

信息论背景下Ia型超新星宇宙学的贝叶斯 统计分析和距离对偶关系检验

马 聪[†]

(中国科学院紫金山天文台 南京 210034)

本文中, 我们系统地表述一种自洽的宇宙学模型分析及其参数估计方法, 并将其用于Ia型超新星(SNIa)光度距离作为宇宙学标准烛光的研究. 我们利用贝叶斯图论方法及其与随机变量间因果性、信息流和概率独立性的对应关系构建我们的分析, 并通过实例展示如何在宇宙学数据分析中恰当地运用联合光变曲线分析(JLA) SNIa数据集中依赖于参数的协方差矩阵. 我们的方法从概念到产出的结果均不同于传统的分析方法. 可以证明, 通常形如 χ^2 最小化的传统分析等效于施加在SNIa标准化模型变量先验分布上的隐含失真. 通过显性地指明这一失真的形式, 我们指出它无论从天体物理还是统计方法角度理解均缺乏合理性. 与之形成对比的是, 我们的贝叶斯统计模型明确地包含各个参数的先验分布. 与先前研究对照, 关于SNIa光度基于光变曲线时间延展和极大时期色指数的标准化改正, 本文的新结果显著支持更小的改正量. 关于受到宿主星系恒星质量控制的SNIa光度改正, 我们得出了更低的统计显著性. 对 w CDM (cold dark matter)和 Λ CDM宇宙学模型, 我们得出了修正的后验概率估计, 并分别得出了对 $w \geq -1$ 和更高 Ω_M 估计值的支持.

这一统计分析方法问题在SNIa数据集合压缩的工作中再次得到体现. 为此, 我们将数据压缩表述为统计推断问题, 并得出其最优化、最大熵解及其高效而准确的高斯近似. 通过多种统计和信息学检验, 这一近似的准确性和稳健性得到了确认. 与JLA原先压缩对比, 我们得出的距离模数数值及其协方差矩阵的结构均呈现出差异. 使用这些新的压缩数据, 我们进行SNIa距离阶梯的内部交叉核实检验, 并对表观上的色改正系数演化趋势给出了新的结果.

在此基础上, 我们论证了SNIa压缩数据用于检验距离对偶关系(DDR)的优越性, 即子样本更良好的代表性和统计不确定度更准确的建模. 与辅助性的观测信息结合后, 压缩的SNIa光度距离可与重子声振荡(BAO)巡天所得的角直径距离数据对应并直接比较. 我们使用蒙特卡罗分析得出DDR判别量 η 的分布性质. 这一计算结果与DDR在统计不确定度限度内一致, 以标准差表示的不确定度约为8%. 假设DDR表观背离的机制来源于对波长选择性低的星系际尘埃吸收造成SNIa亮度系统性偏低, 我们通过结合*Planck*宇宙学后验分布得到更紧密的限制, 并将其表达为星系际消光和光深的红移较差. 其计算结果在改进的不确定度限度内与无消光一致. 我们根据未来巡天观测的性能指标估算得出 η 可达到的统计不确定度约2%. 这一结果显示了系统不确定度对未来相关研究的影响以及针对这些影响采取自洽的统计解决方法的重要性.

[†]2019-01-06获得博士学位, 导师: 中国科学院紫金山天文台冯珑珑研究员; cma@pmo.ac.cn

总结本文的主要发现, 我们指出宇宙学模型统计分析中贝叶斯方法的优势. 关于未来可期待的观测宇宙学数据, 我们强调更深入研究的目标在于提出新的统计方法, 以更完好地保持这些结构更复杂、样本量更大的数据集合中蕴含的信息. 在本文工作的基础上, 我们试探性地提出了未来相关研究计划.

Information-theoretical View of Bayesian Analysis with SNIa in Cosmology and DDR Tests

MA Cong

(Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210034)

In this dissertation, we set out to formalize a self-consistent method of testing cosmological models and to obtain the parameter estimations with the luminosity distance and redshift data from the Type Ia supernova (SNIa) standard candles. Using the Bayesian graphical models and their correspondence with the causality, information-flow, and independence properties among the random variables, we construct a method for properly utilizing the parameter-dependent covariance matrix in cosmological model analyses and demonstrate its application to the Joint Light-curve Analysis (JLA) SNIa distance data set. Our statistical method differs, in both concept and effect, from those used in conventional analyses. It is found that the conventional analysis in the form of χ^2 minimization equivalently incurs an implicit distortion on the prior probability of SNIa standardization parameters. The distortion, once revealed explicitly, cannot find convincing justification from statistical or astrophysical considerations. In contrast, our Bayesian method explicitly incorporates prior probabilities of all the model parameters. Our new analysis favors a lower amount of stretch- and color-based corrections to the SNIa magnitude compared with previous studies, and we report a reduced statistical significance of the step correction based on the host galaxy's stellar mass. For w CDM (cold dark matter) and Λ CDM cosmologies, we obtain the revised posterior estimates, showing an increased favor for $w \geq -1$ and higher estimates of Ω_M , respectively.

The same statistical issue is revisited in the task of compressing the SNIa data set and controlling the loss of cosmological information. To this end, we formalize the data compression as a statistical inference problem, and achieve an efficient and accurate Gaussian approximation to the optimal, maximal-entropy solution. Multiple tests based on statistical and information theories are applied to our proposed solution, and they validate its accuracy and robustness. Discrepancies from the JLA original compression exist in the distance modulus estimates and in the structures of their covariance matrix. We use our compressed data set to perform an internal cross-validation of the SNIa distance ladder, and find a nominal color evolution trend that is different from previous results.

Furthermore, we establish the strengths of compressed SNIa data in the distance-duality relation (DDR) tests in the form of more representative data subsamples and more accurate modelling of statistical uncertainties. In combination with auxiliary observational information, the compressed SNIa luminosity distance moduli provide a subsample of measurements matched to and comparable with the angular-diameter distance data from the baryon acoustic oscillation (BAO) surveys. We perform Monte Carlo analyses to evaluate the distributional properties of the DDR discriminant parameter η . The results are consistent with the DDR, and uncertainty bounds are estimated to be about 8% (standard deviation). Under the additional hypothesis of a grey dust component that may systematically diminish SNIa brightness and induce an apparent deviation from DDR, we improve the constraints by incorporating *Planck* cosmological posteriors, and obtain improved bounds on the intergalactic extinction and optical-depth differential. The results are consistent with zero intergalactic extinction with improved bounds. The statistical uncertainty of η based on future survey specifications is estimated to be $\sim 2\%$, highlighting the importance of systematics and the role of consistent statistical methods in response to them.

In the light of our key discoveries, we advocate the methods of statistical analysis of cosmological models based on the Bayesian formalisms, and call for the further study into their efficacy in preserving the information content of more complex cosmological observational data sets expected in the future. Some research topics are tentatively proposed in relation to the present study.