

利用 MUSIC 算法分析 BL Lac 天体 S5 0716+714 光变周期*

唐 洁[†]

(陕西理工学院物理与电信工程学院 汉中 723001)

摘要 将基于多重信号分类的 MUSIC 谱估计算法引入 BL Lac 天体光变周期分析中, 给出了 MUSIC 算法的基本原理, 利用模拟信号检测了算法的频谱分辨率. 从大量文献中收集了 BL Lac 天体 S5 0716+714 光学 V、R、I 3 个波段从 1994 年到 2008 年的有效观测数据, 用 MUSIC 算法和平均周期图算法分别计算了它们的光变周期, 发现存在两个主要光变周期: 一个是 (3.33 ± 0.08) yr 的周期, 另一个是 (1.24 ± 0.01) yr 的周期. 对这两种算法的周期估计性能进行了比较, 结果表明, MUSIC 谱估计算法对样本长度要求较低, 具有良好的分辨特性和抗噪声能力, 能提高在样本长度较短情况下光变周期分析的准确性.

关键词 蝎虎天体; 个别, 星系; 基本参数, 方法; 分析

中图分类号: P 157; **文献标识码:** A

1 引言

光变是指天体辐射流量随时间的变化, 研究活动星系核的光变是一项极有意义的工作. BL Lac 天体作为活动星系核中重要的一个子类, 光变现象十分激烈, 各个观测波段都观测到了光变, 光变时标跨度从几天到几年. 随着观测技术的进步和高精尖端观测设备的利用, 甚至观测到了小时量级、秒量级的光变时标. 光变时标、不同波段光变的关联性以及时间延迟等为我们提供了许多了解 BL Lac 天体中心结构及其物理过程的重要信息, 如利用光变时标可以估计辐射区半径大小、亮温度等^[1].

分析光变周期的方法有很多, 选择一种有效的周期分析方法显得特别重要. Jurkevich 方法、离散相关分析法、CLEAN 算法、功率谱分析方法、小波分析方法、高阶谱分析方法、结构函数等方法常用来分析光变资料的周期. 经典谱估计法中的周期图谱分析法、现代谱估计法中的自回归模型法、高阶谱分析方法中的双谱分析方法等我们进行过较多对比研究, 结果表明这些方法分析长周期光变是有效的^[2-6], 但这些周期分析方法的采用有一定的条件和局限性, 若能同时采用几种不同的统计分析方法对数据进行对比

2010-11-08 收到原稿, 2011-08-15 收到修改稿

* 陕西省教育厅专项科研基金项目 (11JK0903) 资助

[†] tj168@163.com

分析, 可以获得较为可信的结论. 本文引入利用特征值分解来处理的 MUSIC 算法, 它可用于空间测向和功率谱估计. 最近几年, MUSIC 算法被大量学者成功地应用于波束到达方向的估计, 而另一重要功能功率谱估计, 还很少见于文献. 这种算法所需要的先验信息较少, 具备了现代谱估计算法的高分辨特性, 对样本长度要求较低, 原理上比较简单, 运用这种方法能够在低信噪比下得到高分辨率和高稳定性的谱估计.

S5 0716+714 是一颗著名的 BL Lac 天体, 长期以来一直受到关注, 如 Impey 等^[7]对 S5 0716+714 进行了长达 3 yr 的光学观测, Qian 等^[8]从 1994 年到 2000 年在中国科学院国家天文台兴隆观测站对 S5 0716+714 进行了光学观测, Nesci 等^[9]从 1995 年 2 月到 2003 年 3 月对 S5 0716+714 光学波段的光变进行了观测. 由于其观测数据较多, 已有很多学者用不同方法得到大致相同的长光变周期.

本文概括介绍了 MUSIC 功率谱估计方法的基本原理, 将 MUSIC 算法用低信噪比的模拟信号进行检验, 为了验证 MUSIC 算法用于长周期光变分析的优良性能, 采取了与周期图算法进行比较的方式. 为了对比不同周期分析方法应用效果, 将 MUSIC 谱估计算法和平均周期图算法用在 BL Lac 天体 S5 0716+714 光学 V、R、I 3 个波段的历史光变资料中, 分析可能存在的长光变周期, 并对这两种方法的计算结果进行简单的讨论.

