

基于结构函数方法的类星体证认*

唐洁¹⁾ 傅明星¹⁾ 吴学兵^{2)†}

1) (陕西理工学院物理与电信工程学院, 汉中 723001)

2) (北京大学物理学院天文系, 北京 100871)

(2012年3月13日收到; 2012年5月11日收到修改稿)

光变巡天是搜寻类星体的一个很有效的方法. 应用归一化结构函数对 SDSS stripe 82 中存在光变的 1411 个类星体和 174 个恒星进行分类, 选取幂函数模型拟合这些样本 5 个波段的结构函数. 综合幅值和幂指数两个物理量设置了合适的分类判据, 获得了比较理想的分类结果, 类星体和恒星最高分类正确率分别达到了 92.2% 和 83.6%. 结果表明, 改正后的结构函数对类星体的证认有很好的效果, 能节约大型望远镜宝贵的观测时间, 让类星体巡天获得比较高的观测效率. 分析结论支持大部分类星体的光变缘于盘的不稳定性.

关键词: 结构函数, 类星体, SDSS, 光变

PACS: 95.80.+p, 98.62.Tc, 95.75.-z, 02.50.Fz

1 引言

对活动星系核 (AGNs) 的研究一直是天体物理领域的重要前沿课题之一, 类星体是 AGNs 中光变最为剧烈的一个子类, 具有明亮的致密核区、非热致连续辐射、强发射线、高偏振、激烈光变等一系列极端物理特征^[1-4]. 目前宇宙的起源、结构和演化是天文学研究中最为活跃的领域, 被视为宇宙探针的类星体是目前观测发现距离我们最遥远的极少数天体之一. 类星体的时间和空间跨度非常大, 通过研究它的光度函数、成团性、空间分布等能够揭示类星体演化的规律, 特别是高红移类星体能够提供在星系形成时的信息, 这些信息有助于宇宙大尺度结构和星系形成及演化理论的研究^[5-8]. 在开展这些研究前, 关键问题是寻找到大量的类星体样本.

利用类星体有别于其他天体的特征可以用来寻找类星体. 由于光学波段是最有效的一个巡天波段, 其他波段巡天获得的类星体候选者需要经过光谱证认才能确定, 因此光学波段的巡天是寻找类星

体的主要方法. 类星体的一个重要观测特征是存在光变现象, 由于变星很少, 光变巡天是寻找类星体比较好的方法之一^[8]. 应用结构函数方法从类星体的光变资料可以获得类星体的许多重要参数, 如光变时标. 类星体的结构函数一般存在两个平台, 一个平台在时间间隔大于光变时间尺度处, 另一个平台在时间间隔远小于光变时标处, 在这两个平台之间有一个近似的线性过程^[9].

结构函数的那段线性过程表示辐射流量随时间的一个线性变化过程, 研究发现采用一定的函数模型拟合该线性区域, 对应的斜率可以区分类星体和恒星. 用来拟合的函数模型一般有幂函数模型和指数函数模型. Macleod 等^[10]用随机游走模型模拟 SDSS stripe 82 光谱证认过的 9000 多个类星体和少量恒星的光变曲线, 并用指数函数拟合它们对应的结构函数, 结果能很好地区分类星体和恒星. Schmidt 等^[11]用幂函数来拟合这些样本, 发现用颜色无法区分的恒星和部分高红移类星体, 在用结构函数方法证认时正确率明显提高. Palanque-Delabrouille 等^[12]认为高红移的类星体在用幂函数来拟合它们的结构函数时被恒星污染比

* 国家自然科学基金重点项目 (批准号: 11033001) 和国家重点基础研究发展计划 (批准号: 2007CB815405) 资助的课题.

† E-mail: wuxb@pku.edu.cn

例大, 他们运用神经网络模型来对样本分类时, 分类效果有所改观. 我们也利用其他方法得到了类星体选源判据, 并应用这些判据选取了一些类星体候选体, 这些候选体是否为类星体需要以后的光谱观测去证认 [13,14].

