and I. Obviously, better results are obtained when all colours can be combined in the way we have described.

Acknowledgements

This work would not have been possible without the use of the New Jersey/Warsaw evolutionary code and W Dziembowski's non-adiabatic pulsation code. We are very grateful to both Drs A Pamyatnykh and W Dziembowski for permission to use these codes. We are also very grateful to Dr R L Kurucz for sending us a CDROM with the latest model atmospheres. We would also like to thank Dr C Sterken for allowing us to use his unpublished data for V853 Cen Drs T Moon and M. Dworetsky for supplying their computer code.

References

Balona L.A., Evers E.A., 1999, MNRAS, in press

Breger M., Handler G., Nather R.E., et al., 1995, A&A 297, 473

Breger M., Zima W., Handler G., Poretti E., Shobbrook R.R., Nitta A., Prouton O.R., Garrido R., Rodriguez E., Thomassen T., 1998, A&A 331 271

Fontaine G., Brassard P., Bergeron P., Wesmael F., 1996, ApJ 469, 320

Heynderickx D., Waelkens C., Smeyers P., 1994, A&AS, 105, 447

Kurucz R.L., 1979, ApJS, 40, 1

Viskum M., Kjeldsen H., Bedding T.R., Dall T.H., Baldry I.K.,Bruntt H., Frandsen S., 1998, A&A 335 549

Watson R.D., 1988, ApSS 140, 255

A derivation of the errors for least squares fitting to time series data

M. H. Montgomery[1] and D. O'Donoghue[2]

[1]Astronomisches Institut der Universität Wien, Türkenschanzstr. 17, A–1180 Wien, Austria
email: montgomery@astro.univie.ac.at

[2]South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, South Africa

Abstract

We present analytical results of the effect which random, uncorrelated noise has on a least squares fit of a sinusoidal signal. Our results are as follows:

$$\sigma(a) = \sqrt{\frac{2}{N}}\,\sigma(m)$$

$$\sigma(\phi) = \sqrt{\frac{2}{N}}\,\frac{\sigma(m)}{a}$$

$$\sigma(f) = \sqrt{\frac{6}{N}}\,\frac{1}{\pi T}\,\frac{\sigma(m)}{a},$$

where $\sigma(a)$, $\sigma(\phi)$, and $\sigma(f)$ are the one-sigma errors in the amplitude, phase, and frequency, respectively, due to noise in the observed magnitudes with a root-mean-square deviation of $\sigma(m)$. The variables N and T are the total number of data points and the total time baseline of the observations, respectively. We demonstrate that these formulae yield essentially the same error estimates as those obtained from numerical nonlinear least squares fits when applied to a real data set.

Finally, we discuss the applicability and limitations of these formulae. In particular, we show that correlations in the noise can modify these relations by multiplying the error estimates by a factor greater than unity, with the result that the above formulae may severely underestimate the sizes of the true errors.

Motivation

The calculation of the way in which experimental errors affect the amplitude, frequency, and phase determination of a time series signal is a nontrivial problem. In the general case, the data may not be evenly sampled in time and the noise superimposed on the (assumed) sinusoidal signal may be correlated in time as well as being non-Gaussian. In addition, there may be more than one such sinusoidal signal present, which, depending upon the aliasing, will also add uncertainty to the parameter determination

To treat adequately and systematically the realistic cases which are encountered in astronomy, for example, it is probably necessary to perform numerical simulations which use the same time sampling as the data set and which incorporate a realistic model for the "noise". While this is certainly a project worth doing, in this present paper we wish to compute analytically the best case scenario, i.e., a lower limit on the size of the errors one can expect. Our goal is to collect in one place all the simple formulae which will allow the reader to make these estimates. For more detailed and statistically rigorous treatments of aspects of this problem, we refer to the papers by Schwarzenberg-Czerny (1991, 1999) and Koen (1999, in press).

In the following section, we derive the errors in the amplitude, phase, and frequency for the case when these parameters are determined from a single data set. Another commonly encountered case is one in which the frequency is well-constrained by previous observations and can be considered known, but where the amplitude and phase are yet to be determined. In this case the derivation is somewhat simpler, and it is treated in the appendix of Breger et al. (1999, in press). The formulae derived for the errors in amplitude and phase are identical in both cases.

Derivation of the uncertainties in the amplitude, phase, and frequency

As stated above, we wish to assume the ideal case for our error computations. We therefore assume that we have N measurements of the magnitudes, m_i, which are taken at times t_i, each of which are evenly separated by a time Δt. We assume that the times of the observations are error free, but that the brightness measurements m_i are subject to random errors, Δm_i. These errors are assumed to have an average of zero ($\langle \Delta m_i \rangle = 0$), to have a root-mean-square amplitude which is constant in time ($\langle \Delta m_i^2 \rangle = \sigma^2(m)$), and to be uncorrelated in time ($\langle \Delta m_i \Delta m_j \rangle = 0$ for $i \neq j$).