2 MUSIC 谱估计法的基本原理

MUSIC 算法是多重信号分类 (Multiple Signal Classification) 的英文缩写, MUSIC 谱估计法是空间谱估计技术的典型代表, 它根据矩阵特征分解理论, 将信息空间分成两个正交的子空间即信号子空间和噪声子空间, 通过噪声子空间的矢量和信号子空间的矢量的正交性, 可以估计信号中所包含的频率成分, 从而得到信号周期^[10].

信号中总存在我们感兴趣的信号和干扰噪声, 我们通常要对这些信号进行分离, 抑制那些干扰的噪声, MUSIC 算法就是通过特征分解将有用的信号从干扰的噪声中提取出来, 在信号频率处功率谱出现峰值, 而在其他地方值较小. 只要观测样本的长度 N 大于或等于 $P+1$ (P 为信号的个数), 就可以估计出样本中所有的频率, 再根据周期是频率的倒数可得到周期.

将要分析的信号构建成数据向量 $\mathbf{x} = [x_0, x_1, \dots, x_N]^T$, 式中 T 表示转置运算符, 它的相关矩阵为

$$\mathbf{R}_x = E[\mathbf{x}^* \mathbf{x}^H], \quad (1)$$

式中 $*$ 表示复共轭运算符, H 表示共轭转置运算符. 矩阵 \mathbf{R}_x 的特征值分解可表示为

$$\hat{\mathbf{R}}_x = \sum_{i=1}^N \lambda_i \boldsymbol{\nu}_i^* \boldsymbol{\nu}_i^H, \quad (2)$$

式中 λ_i 为第 i 个特征值, $\boldsymbol{\nu}_i$ 为对应于 λ_i 的特征向量. 信号由 p 个信号源和方差为 σ_n^2 的内部噪声组成, 且

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > \sigma_n^2, \lambda_{p+1} = \lambda_{p+2} = \dots = \lambda_N = \sigma_n^2. \quad (3)$$

这些特征值 $\lambda_1, \lambda_2, \dots, \lambda_p$ 对应的特征向量组成信号空间, 它的相关矩阵为 \mathbf{R}_s , 特征值 $\lambda_{p+1}, \lambda_{p+2}, \dots, \lambda_N$ 对应的特征向量组成噪声空间, 它的相关矩阵为 \mathbf{R}_n , 因而有

$$\mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_n = \sum_{k=1}^p \lambda_k \boldsymbol{\nu}_k \boldsymbol{\nu}_k^H + \sigma_n^2 \mathbf{I}. \quad (4)$$

由信号子空间与噪声子空间的正交性, 可以得到 MUSIC 算法的谱估计^[10]:

$$\hat{P}_{\text{MUSIC}}(f) = \frac{1}{\sum_{i=p+1}^N |\mathbf{e}^H \boldsymbol{\nu}_i|^2} = \frac{1}{|\mathbf{e}^H \mathbf{V}_n|^2}, \quad (5)$$

$\mathbf{e} = [1, e^{j2\pi f}, \dots, e^{j(N-1)2\pi f}]$ 表示信号向量, $\mathbf{V}_n = [\boldsymbol{\nu}_{p+1}, \boldsymbol{\nu}_{p+2}, \dots, \boldsymbol{\nu}_N]$ 表示与 $N - P$ 个零特征值对应的 \mathbf{R}_x 的特征向量, $\mathbf{e}^H \boldsymbol{\nu}_i$ 是 \mathbf{e} 在噪声子空间上的投影. 对应于信号中包含的每个频率 f_i , P_{MUSIC} 都会出现一个峰值

$$\hat{P}_{\text{MUSIC}}(f = f_i) = \max_{f=f_i} \frac{1}{|\mathbf{e}^H \mathbf{V}_n|^2} = \hat{P}_{\max}. \quad (6)$$

在实际应用中, 通常将 f 划分为数百个等间距单位, 将每个 f_i 分别代入上式计算 $\hat{P}_{\text{MUSIC}}(f_i)$, 求出所有峰值所对应的 f_i , 找到频谱的峰值也就确定了信号频率.