利用结构函数方法把类星体从众多的恒星中分离出来, 对于部分类星体效果不是太好, 主要原因是无论用幂函数模型还是指数函数模型, 都只是用了一个参量来区分. 幂函数模型直接用幂函数拟合结构函数值和时间延迟, 方法简便, 但在拟合方程中没有考虑结构函数的极大值, 而这个极大值具有重要的物理意义, 它代表光变大小. 指数函数模型尽管将结构函数的极大值引进了方程, 但其方程中缺少了一个像幂函数模型中那样的物理量: 幂指数就是结构函数线性过程的斜率. 本文结合这两个模型各自的优点, 选取 SDSS stripe 82 中光谱证认过的 1411 个类星体和 174 个恒星, 将结构函数的极大值也引进幂函数模型的方程中, 通过数据预处理中的归一化方法将类星体和恒星的结构函数值归一到一个区域, 避免了由于其结构函数值大小相差太大带来的影响, 同时幂指数的值和数据变换前一样. 这样不仅幂指数代表结构函数线性过程的斜率, 而且方程里的另一个代表幅值的物理量也具有物理意义, 并且这两个物理量都可以区分类星体和恒星. 综合考虑这两个量, 类星体的证认率比单一参量证认的正确率将得到大幅度的提高.

2 结构函数方法

结构函数方法是类星体光变研究中比较常见的一种方法, 在结构函数和时间延迟图中的近似直线部分的斜率表示天体辐射流量的变化率, 被许多研究者用来分析类星体的内禀光变, 并用适当的理论模型来解释类星体的光变 [15,16].

对于某一天体的观测量 $m(t_i)$, $i = 1, 2, \dots, N$, 任意时间延迟 τ , 其一阶结构函数 $SF(\tau)$ 为 [15]

$$SF(\tau) = \frac{1}{N(\tau)} \sum_{i=1}^N w(t_i)w(t_i + \tau)[m(t_i + \tau) - m(t_i)]^2, \quad (1)$$

$$N(\tau) = \sum_i w(t_i)w(t_i + \tau).$$

如果存在观测量 $m(t_i + \tau)$, 权重因子 $w(t_i) =$

$w(t_i + \tau) = 1$, 反之则 $w(t_i) = w(t_i + \tau) = 0$. 对于任意天体, 存在多阶结构函数, 阶数越多, 结构函数值越精确, 但计算量也越大, 一般只计算到一级结构函数.

结构函数 $SF(\tau)$ 与时间延迟 τ 满足一个很好的幂函数和指数函数拟合形式:

$$[SF(\tau)]^{1/2} = SF = A\tau^\beta, \quad SF = SF_\infty(1 - e^{-\tau/\Delta t})^{1/2}, \quad (2)$$

其中 β 表示幂指数, A 表示幅值, SF_∞ 表示结构函数的极大值, Δt 表示时间间隔.

用幂函数和指数函数拟合都有长处, 幂函数拟合直接拟合结构函数值与时间延迟, 但一个极重要的物理量, 即结构函数极大值在方程中并没有体现, 而指数函数拟合需要选择一个恰当的时间间隔 Δt , 如果选得不好, 拟合效果差, 且没有一个拟合参数表示结构函数线性部分的斜率. 本文结合这两种方法各自的优点, 将拟合方程变为

$$SF' = A'\tau^\beta = (SF_\infty \cdot A)\tau^\beta, \quad SF = \frac{SF'}{SF_\infty} = A\tau^\beta. \quad (3)$$

幂函数拟合方程经过这样变换后, 结构函数值都归一到 0—1 这个区间. 由于恒星的结构函数值比类星体的结构函数值要小很多, 现在将它们全归一化了, 可以使表示幅值的 A 具有可比性. 另外, 由于天气或采集数据的设备问题带来的离群观测值, 经过归一化后转换成适合数据处理的形式, 且也能用这个参量来对恒星和类星体进行区分, 同时改变后的方程不影响幂指数的值.

