Astrophys Space Sci (2012) 338:35–48
DOI 10.1007/s10509-011-0922-4

ORIGINAL ARTICLE

MaxEnt power spectrum estimation using the Fourier transform for irregularly sampled data applied to a record of stellar luminosity

Robert W. Johnson

Received: 23 September 2011 / Accepted: 5 November 2011 / Published online: 30 November 2011

Abstract The principle of maximum entropy is applied to the spectral analysis of a data signal with general variance matrix and containing gaps in the record. The role of the entropic regularizer is to prevent one from overestimating structure in the spectrum when faced with imperfect data. Several arguments are presented suggesting that the arbitrary prefactor should not be introduced to the entropy term. The introduction of that factor is not required when a continuous Poisson distribution is used for the amplitude coefficients. We compare the formalism for when the variance of the data is known explicitly to that for when the variance is known only to lie in some finite range. The result of including the entropic measure factor is to suggest a spectrum consistent with the variance of the data which has less structure than that given by the forward transform. An application of the methodology to example data is demonstrated.

Keywords Fourier transform · Power spectral density · Irregular sampling · Maximum entropy data analysis

1 Introduction

The analysis of imperfect data is a common task in science. Given a set of measurements sampled over time, one commonly uses the Fourier transform to estimate the power carried by the signal as a function of frequency. The forward transform is commonly viewed as the best estimate of the amplitude and phase associated with basis functions of independent frequency; however, the indiscriminate use of the forward transform is not appropriate when the data are known to be subject to measurement error, and the problem of irregular sampling is often addressed by *ad hoc* methods of varying subjectivity, such as interpolation (Cenker et al. 1991; Malik et al. 2005) or zero-padding (Boyd 1992).

Bayesian statistical inference is a well-established methodology for dealing with imperfect data (Bretthorst 1988; Sivia 1996; Gregory 2005). The parameters of interest, here the amplitude and phase comprising the power spectral density, are related to the data through a model function which may be nonlinear. When that function is invertible, its inverse is usually called the forward transform of the data, but the methodology applies as well to model functions which are not invertible. The most likely values for the parameters are given by those which maximize their joint distribution, which takes into account both their likelihood as measured by the discrepancy between the model and the data and their possibly non-uniform prior distribution. A non-uniform prior commonly represents the invariant Haar measure under a change of variables, and when the number of parameters exceeds the number of data, a prior based on entropic arguments is often employed. These methods generally fall under the rubric of MaxEnt data analysis, as the optimum of the joint distribution minimizes the residual while simultaneously maximizing the entropy distribution. Essentially, we are extending the Lomb-Scargle method (Lomb 1976; Scargle 1982) to incorporate the effect of the measurement errors on the estimate of the most likely amplitude and phase coefficients.

After a brief review of Bayesian data analysis, we investigate the details of its application to power spectral density estimation using Fourier basis functions. The problem of missing data values for otherwise regular sampling is easily addressed by working with basis functions defined only at the measurement times. We will compare the methodol-

---

R.W. Johnson (✉)
Alphawave Research, Atlanta, GA 30238, USA
e-mail: robjohnson@alphawaveresearch.com