3 天文观测周期信号的模拟检验

为了检验 MUSIC 谱估计法的可靠性和优良性能, 用一个模拟信号来验证它, 并采取与平均周期图算法进行比较的方式. 由于实际天文观测的信号总存在着噪声, 模拟信号由包含了 3 个频率成分的信号和一噪声组成, 即

$$x(t) = \sum_{i=1}^3 A_i \sin(2\pi f_i t) + \xi_i, \quad (7)$$

其中 $A_1 = A_2 = A_3 = 6$, $f_1 = 0.1$ Hz, $f_2 = 0.15$ Hz, $f_3 = 0.3$ Hz. 为了更具代表性, 取频率 f_1 和 f_2 相差不多, 和 f_3 相差较大. 观测噪声 ξ_i 是高斯白噪声. 为了对比两谱分析方法抗噪性能, 逐步降低信噪比, 在本试验中, 抽样频率为 1 000 Hz, 起始采样点数为 256, 对采样所得序列分别进行平均周期图和 MUSIC 谱估计, 信号及其谱估计结果如图 1 所示, 图中横轴表示频率, 纵轴表示功率谱估计的幅度值, 为了突出峰值, 本文采用归一化功率谱图.

对模拟信号数据采样点数取 256 时, 两种方法都能准确地检测出信号频率, 如图 1(b) 和图 1(c). 当逐步降低采样点数时, MUSIC 算法得到的谱图 1(b) 基本不变, 当样本量较少, 采样点数降到 64 时, 平均周期图算法由于主瓣变宽, 频率相隔很近的 0.10 Hz 和 0.15 Hz 的两个信号尽管能分辨出来但谱图明显有重叠现象, 如图 1(d) 所示. 如果采样点数再减少, 这两个信号便分辨不开了. 当降低信噪比时, MUSIC 算法得到的谱图会产生一些毛峰现象, 但不会干扰信号峰值的正确判断, 平均周期图算法效果更差, 会出现很多伪峰, 干扰主峰频率的判断.

通过以上模拟信号仿真可知, MUSIC 算法即使样本量较少时,也可以通过降低算法的实时性来提高频率估计的分辨率. 能够从其频域峰值找到对应的频率, 频率的倒数就是信号的周期. 在低信噪比的情况下, 也能通过抑制噪声将信号准确检测出来. 在天文观测过程中, 由于多方面的影响, 观测资料受到背景噪声干扰较多, 周期图谱分析方法有可能得到伪周期, 无法准确识别出信号频率, 而 MUSIC 谱估计仍能在低信噪比情况下准确地检测出信号频率. 上面的仿真结果证实了 MUSIC 算法检测周期的可行性和优势.

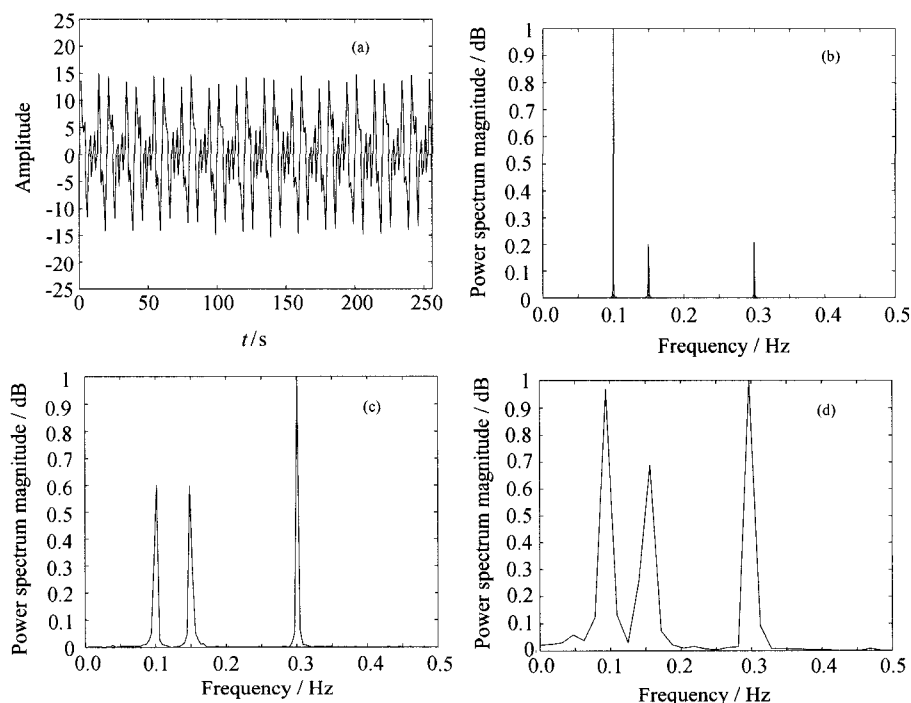


图 1 模拟信号. (a) 原始波形图; (b) MUSIC 谱; (c) 平均周期图谱 ($N=256$); (d) 平均周期图谱 ($N=64$)