3 类星体的证认

斯隆数字巡天 (sloan digital sky survey, SDSS) 使用美国新墨西哥州 Apache Point 天文台的 2.5 m 口径地平式望远镜进行大面积天区数字巡天. SDSS 巡天项目从 2000 年开始启动, 巡天面积已达四分之一天区, 测光系统由 u, g, r, i, z 五种颜色的宽带滤光片组成, 这 5 个波段的中心波长分别为 3551, 4686, 6165, 7481 和 8931 Å, 极限星等分别为 22.0, 22.2, 22.2, 21.3, 20.5. 测光参数包括星等、颜色、轮廓、大小等. SDSS 计划测光系统能获得近亿个天体准确的坐标和星等, 对 100 多万个恒星和星系进行光谱证认. SDSS 现在已经公开释放了第 8 批数据 (SDSS DR8).

本文分析的类星体和恒星样本数据来自于 SDSS DR7 中的 SDSS stripe 82 已经被光谱证认过的类星体和恒星. stripe 82 光谱证认过的类星体为 9275, 为了分析方便, 我们从中选取 1411 个类星体和 174 个恒星, 包括 50 个红移大于 2.2 的高红移类星体. 要求这些样本的赤经、赤纬覆盖 stripe 82 整个区域, 且样本观测时间跨度在 5 年以上. 这是由于若观测资料的时间跨度不够长, 有些光变就探测不到, 可利用的观测数据超过 10 个, 这样挑出来的样本具有代表性.

由于望远镜的观测时间有限, 不可能对每一个天体进行长时间观测, 因此 SDSS 测光数据之间存在 90 到 270 d 的间隔没有观测数据. 为了得到天体的结构函数值, 我们用三次样条插值来代替这些缺失值. 时间延迟取 50 d 的倍数, 由于结构函数的最大值对应的特征时标为光变时标, 将最大的时间延迟取到大于光变时间尺度. 需要说明的是, 这里的时间延迟取观测者坐标, 不取为静止坐标, 这是由于当选择某一天体仅进行图像观测时, 并不知道它的红移. 图 1 给出了类星体 SDSS J033004.34 + 000901.7 的 i 波段光变曲线和结构函数图. 在图 1(b) 中, 该天体的结构函数值和拟合值相差很少, 拟合得非常好, 在时间延迟为 1.5 年以后, 结构函数值增加缓慢近似出现一个平台, 拟合函数的幂指数为 0.37, 在时间延迟 150 d 到 1.5 年之间结构函数有段线性区域, 其斜率和幂指数近似相等为 0.37. 在幅值 0.70 的位置, 其对应的时间延迟约为 1 年, 时间延迟大于 1 年后, 基本出现平台区. 由此看来, 引进归一化结构函数后, 拟合方程的参

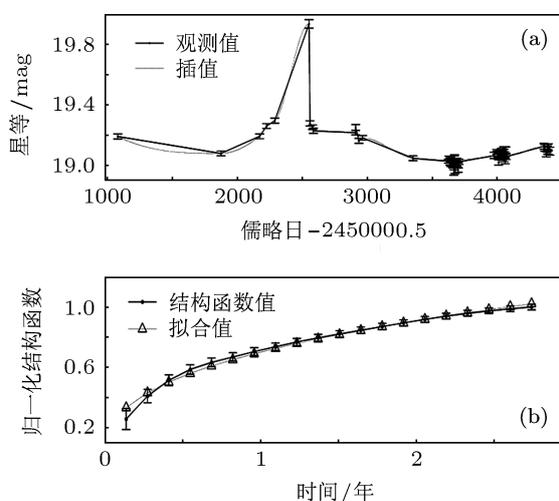


图 1 类星体 SDSS J033004.34 + 000901.7 的 i 波段 (a) 光变曲线; (b) 结构函数

数幅值也具有明显的物理意义.

大多数类星体的光变曲线是不规则的, 它们的结构函数形状也是不规则的, 出现结构函数最大值的平台的延迟时间也会不一样, 因此我们在取每个天体最大延迟时间时要使该平台出现. 有的天体由于其某一波段光变曲线的复杂性, 其结构函数不能用幂函数拟合, 所以我们选的这些类星体和恒星在分析该波段时要去掉. 5 个波段的结构函数拟合参数分布分别如图 2—6 所示.

