Detection of weak periodic signals from irregularly spaced observations

Jörg Pfleiderer and Martin Mössner[1,2]

Inst. of Astronomy, Leopold-Franzens-University, Innsbruck, Austria

Abstract. We perform time series analysis on irregularly spaced data with large gaps. Our algorithm is based on multidimensional nonlinear optimization. Particular emphasis is given to the selection and deletion of frequencies. The final result is checked by various statistical tests.

1 Method. The algorithm performs in an infinite loop, consisting of the steps: Selection, adjustment, and deletion of frequencies. The algorithm stops when the standard deviation of the given data is reached. The computed results are tested by applying various statistical tests.

A set of M data $\mathbf{m} = (m_j)_j$, such as time-dependent brightness data, is to be approximated by $\mathbf{s} : s_j = A_0 + \sum_{i=1}^{N} A_i \cos(2\pi\, f_i t_j - \varphi_i)$. The amplitudes $\mathbf{A}$, frequencies $\mathbf{f}$, phases $\boldsymbol{\varphi}$, as well as the number N of frequencies considered are unknown. The data and, perhaps, the times contain observational errors. The problem is how to choose N and $\mathbf{f}$ and then to compute $\mathbf{A}$ and $\boldsymbol{\varphi}$.

1.1 Selection of Frequencies. Our main criterion for selecting new frequencies is the power function $p(f)$ of the data minus the signal expected from the last iteration. Since $p(f)$ is rapidly oscillating, high resolution is necessary for a clear image. Optimal resolution is given by $\Delta f = 0.1\,\frac{1}{T}$ with $T = t_M - t_1$.

In order to overcome local oscillations we smooth $p(f)$ by computing the maximum within small windows. From this we choose the largest contributions and add them to the frequency vector of the last iteration. The new frequency vector will be used as initial guess in the adjustment step. If the number of frequencies gets too large only the strongest contributions will be accepted.

1.2 Adjustment of Frequencies. Given a start frequency vector $\mathbf{f}$, we use a damped Newton solver (implementation of Deuflhard 1974), to minimize $\|\mathbf{g}(\mathbf{f})\|_2$

$$\text{with} \quad g_j(\mathbf{f}) = a_0(\mathbf{f}) + \sum_{i=1}^{N} a_i(\mathbf{f})\, \cos(2\pi\, f_i t_j) + \sum_{i=1}^{N} b_i(\mathbf{f})\, \sin(2\pi\, f_i t_j) - \mathbf{m}_j .$$

For this we need the amplitudes a_0, a_i, b_i as function of $\mathbf{f}$. The derivatives with respect to $\mathbf{f}$ are computed by numerical differentiation.

The amplitudes for a particular $\mathbf{f}$ can be computed as the solution to the overdetermined linear system $\mathbf{A} \cdot \mathbf{a} \approx \mathbf{m}$ with $\mathbf{A} = (1, \cos(2\pi\, f_i t_j), \sin(2\pi\, f_i t_j))_{j,i}$ and $\mathbf{a} = (a_0(\mathbf{f}), a_i(\mathbf{f}), b_i(\mathbf{f}))_i$. In the full rank case, an unique $\mathbf{a}$, the so-called

---

[1]email: joerg.pfleiderer@uibk.ac.at and martin.moessner@uibk.ac.at

[2]This work was supported by the Austrian Research Foundation under grant P8568-PHY and by the Austrian Academy of Sciences (Space Research / National Programs).

least squares solution (LS), exists, which minimizes $\|\mathbf{Aa} - \mathbf{m}\|_2$. For solving the linear system, we use either the complete orthogonal factorization (COF, e.g. Golub and van Loan 1991) or the total least squares method (TLS, van Huffel and Vandewalle 1991). We do not use normal equations since this leads to bad conditioning. TLS has the advantage that it considers errors both in $\mathbf{A}$ and $\mathbf{b}$.

1.3 Deletion of Frequencies. In practical examples, it happens that the strongest power-amplitude is caused by aliasing effects. Nevertheless, if the correct frequencies are chosen, amplitudes for aliased frequencies get small and therefore can be deleted. In such cases we insert some test values and perform one or two Newton iterations. Frequencies which are *very* closely spaced are substituted by their mean value.

1.4 Grid Search. The main break criterion is the norm of the residual vector. Since it shows strong oscillations in each component of the frequency vector, most optimization techniques tend to catch one of the local minima instead of the absolute minimum. The typical frequency for these oscillations is well known ($\Delta f = \frac{1}{2T}$). Therefore, we do, after each major iteration, some grid search in each component of $\mathbf{f}$.

1.5 Statistical Package (see Dougherty 1990). We begin by computing standard statistics, such as mean, s.d., etc. To test the randomness of the residual vector, we apply the signed-rank or runs test and the Anderson-Darling test (Mössner et al. 1995). For estimating the accuracy of the chosen frequencies we compute linear confidence intervals. The reliability of weak signals can be tested by Fisher's test.

2 Results. The method was assessed by test-data and real astrophysical data: Delta Scuti star θ^2 Tauri. The results of an early version of our method are given in Mössner and Pfleiderer (1993). The results of θ^2 Tauri compare well to the results of Breger (1989) but contain additional periods.

References

Breger, M. et al. 1989, "Multiple close frequencies of the Delta Scuti star θ^2 Tauri. II. The second multisite campaign", A&A, **214**, 209–219

Deuflhard, P. 1974, "A Modified Newton Method for the Solution of Ill-Conditioned Systems of Nonlinear Equations with Applications to Multiple Shooting", Numerische Mathematik, **22**, 289–315

Dougherty, E.R. 1990, "Probability and statistics for the engineering, computing and physical sciences", Prentice Hall International, London

Golub, G.H. and van Loan, Ch.F. 1991, "Matrix Computations", The Johns Hopkins University Press, Baltimore

Mössner, M., Netzer, N., and Pfleiderer, J. 1995, "Efficient computing of the Anderson-Darling Test Statistic", submitted

Mössner, M. and Pfleiderer, J. 1993, "Automatic Analysis for Time-Series with Large Gaps", Proc. of the 5th ESO/ST-ECF Data Analysis Workshop (P.J. Grosbøl and R.C.E. de Ruijsscher, eds.), ESO, 197–202

van Huffel, S. and Vandewalle, J. 1991, "The Total Least Squares Problem. Computational Aspects and Analysis", SIAM, Philadelphia