Fig. 1 The modeling signal. (a) The original oscillogram; (b) MUSIC power spectrum; (c) Average periodogram power spectrum ($N=256$); (d) Average periodogram power spectrum ($N=64$)

4 计算和分析 BL Lac 天体 S5 0716+714 的光变周期

BL Lac 天体 S5 0716+714 各波段的变化十分激烈, 从被发现以来保存有较多的光学观测数据, 已发表的文献表明它具有短时标光变、中等时标光变和长时标光变 [7-9,11-23]. 为了确定真实周期, 在长时间内观测到足够多的数据资料是必要的. 本文收集了从 1994 年到 2008 年期间 S5 0716+714 光学 V、R、I 3 个波段的观测数据, 其中 Qian 等 [8]、Raiteri 等 [11] 给出了 1994 年到 2001 年的观测数据, Montagni 等 [12] 给出了 1996 年到 2003 年的观测数据, 最近几年的观测数据主要有 Wu 等 [13-14] 提供的 2003 年到 2006 年

的观测数据, 以及 Poon 等^[15] 提供的 2008 年的观测数据. 收集的观测数据比张皓晶等^[16-18] 分析 S5 0716+714 的光变周期数据要多, 他们给出了 S5 0716+714 光学 V、R、I 3 个波段的光变曲线图.

在实际天文观测中, 通常一个月内会对观测对象进行多次测量, 因而一个月的数据点会比较多, 而这也是造成噪声的原因之一, 对周期分析不利, 为了消除这一影响, 将收集的 S5 0716+714 在光学 I、R、V 3 个波段的历史光变资料的数据进行一个月的算术平均处理. 对于没有观测数据的点用内插法求得的数据替代, 再将这些数据按发生的先后构成新的时间序列. 用 MUSIC 算法和平均周期图谱分析方法分别对这些进行过预处理的等间隔观测数据进行长周期性光变分析.

平均周期图算法和 MUSIC 算法得到的 S5 0716+714 光学 I、R、V 3 个波段的长周期光变图如图 2 所示, 图中横轴表示频率, 纵轴表示谱估计的幅度值. 图 2(a)、图 2(c)、图 2(e) 是用 MUSIC 算法得到的结果, 图 2(b)、图 2(d)、图 2(f) 是平均周期图算法得到的结果, 图 2(a)、图 2(b) 是 I 波段的分析结论, 图 2(c)、图 2(d) 是 R 波段的分析结论, 图 2(e)、图 2(f) 是 V 波段的分析结论. 若在某频率处有明显的峰值, 则可判定原观测序列中该频率分量的功率谱强度较大, 是序列的一个主要周期分量. 通过 6 个谱图可以看出, 在频率约 $(0.025 \pm 0.0006) \text{ month}^{-1}$ 处, 功率谱有一个明显的峰值, 说明观测数据中含有约 $(3.33 \pm 0.08) \text{ yr}$ 的周期成分, $(3.33 \pm 0.08) \text{ yr}$ 这个主周期与其他文献分析的结论相吻合^[6,11,17-18].

对 S5 0716+714 进行谱分析的目的是要找出隐藏在历史光变资料数据中的各主要周期分量, 但由于各主周期分量的作用强度不同, 在谱图上作用强度最大的周期分量即对应谱峰的主要周期分量一般较易分辨, 而某些强度较小的周期分量则易被掩盖.

图 2(a) 在 I 波段主峰幅度非常大, 峰值比较尖锐, 其它地方都是平坦的, 不存在明显的次峰, 数据周期性非常明显, 在图 2(b) 谱线中还有个次峰, 频率是 $(0.032 \pm 0.0008) \text{ month}^{-1}$, 其对应的周期为 $(2.60 \pm 0.07) \text{ yr}$, 这一结果其他作者都未提到, 对应的振幅很小, 是否是真实的光变周期还需要更多的资料来验证. 在 I 波段, MUSIC 算法和周期图算法具有几乎相同的频率估计精度.