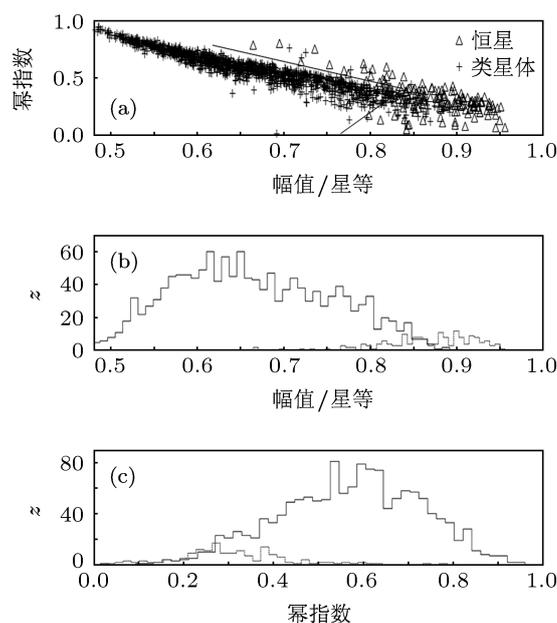


图 2 SDSS stripe 82 的类星体和恒星 i 波段结构函数拟合参数分布 (a) 幅值-幂指数分布图; (b) 幅值直方图; (c) 幂指数直方图

图 2 是类星体和恒星 i 波段的结构函数拟合参数分布图, 1400 个类星体、168 个恒星能用幂函数拟合, 拟合率超过 96%. 从图 2(a) 幅值-幂指数分布图中可见, 类星体和恒星有明显的界限, 相连的两条直线把它们分开在两个不同的区域; 从图 2(b) 幅值直方图可以看出 (实线代表类星体, 虚线代表恒星), 类星体的幅值大部分在 0.51 到 0.8 之间, 有 1288 个, 恒星小于 0.8 的数量很少, 大多数分布在 0.8 到 0.95 这个区域; 图 2(c) 是幂指数直方图, 类星体的幂指数大多数分布在 0.27 到 0.80 之间, 有 1311 个, 出现的峰值范围是 0.50 到 0.64. 而恒星完全不同于类星体, 幂指数小于 0.6 的恒星仅有几个, 分布在 0.19 到 0.43 这个范围有 145 个, 在 0.39 后与类星体重叠的少. 尽管用幅值或幂指数中的一个参量就可以区分类星体和恒星, 但为

了提高证认率,选择的判据将幅值和幂指数结合起来.我们以幅值 0.8,幂指数 0.4 为分类分割点,幅值小于 0.8,且幂指数大于 0.4 的判为类星体.幅值大于 0.8,且幂指数小于 0.4 的判为恒星.按照这个判据,有 1190 个类星体判对,只有 98 个判错,正确率高达 92.2%,有 39 个恒星判为类星体,正确率为 76.8%.

图 3—6 分别是类星体和恒星 g, r, z, u 波段的结构函数拟合参数分布图.这些波段的结构函数能用幂函数拟合的类星体个数分别为 1388, 1409, 1379, 1342 个,能拟合的恒星数分别为 161, 168, 165, 139 个,拟合率最高的 99.8%,最低的也有 80.0%.从这四个波段的幅值-幂指数分布图中的两根相连的直线能将类星体和恒星区分开,只有极少量的恒星污染类星体.由四个波段的幅值直方图可以看出,幅值小于 0.8 的恒星很少,小于 0.75 的极少,类星体幅值大部分小于 0.8,超过 0.85 的极少.从四个波段的幂指数直方图可知,类星体的幂指数普遍要大于恒星的幂指数,类星体幂指数从 0 到 0.95 之间的均有,但大部分分布在 0.4 到 0.75 这个范围,恒星幂指数超过 0.5 的不多,大部分分布在 0.1 到 0.4 这个狭窄的范围.因此,如 i 波段那样,我们将幅值 0.8 和幂指数 0.4 作为分割点,规定幅值小于 0.8,且幂指数大于 0.4 的天体判为类星体,反之幅值大于 0.8,且幂指数小于 0.4 的判为恒星.根据这个判据,四个波段的天体判为类星体的数目分别为 1198, 1267, 877, 956 个,正确率为 86.3%, 89.9%, 63.3%, 71.2%;判为恒星的数目分别为 125, 122, 138, 114 个,正确率为 77.6%, 72.6%, 83.6%, 82.0%.