We now wish to analyze our time series data by fitting a sinusoid to it. Specifically, we fit the function

$$f(t) = a \sin(\omega t_i + \phi),$$

where the amplitude a, phase ϕ, and (angular) frequency ω are yet to be determined. We define

$$\begin{aligned} \chi^2 &\equiv \sum_{i=1}^{N} [m_i - f(t_i)]^2 \\ &= \sum_{i=1}^{N} [m_i - a \sin(\omega t_i + \phi)]^2 , \end{aligned}$$

where the minimum in χ^2 corresponds to the best fit solution of the model parameters.

Minimizing χ^2 with respect to a, ϕ, and ω, we obtain the following three relations:

$$\frac{\partial \chi^2}{\partial a} = 0 \quad \Rightarrow a = \frac{2}{N} \sum_{i=1}^{N} m_i \sin(\omega t_i + \phi) \qquad (1)$$

$$\frac{\partial \chi^2}{\partial \phi} = 0 \quad \Rightarrow 0 = \sum_{i=1}^{N} m_i \cos(\omega t_i + \phi), \qquad (2)$$

$$\frac{\partial \chi^2}{\partial \omega} = 0 \quad \Rightarrow 0 = \sum_{i=1}^{N} m_i t_i \cos(\omega t_i + \phi), \qquad (3)$$

where we have assumed that the time distribution of the data is such that the orthogonality relations $\sum_{i=1}^{N} \sin^2(\omega t_i + \phi) = N/2$ and $\sum_{i=1}^{N} \sin(\omega t_i + \phi) \cos(\omega t_i + \phi) = 0$ represent valid approximations. The above three equations must be satisfied simultaneously by the best fit solution.

In general, the random errors in magnitude, Δm_i, produce small variations in the fit parameters $(\Delta a, \Delta \phi, \Delta \omega)$ from their "true" values. If we take a total differential of equation 1 with respect to (m_i, a, ϕ, ω), we obtain

$$\begin{aligned} \Delta a &= \frac{2}{N} \sum_{i=1}^{N} [\Delta m_i \sin(\omega t_i + \phi) \\ &\quad + m_i \cos(\omega t_i + \phi) \Delta \phi + m_i t_i \cos(\omega t_i + \phi) \Delta \omega] \\ &= \frac{2}{N} \sum_{i=1}^{N} \Delta m_i \sin(\omega t_i + \phi), \end{aligned}$$

where the second and third terms have vanished through the application of equations 2 and 3, respectively. If we square this expression and then take a statistical average, we find

$$\begin{aligned} \langle (\Delta a)^2 \rangle &= \frac{4}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \Delta m_i \Delta m_j \rangle \sin(\omega t_i + \phi) \sin(\omega t_j + \phi) \\ &= \frac{4}{N^2} \sum_{i=1}^{N} \langle (\Delta m_i)^2 \rangle \sin^2(\omega t_i + \phi) \\ &= \frac{4}{N^2} \sigma^2(m) \sum_{i=1}^{N} \sin^2(\omega t_i + \phi) \\ &= \frac{2}{N} \sigma^2(m), \end{aligned}$$

where we have made use of the assumed statistical properties of Δm_i. Writing $\sigma^2(a) = \langle (\Delta a)^2 \rangle$, we find

$$\sigma(a) = \sqrt{\frac{2}{N}} \cdot \sigma(m). \qquad (4)$$

We repeat this analysis in order to find the errors in ϕ and ω. From equation 2, we have

$$\begin{aligned} 0 &= \sum_{i=1}^{N} [\Delta m_i \cos(\omega t_i + \phi) - \\ &\quad m_i \Delta \phi \sin(\omega t_i + \phi) - m_i t_i \Delta \omega \sin(\omega t_i + \phi)] \, . \qquad (5) \end{aligned}$$

Now we must make use of the fact that the signal without noise is just the sinusoidal solution which we are seeking, i.e., $m_i = a\sin(\omega t_i + \phi)$. Using this, we find that equation 5 reduces to

$$\sum_{i=1}^{N} \Delta m_i \cos(\omega t_i + \phi) = \Delta\phi \frac{a}{2} N + \Delta\omega \frac{a}{2} \sum_{i=1}^{N} t_i. \quad (6)$$

Applying this same set of steps to equation 3, we obtain

$$\sum_{i=1}^{N} \Delta m_i t_i \cos(\omega t_i + \phi) = \Delta\phi \frac{a}{2} \sum_{i=1}^{N} t_i + \Delta\omega \frac{a}{2} \sum_{i=1}^{N} t_i^2. \quad (7)$$

Equations 6 and 7 form a system of linear equations for the errors $\Delta\phi$ and $\Delta\omega$. These equations can be greatly simplified by a specific choice of the zero point in time. We choose the zero point to be the "average time", so we require

$$\sum_{i=1}^{N} t_i = 0.$$

This choice has the advantage of decoupling equations 6 and 7. Using this zero point, the sum in the coefficient of $\Delta\omega$ in equation 7 becomes

$$\sum_{i=1}^{N} t_i^2 = \Delta t^2 \sum_{i=1}^{N} (i - N/2)^2 = \frac{\Delta t^2}{12} N^3 + O(N^2).$$