图 2(c) 主峰频率不完全为 $(0.025 \pm 0.0006) \text{ month}^{-1}$, 存在少量的谱线漂移现象, 谱线比较光滑, 主瓣宽度比较宽, 造成估计误差. 另有一次峰频率是 $(0.067 \pm 0.001) \text{ month}^{-1}$, 次峰幅度相对主峰来说不是太小, 对应的周期为 $(1.24 \pm 0.01) \text{ yr}$, 这一结果与张皓晶等^[16] 得到的 1.10 yr 较为一致, 并且张皓晶等认为 1.10 yr 可能和 Raiteri 等^[11] 得到的 3.30 yr 是倍周期关系, 由此看来 1.10 yr 的周期可能才是真正的周期. 如果这样的话, 这一周期可能并不是这一天体的真正周期, 而是由于地球自转引入的, 1.10 yr 到底是否是真正的周期需要以后的观测去证实.

从图 2(d) 可以看出, 也存在一个次周期为 $(1.24 \pm 0.01) \text{ yr}$, 周期图算法得到的谱图存在许多伪峰, 原因是这些方法所用的观测数据的质量依赖于信噪比、系统误差以及数据分布特性. 图 2(e) 有一次峰频率为 $(0.062 \pm 0.007) \text{ month}^{-1}$, 周期是 $(1.34 \pm 0.01) \text{ yr}$, 和图 2(c)、图 2(d) 得到的 $(1.24 \pm 0.01) \text{ yr}$ 存在差异, 原因是由于观测数据总是有限长, 在功率谱估计中, 信号空间的相关矩阵是由有限长数据估计出来的, V 波段观测数据存在