利用类星体与恒星的颜色差异是认证类星体的一个重要的常用方法.为了比较颜色分类法与本文采用的改进的幂函数拟合结构函数方法的优劣,我们对同一类星体和恒星样本用颜色分类法进行分类.不同的颜色组合可以获得不同的两色图,本文选取 u, g, r 三色测光资料.图 7 是类星体和恒星的 (u-g)-(g-r) 双色图.从图中可以看出,类星体主要分布在标记为 A 和 D 的区域,在该区域受恒星污染的少,在选取类星体候选者时,位于区域 A 和 D 的天体入选.恒星大部分在标记为 B 的区域,在 C 区域类星体和恒星难以区分开, E 和 F 区域只有极少数的类星体和恒星的色落在该区域.如果选取 $u - g \leq 0.9$ 且 $g - r \leq 0.6$ 和 $u - g \geq 2.3$

且 $g - r \geq 1.2$ 为类星体的判据,则只有 66.9% 的类星体判对.将范围是 $0.9 \leq u - g \leq 1.4$ 且 $g - r \leq 0.3$ 判为恒星,其正确率也仅有 58.6%.与本文采用结构函数方法相比,颜色分类法的类星体证认率明显偏低.

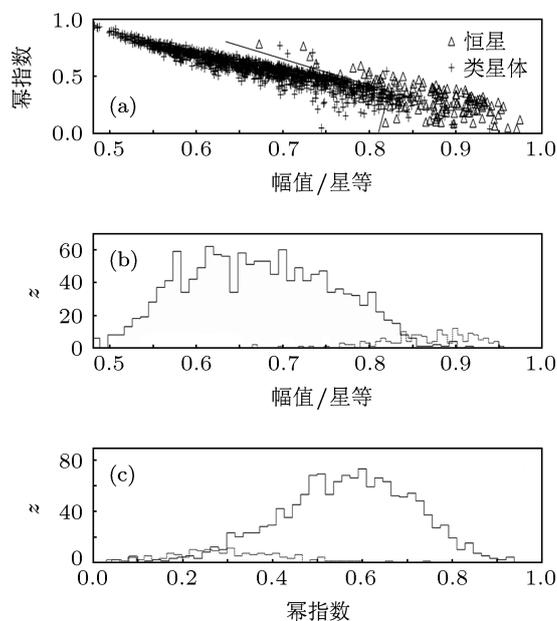


图 3 SDSS stripe 82 的类星体和恒星 g 波段结构函数拟合参数分布 (a) 幅值-幂指数分布图; (b) 幅值直方图; (c) 幂指数直方图

