1 月 1~3 日 EW 向和 NS 向时序观测数据,得到了拟合函数解析式,通过对误差的比较分析讨论了参数选择,讨论了其物理内涵并与多项式拟合的结果进行了简要对比。

分析表明,昭通台的钻孔应变仪观测环境稳定,信息丰富,频域分布广,人为干扰少,数据质量好;环境干扰噪声周期主要集中在 90~240min(1.5~4h)内;NS 方向上可能存在某种低频长周期的地壳运动趋势。综上所述,多核函数法在钻孔应变时序资料的拟合、信息分离等处理上有精度高、计算简单可靠性好、功能完备等钻孔应变数据处理要求的特点,同时相对于多项式拟合等方法它还兼有如下优点:①时序资料处理中很容易完成不同频段信息的提取与分离;②能得到数学解析式,参数及拟合值精度评定严谨,可在此基础上进行进一步分析;③计算结果稳定,拟合残差分布不受图形形状影响。因此,多面函数法不失为一种处理钻孔应变连续观测时序资料较优秀的方法。

参考文献

- 刘经南、施闯、姚宜斌等,2001,多面函数拟合法及其在建立中国地壳平面运动速度场模型中的应用研究,武汉大学学报(信息科学版),**26**(6),500~503及508。
- 刘经南、姚宜斌、施闯,2002,中国地壳运动整体速度场模型的建立方法研究,武汉大学学报(信息科学版),**27**(4),331~336。
- 邱泽华、石耀霖,2004,国外钻孔应变观测的发展现状,地震学报,**26**(增刊 I),162~168及176。
- 邱泽华、阚宝祥、唐磊,2009,四分量钻孔应变观测资料的换算和使用,地震,**29**(4),83~89。
- 齐娜、胡良柏,2010,关于多面函数拟合方法中平滑因子的研究,工程勘察,**10**(9),77~79。
- 陶本藻、姚宜斌,2003,基于多面核函数配置型模型的参数估计,武汉大学学报(信息科学版),**28**(5),547~550。
- 杨博、张风霜、韩月萍,2010,多核函数法在GPS时序资料处理中的应用,大地测量与地球动力学,**30**(4),137~141。
- 杨国华、黄立人,1990,速率面拟合法中多面函数几个特性的初步数值研究,地壳形变与地震,**10**(4),70~82。
- Hardy R L,1978,The application of multi-quadric equations and point mass anomaly models to crustal movement studies,NOAA Technical Report.

Method of multi-Kernel function and its application in borehole strain timing data processing

Peng Zhao Chen Zhiyao Lv Pinji

Institute of Seismology, China Earthquake Administration, Wuhan 430071, China

Abstract Based on the existing borehole strain continuous observations, we use the multi-Kernel function to deal with the timing data and discuss the impact of Kernel function on fitting with different parameter selections. We analyze the physical connotation of the original data and its possible applications. Meanwhile, a brief comparison is done between the fitting results of multi-Kernel function method and polynomial fitting method.

Key words: Method of multi-Kernel function Kernel function Borehole strain timing data