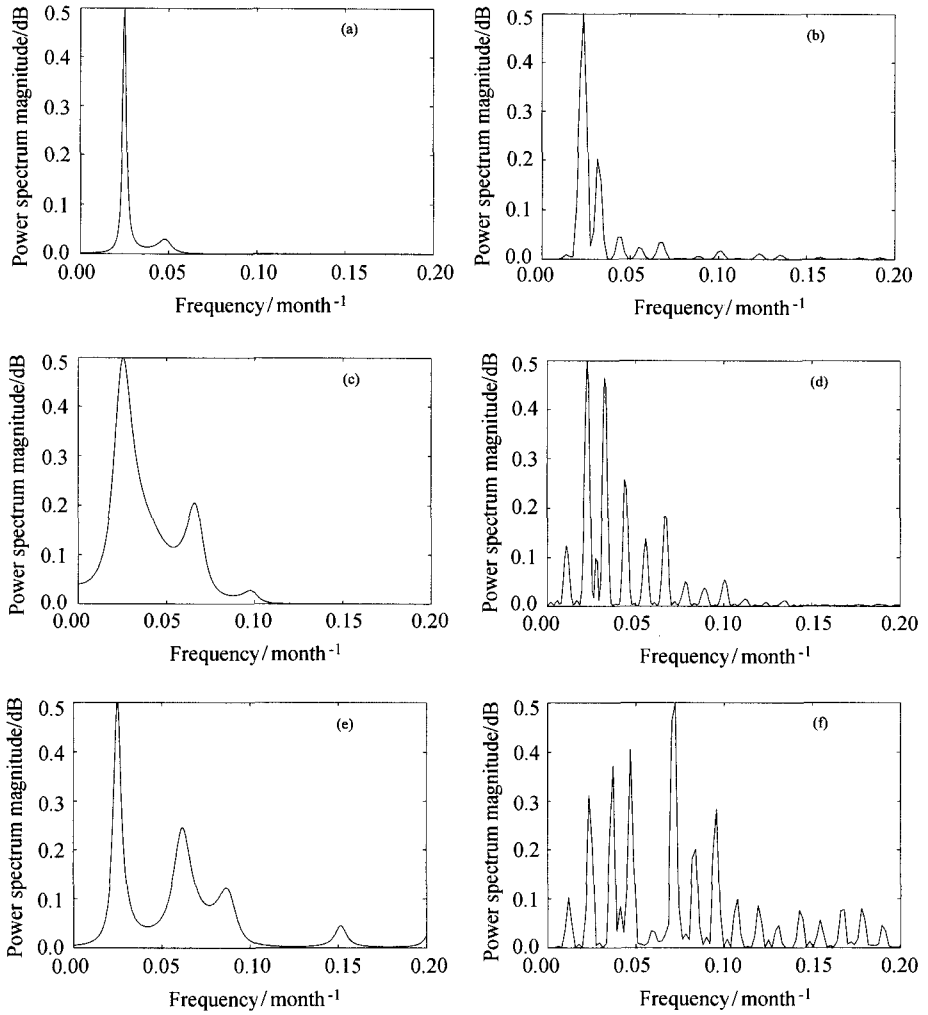


图 2 S5 0716+714 谱估计图. (a) MUSIC 算法 (I 波段); (b) 平均周期图 (I 波段); (c) MUSIC 算法 (R 波段); (d) 平均周期图 (R 波段); (e) MUSIC 算法 (V 波段); (f) 平均周期图 (V 波段)

Fig. 2 The power spectrum estimation of S5 0716+714. (a) MUSIC power spectrum (I band); (b) Average periodogram power spectrum (I band); (c) MUSIC power spectrum (R band); (d) Average periodogram power spectrum (R band); (e) MUSIC power spectrum (V band); (f) Average periodogram power spectrum (V band)

较强的噪声分量, 当有较强的噪声分量影响, 观测数据样本量较少时, 必然会产生一定的误差^[24]. 图 2(f) 从周期图谱图上看到很多峰, 次峰强度在一定程度上影响了主峰的判断, 出现较为严重的栅栏效应和谱线分裂现象, 从而产生虚假的峰, 其谱估计质量明显下降, 伪峰出现的几率增加, 其位置的随机性很强, 伪峰干扰了周期峰的判断.

5 讨论

BL Lac 天体 S5 0716+714 的光变周期分析研究现在已获得许多结果. 对短时标光变的研究显示出存在短时标周期, 而对长时标光变的分析研究也显示了长周期性. Heidt 等^[19]得到的光学波段可能周期为 4 d, Wu 等^[13-14]发现它有 0.17 d、0.11 d、0.23 d 的周期, Qian 等^[20]认为周期是 10 d, Quirrenbach 等^[21]推测它的周期从 1 d 变化到 7 d, Ma 等^[22]获得了 14 d 的周期, Sagar 等^[23]总结出的周期约为 30 d, Katajainen 等^[1]观测到它的两个最大流量的时间间隔为 60~70 d. Raiteri 等^[11]采用离散傅立叶变换、离散相关分析法、结构函数法发现它的光学波段光变周期为 3.3 yr, 射电波段的周期为 5.5~6 yr. 张皓晶等^[16]得到的光学 B、V、R、I 波段周期为 1.1 yr, 他们用小波分析方法得到的周期为 1 160 d^[17-18], Qian 等^[8]发现其有 5.3 yr 的光变周期. 我们在收集最新的观测数据的基础上, 曾先后用自回归方法和双谱方法对 S5 0716+714 的光变周期在光学波段重新进行分析, 结果发现 S5 0716+714 存在大约 3.33 yr 和 3.47 yr 的周期^[6]. 由此看来此源的变化特性是否具有一个或多个周期成份还存在争议, 即使用同一个数据样本, 其光变曲线中推出的周期性也有不同.

本文将 MUSIC 谱估计算法和平均周期图算法用在 BL Lac 天体 S5 0716+714 光学 V、R、I 3 个波段的历史光变资料中, 得到它们的周期为 (3.33 ± 0.08) yr、 (1.24 ± 0.01) yr. 这个结论和 Raiteri 等^[11], 张皓晶等^[16-17]采用其他方法得到的结论基本是一致的. 结合仿真试验和 BL Lac 天体 S5 0716+714 光变周期分析, 说明了周期图法以其物理概念清晰、方法简便和可利用 FFT 计算效率高等优点, 已成为功率谱估计的一种基本方法, 但当观测数据不是太多、信噪比较低的情况下, 可能出现虚假谱峰现象, 导致得到错误的周期, 影响了其在实际中的应用. 总体来说, MUSIC 算法利用了特征值分解处理, 显著改善了低信噪比下的谱估计性能, 具有很高的分辨率, 从理论上说可以达到无限的精度, 是现代谱估计的代表, 样本长度较短时也可以较准确估计信号中的周期成分, 这在利用天文观测数据分析周期有重要的应用价值, 它的估计精度优于周期图算法, 而且其频率分辨率也明显好于周期图算法.