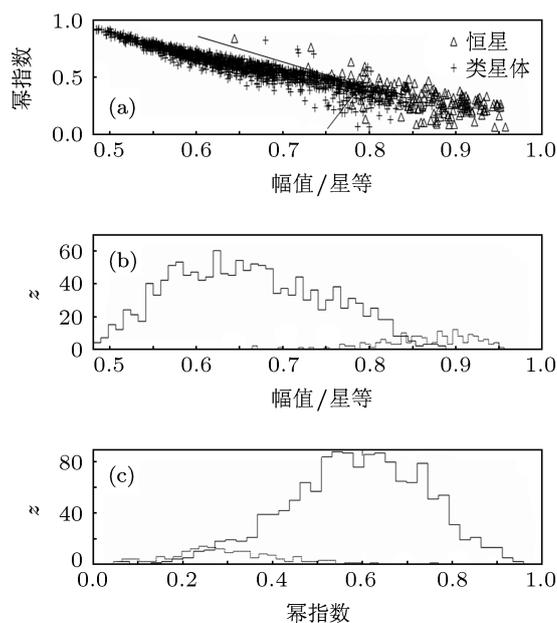


图 4 SDSS stripe 82 的类星体和恒星 r 波段结构函数拟合参数分布 (a) 幅值-幂指数分布图; (b) 幅值直方图; (c) 幂指数直方图

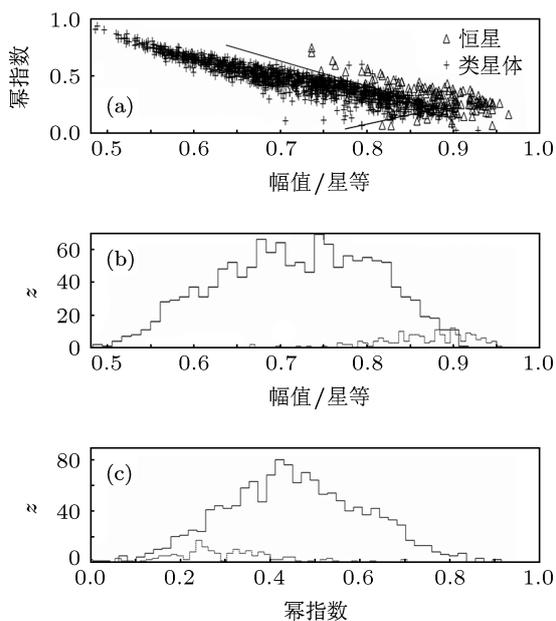


图5 SDSS stripe 82 的类星体和恒星 z 波段结构函数拟合参数分布图 (a) 幅值-幂指数分布图; (b) 幅值直方图; (c) 幂指数直方图

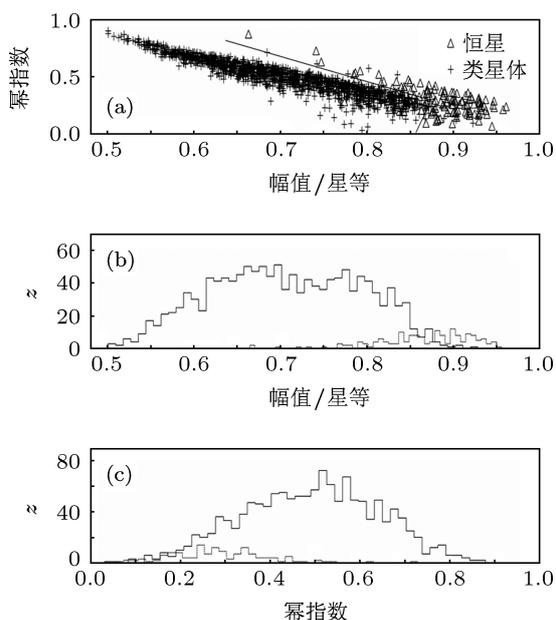


图6 SDSS stripe 82 的类星体和恒星 u 波段结构函数拟合参数分布图 (a) 幅值-幂指数分布图; (b) 幅值直方图; (c) 幂指数直方图

4 讨论和结论

到目前为止, 类星体的光变产生原因还不是十分清楚, 有三种理论模型可以解释类星体的光变: 盘不稳定模型, 星爆模型和引力透镜模型. 盘不稳定模型和星爆模型认为光变原因是内禀的, 是类

星体自身的变化导致光变的产生; 引力微透镜模型则认为光变的产生是外在因素引起的^[16]. 这三种理论模型都能一定程度上解释实际观测到的类星体光变, 尤其是盘不稳定模型得到很多观测上的支持. 类星体的高能辐射说明其中心可能存在大质量的黑洞, 长期光变可能源于黑洞吸积盘的热不稳定性^[17,18].

