Statistical Methodology 9 (2012) 101–114

Contents lists available at SciVerse ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

# Bayesian methods for analysis and adaptive scheduling of exoplanet observations

Thomas J. Loredo [a,*], James O. Berger [b], David F. Chernoff [a], Merlise A. Clyde [b], Bin Liu [c]

[a] Department of Astronomy, Cornell University, United States
[b] Department of Statistical Science, Duke University, United States
[c] Kuang-Chi Institute of Advanced Technology, China

ARTICLE INFO

*Article history:*
Received 15 October 2010
Received in revised form
30 July 2011
Accepted 30 July 2011

*Keywords:*
Bayesian statistics
Sequential design
Active learning
Monte Carlo methods
Extrasolar planets

ABSTRACT

We describe work in progress by a collaboration of astronomers and statisticians developing a suite of Bayesian data analysis tools for extrasolar planet (exoplanet) detection, planetary orbit estimation, and adaptive scheduling of observations. Our work addresses analysis of stellar reflex motion data, where a planet is detected by observing the "wobble" of its host star as it responds to the gravitational tug of the orbiting planet. Newtonian mechanics specifies an analytical model for the resulting time series, but it is strongly nonlinear, yielding complex, multimodal likelihood functions; it is even more complex when multiple planets are present. The number of dimensions in the model parameter space ranges from a few to dozens, depending on the number of planets in the system, and the type of motion measured (line-of-sight velocity, or position on the sky). Since orbits are periodic, Bayesian generalizations of periodogram methods facilitate the analysis. This relies on the model being linearly separable, enabling partial analytical marginalization, reducing the dimension of the parameter space. Subsequent analysis uses adaptive Markov chain Monte Carlo methods and adaptive importance sampling to perform the integrals required for both inference (planet detection and orbit measurement), and information-maximizing sequential design (for adaptive scheduling of observations). We present an overview of our current techniques and highlight directions being explored by ongoing research.

© 2011 Elsevier B.V. All rights reserved.

* Corresponding author.
*E-mail address:* loredo@astro.cornell.edu (T.J. Loredo).

1572-3127/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.stamet.2011.07.005

## 1. Introduction

In the last fifteen years astronomers have discovered about 500 planetary systems hosted by nearby stars; new systems are announced almost weekly and the pace of discovery is accelerating. The data are now of sufficient quantity and quality that exoplanet science is shifting from being discovery-oriented to focusing on detailed astrophysical modeling and analysis of the growing catalog of observations. Making the most of the data requires new statistical tools that can fully and accurately account for diverse sources of uncertainty in the context of complex models.

Planets are small and shine in reflected light, making them much dimmer than their host stars. With current technology it is not possible to directly image exoplanets against the glare of their hosts except in rare cases of nearby systems with a large planet far from its host star. The vast majority of exoplanets are instead detected indirectly. The most productive technique to date is the *radial velocity (RV) method*, which uses Doppler shifts of lines in the star's spectrum to measure the line-of-sight velocity component of the reflex motion of the star in response to the planet's gravitational pull. Such "to-and-fro wobble" can be measured for planets as small as a few Earth masses, in close orbits around Sun-like stars. Space-based telescopes will soon enable *high-precision astrometry* capable of measuring the side-to-side reflex motion (the *Hubble Space Telescope* Fine Guidance Sensors have already performed such measurements for a few exceptional systems). Another indirect technique measures the diminution of light from the star when a planet *transits* in front of the stellar disk. This requires a fortuitous orbital orientation, but large, space-based transit surveys monitoring many stars are making the transit method increasingly productive.[1] Astronomers are using these techniques to build up a census of the nearby population of extrasolar planetary systems, both to understand the diversity of such systems, and to look for potentially habitable "Earth-like" planets.

Here we focus on analysis of reflex motion data. Motivated by the needs of astrometric and RV campaigns being planned in 2000, Loredo and Chernoff [17,18] described Bayesian approaches for analyzing observations of stellar reflex motion for detection and measurement of exoplanets, and for adaptive scheduling of observations (see also [16] for a pedagogical treatment of adaptive scheduling). This work advocated a fully nonlinear Bayesian analysis based on Keplerian models for the reflex motion (i.e., objects in elliptical orbits around a center of mass). Loredo and Chernoff [17] noted that a subset of the orbital parameters appear linearly and (with appropriate priors) could be marginalized analytically, reducing dimensionality and simplifying subsequent analysis. Building on the work of Jaynes and Bretthorst (see [3] and references therein), they described the close relationship between this approach and periodogram-based methods already in use for planet detection, dubbing the Bayesian counterpart a *Kepler periodogram* (Scargle independently developed similar ideas at the same time, calling the resulting tool a *Keplerogram*; see [20]). They also described how Monte Carlo posterior sampling algorithms could enable implementation of fully nonlinear Bayesian experimental design algorithms for adaptive scheduling of exoplanet observations. Unfortunately, although numerous exoplanets were discovered by the time of [17,18], none of the data were publicly available, and the virtues of the approach could not be demonstrated with real data.

By 2004 several observing teams were at last releasing high-quality exoplanet RV data, fueling new interest in Bayesian algorithms. Independently of earlier work, [7] explored the connections between conventional periodogram methods and the Bayesian approach (albeit hampered by an incorrect marginal likelihood calculation); Cumming and Dragomir [8] later extended this work, reproducing the Kepler periodogram of [17]. Ford [9,10] and Gregory [12] applied classic posterior sampling techniques (Metropolis random walk (MRW) and parallel tempering Markov chain Monte Carlo (MCMC), respectively) to orbit modeling; Gregory's approach uses a control system to tune the proposal parameters in a pilot run, and he has recently augmented his algorithm to include parameter updates based on genetic algorithms [13]. Balan and Lahav [1] applied an early, approximate

[1] During the review of this manuscript, the *Kepler* mission announced the discovery of over 1000 candidate exoplanet systems, observed to have periodic, transit-like events; it is expected that over 90% of these may be confirmed as exoplanets by follow-up RV and other observations.

Download English Version:

https://daneshyari.com/en/article/1151188

Download Persian Version:

https://daneshyari.com/article/1151188

Daneshyari.com