致谢 感谢审稿人仔细阅读本文并提出了许多宝贵的审稿意见, 感谢北京大学天文系吴学兵教授提供跟他学习的机会和指导, 感谢陕西理工学院张义明、何宁、傅明星、庞桥、仁亚杰、何军锋等教授对去北京大学学习的支持, 感谢陕西理工学院计算物理创新团队基金资助.

参 考 文 献

[1] Katajainen S, Takalo L O, Sillanpää A, et al. A&AS, 2000, 143: 357

- [2] 唐洁, 张雄, 伍林. 天文学报, 2007, 48: 488
- [3] Tang J, Zhang X, Wu L. ChA&A, 2008, 32: 151
- [4] 唐洁, 张晓娟, 庞桥, 等. 天文学报, 2009, 50: 364
- [5] Tang J, Zhang X J, Pang Q, et al. ChA&A, 2010, 34: 121
- [6] 唐洁, 张雄. 物理学报, 2010, 59: 7516
- [7] Impey C D, Bychkov V, Tapia S, et al. AJ, 2000, 119: 1542
- [8] Qian B C, Tao J, Fan J H. AJ, 2002, 123: 678
- [9] Nesci R, Massaro E, Rossi C, et al. AJ, 2005, 130: 1466
- [10] 张明友, 吕明. 近代信号处理理论与方法. 北京: 国防工业出版社, 2005: 121-180
- [11] Raiteri C M, Villata M, Tosti G, et al. A&A, 2003, 402: 151
- [12] Montagni F, Maselli A, Massaro E, et al. A&A, 2006, 451: 435
- [13] Wu J H, Zhou X, Ma J, et al. AJ, 2007, 133: 1599
- [14] Wu J H, Peng B, Zhou X, et al. AJ, 2005, 129: 1818
- [15] Poon H, Fan J H, Fu J N. ApJS, 2009, 185: 511
- [16] 张皓晶, 张雄. 物理学报, 2007, 56: 4305
- [17] 张皓晶, 张雄, 董富通, 等. 天文学报, 2009, 50: 141
- [18] Zhang H J, Zhang X, Dong F T, et al. ChA&A, 2009, 33: 373
- [19] Heidt J, Wagner S J. A&A, 1996, 305: 42
- [20] Qian B C, Tao J, Fan J H. PASJ, 2000, 52: 1075
- [21] Quirrenbach A, Witzel A, Wagner S, et al. ApJ, 1991, 372: L71
- [22] Ma L, Xie G Z, Zhou S B, et al. IJMPD, 2004, 13: 659
- [23] Sagar R, Gopal-Krishna, Mohan V, et al. A&AS, 1999, 134: 453
- [24] 胡广书. 数字信号处理 - 理论、算法与实现. 北京: 清华大学出版社, 1997: 37-62

An Analysis of Periodic Components in BL Lac Object S5 0716 +714 with MUSIC Method

TANG Jie

(School of Physics and Telecommunication Engineering, Shaanxi University of Technology,
Hanzhong 723001)

ABSTRACT Multiple signal classification (MUSIC) algorithms are introduced to the estimation of the period of variation of BL Lac objects. The principle of MUSIC spectral analysis method and theoretical analysis of the resolution of frequency spectrum using analog signals are included. From a lot of literatures, we have collected a lot of effective observation data of BL Lac object S5 0716 + 714 in V, R, I bands from 1994 to 2008. The light variation periods of S5 0716 +714 are obtained by means of the MUSIC spectral analysis method and periodogram spectral analysis method. There exist two major periods: (3.33 ± 0.08) years and (1.24 ± 0.01) years for all bands. The estimation of the period of variation of the algorithm based on the MUSIC spectral analysis method is compared with that of the algorithm based on the periodogram spectral analysis method. It is a super-resolution algorithm with small data length, and could be used to detect the period of variation of weak signals.

Key words BL Lacertae objects: individual, galaxies: fundamental parameters, methods: analytical