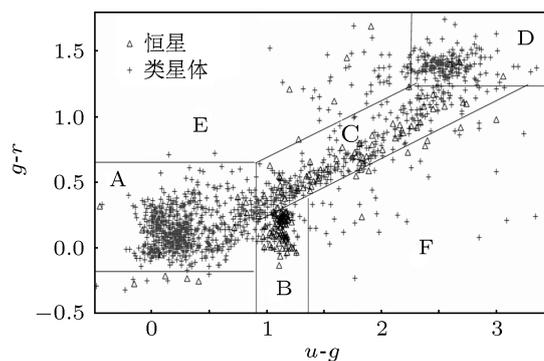


图7 类星体和恒星的 (u-g)-(g-r) 双色图

Hawkins^[15] 利用这三种理论模型模拟类星体的光变曲线, 计算出它们相对应的结构函数, 所有的结构函数都具有一个共同的特征, 即从某一时间延迟开始会有一段幂律形式的线性增长过程, 在一较大的时间尺度处逐渐变平. 线性增长过程用幂函数拟合得到的斜率中星爆模型最大, 范围是 0.74 到 0.90, 盘不稳定模型是 0.41 到 0.49 之间, 引力透镜模型是 0.22 到 0.28. 本文选取的类星体样本的结构函数用幂函数拟合得到的幂指数小于 0.2 的极少, 大多数分布在 0.38 到 0.8 的范围, 这个区域比较适合用盘不稳定模型和星爆模型来解释类星体的光变, 这两个模型都支持光变由随机事件引起的. 幂指数从 0.22 到 0.28 的类星体数目不是太多, 说明大部分类星体的光变是内禀的, 受外部影响引起光变的类星体比例较少, 一般来说这些类星体是高光度类星体.

用幂函数来拟合类星体光变的结构函数已有许多研究, Meusinger 等^[19] 进行了总结, 结构函数形式有三种, 利用不同形式的结构函数分析不同的类星体样本, 得到的结构函数线性部分的斜率也不一样, 我们的结论与 Wilhite 等^[20] 和 Rengstorf 等^[21] 的结果比较一致, 原因是采取的拟合方程相同, 只不过我们用的是归一化结构函数, 且计算结构函数的形式与文献^[20, 21] 不同. 由于本文将结构函数进行了归一化处理, 在不影响结构函数中线

性部分的斜率条件下,比其他文献增加了结构函数的幅值来区分类星体和恒星.通过这两个物理量将类星体和恒星区分在不同区域,并提出了一个选取类星体候选者的判据:幅值小于 0.8,且幂指数大于 0.4 的天体为类星体;幅值大于 0.8,且幂指数小于 0.4 的为恒星.利用该判据,证认类星体最高的正确率高达 92.2%,比常用的颜色分类法正确率高很多.因此尽管我们加强了选取类星体候选者的选择条件,在类星体候选者中排除掉更多污染的恒星,但却提高了证认类星体的效率,同时兼顾了样本的完备性.

本文用 SDSS 获得的类星体光变资料开展类星体遴选方法的研究,对我国 LAMOST 巡天望远镜

的类星体巡天观测计划具有重要意义.由于证认类星体效率的提高,为望远镜观测节约了大量的时间,可以对感兴趣的类星体进行长时间观测.为了避免利用各种方法挑选出来的类星体候选体测光数据之间存在较长的时间间隔没有观测数据,导致一些类星体光变由于没有观测而遗漏掉,建议 LAMOST 望远镜在类星体巡天观测时,类星体候选体观测时间间隔控制在 6 个月内,尽量将光变幅较小的类星体光变也观测到.

感谢审稿专家仔细阅读本文并提出许多宝贵的审稿意见,感谢美国斯隆数字巡天项目提供的数据.

-
- [1] Fan J H, Zhang Y W, Qian B C, Tao J, Liu Y, Hua T X 2009 *Astrophys. J. Suppl. Ser.* **181** 466
- [2] Fan J H, Liu Y, Qian B C, Tao J, Shen Z Q, Zhang J S, Huang Y, Wang J 2010 *Res. Astron. Astrophys.* **10** 1100
- [3] Tang J, Zhang X 2010 *Acta Phys. Sin.* **59** 7516 (in Chinese) [唐洁, 张雄 2010 物理学报 **59** 7516]
- [4] Tang J, Wu X B 2011 *Acta Phys. Sin.* **60** 119801 (in Chinese) [唐洁, 吴学兵 2011 物理学报 **60** 119801]
- [5] Cao X W, Li F 2008 *Mon. Not. R. Astron. Soc.* **390** 561
- [6] Li S L, Cao X W 2008 *Mon. Not. R. Astron. Soc.* **387** 41
- [7] Cao X W, Bai J M 2008 *Astrophys. J.* **673** 131
- [8] Huang K L 2005 *Quasars and Active Galactic Nucleus* (Beijing: Science Press) p33 (in Chinese) [黄克谅 2005 类星体与活动星系核 (北京: 科学出版社) 第 33 页]
- [9] Crisiani S, Trentini S, la France F 1996 *Astron. Astrophys.* **306** 395
- [10] MacLeod C L, Ž Ivezić C S, Kochanek S, Koztowski B, Kelly E, Bullock A, Kimball B, Sesar D, Westman K B 2010 *Astrophys. J.* **721** 1014
- [11] Schmidt K B, Rix H W, Shields J C, Knecht M, Hogg D W, Maoz D, Bovy J 2012 *Astrophys. J.* **744** 147
- [12] Palanque-Delabrouille N, Yèche C, Myers A D, Petitjean P, Nicholas P R, Sheldon E, Aubourg E 2011 *Astron. Astrophys.* **530** 122
- [13] Wu X B, Jia Z D 2010 *Mon. Not. R. Astron. Soc.* **406** 1583
- [14] Wu X B, Wang R, Schmidt K B, Bian F Y, Jiang L H, Fan X H 2011 *Astron. J.* **142** 78
- [15] Hawkins M R S 1996 *Mon. Not. R. Astron. Soc.* **278** 787
- [16] Hawkins M R S 1993 *Nature* **366** 242
- [17] Zhang J S 2004 *Chin. Phys.* **13** 1177
- [18] Deng J G, Huang R T 2008 *Acta Phys. Sin.* **57** 1285 (in Chinese) [邓家干, 黄仁堂 2008 物理学报 **57** 1285]
- [19] Meusinger H, Hinze A, Hoon A D 2011 *Astron. Astrophys.* **525** 37
- [20] Wilhite B C, Brunner R J, Grier C J 2008 *Mon. Not. R. Astron. Soc.* **383** 1232
- [21] Rengstorf A W, Brunner R J, Wilhite B C 2006 *Astron. J.* **131** 1923

Identification of quasars based by the variability structure function method*

Tang Jie¹⁾ Fu Ming-Xing¹⁾ Wu Xue-Bing²⁾†

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

2) (*Department of Astronomy, Peking University, Beijing 100871, China*)

(Received 13 March 2012; revised manuscript received 11 May 2012)

Abstract

The variability sky survey is a very effective method of searching for quasars. In the paper, we present a novel method to discriminate quasars from variable stars based on their intrinsic variabilities. The power-law model is used to fit the light-curve structure function in five wavebands, but the value of the structure function is normalized. The developed method is applied to 1411 spectroscopically confirmed quasars and 174 stars in SDSS stripe 82. We set a suitable criterion, then obtain a reasonable classified result. The classification accuracies about quasars and stars reach 92.2% and 83.6% respectively. Compared with the structure function only based on one physical parameter, the normalized structure function with amplitude and power-law index has a good efficiency to deal with these data. The method for selecting quasar candidates ensures big sky survey telescope a high survey efficiency and saving valuable astronomical observation time. The results support that the optical variabilities of most of quasars originate from the instabilities of accretion disks.

Keywords: structure function, quasars, SDSS, light curve

PACS: 95.80.+p, 98.62.Tc, 95.75.-z, 02.50.Fz

* Project supported by the Key Program of the National Natural Science Foundation of China (Grant No. 11033001) and the National Basic Research Program of China (Grant No. 2007CB815405).

† E-mail: wuxb@pku.edu.cn