Assuming that $N \gg 1$, we retain only the leading term in N. It is now straightforward to solve equations 6 and 7 for $\Delta\phi$ and $\Delta\omega$:

$$\Delta\phi = \frac{2}{aN} \sum_{i=1}^{N} \Delta m_i \cos(\omega t_i + \phi)$$

$$\Delta\omega = \frac{24}{aN^3 \Delta t^2} \sum_{i=1}^{N} \Delta m_i t_i \cos(\omega t_i + \phi)$$

If we now square both sides of the above equations, perform a statistical average, and then a summation over i, we find that

$$\sigma^2(\phi) = \frac{2}{N} \frac{\sigma^2(m)}{a^2} \quad (8)$$

$$\sigma^2(\omega) = \frac{24}{N^3 \Delta t^2} \frac{\sigma^2(m)}{a^2}, \quad (9)$$

where we have written $\sigma^2(\phi) \equiv \langle (\Delta\phi)^2 \rangle$ and $\sigma^2(\omega) \equiv \langle (\Delta\omega)^2 \rangle$. Rewriting these relations in terms of the actual frequency ($\omega = 2\pi f$) and taking a square root, we find that

$$\sigma(\phi) = \sqrt{\frac{2}{N}} \frac{\sigma(m)}{a} \quad (10)$$

$$\sigma(f) = \sqrt{\frac{6}{N}} \frac{1}{\pi T} \frac{\sigma(m)}{a}, \quad (11)$$

where $T \equiv N\Delta t$ is the total time length of the data set.

We note that if a different zero point for the time t_i is chosen, then the errors in ϕ and ω are no longer uncorrelated, which has the effect of changing the derived errors in ϕ. For instance, if we choose the beginning of the run to be the zero point (i.e., $t_1 = 0$), then the derived error for the phase, $\sigma(\phi)$, is exactly twice the value given by equation 10. However, the relations for the errors in the amplitude and frequency are unchanged, as must be the case.

Equivalence to Numerical Least Squares for a Real Data Set

These simple formulae should of course be equivalent to the errors obtained from a numerical nonlinear least squares analysis. Put another way, such an analysis should have all the shortcomings of the one we performed in the previous section, with respect to the assumptions about the ideal nature of the errors.

To demonstrate this, we therefore compared the above formulae with the results of a routine based on the Levenberg-Marquardt method, as described in Press et al. (1992). And to add a dose of "realism", we took the times $\{t_i\}$ from the Strømgren y observations of the star 4CVn obtained in the 1996 Delta Scuti campaign.

In Table 9, we show the ratio of the errors computed with our formulae to those computed with the nonlinear least squares routine. The columns labeled `freq`, `amp`, and `phase` are the relevant ratios of the errors for the frequencies, amplitudes, and phases. The nonlinear analysis involved a simultaneous fit for the 6 largest amplitude frequencies in this data set. This is technically not a completely valid comparison with our formulae, since we considered fits for only one sinusoidal component. Furthermore, our formulae were derived assuming that no other sinusoidal signals were present, which is clearly not the case in this data set. By fitting the 6 largest frequencies simultaneously, we believe that we actually do a better job of removing the different frequencies' effects on each other, and therefore of sim-

Table 9: Comparison of analytical and numerical results

Mode Frequency (c/d)	Ratio of Analytic to Numerical Errors freq	amp	phase
8.59	0.912	1.001	0.934
7.37	0.887	0.962	0.925
5.05	0.924	0.953	0.924
6.12	0.886	0.951	0.907
5.85	0.916	0.948	0.904
5.53	0.904	0.957	0.919

ulating the conditions under which our expressions for the errors are valid.

As can be seen from Table 9, the ratio of the errors calculated in these two ways is fairly close to 1.0, and in fact deviates no more than 12% from this value. First of all, this shows that the coverage in the 1996 campaign was sufficient so that our assumptions about the orthogonality of the sinusoids was valid. Second, it serves as a reminder that the simplified assumptions about the noise which we have made are also being made by the nonlinear least squares procedure. Therefore, the numerically obtained results should not be regarded in any sense as more realistic or "better". Thus, the errors computed with either method just provide a lower limit on the size of the true errors. We therefore urge extreme caution in their application and use.

Correlated Noise

The reason that our formulae represent only lower bounds on the errors is that our assumptions about the properties of the noise may be false. In particular, we expect that in general the errors in the observed magnitudes will be correlated in time, due to transparency variations in the Earth's atmosphere, for instance. Schwarzenberg-Czerny (1991) shows how the error formulae may be modified in this case, and we re-derive his result below.

In section 2, the assumption of uncorrelated noise was expressed by the relation $\langle \Delta m_i \Delta m_j \rangle = 0$ for $i \neq j$. For correlated noise this expression can be non-zero if i and j are *close enough* to one another, i.e., within a "correlation length". One simple choice for the correlation function would be

$$\langle \Delta m_i \Delta m_j \rangle = \begin{cases} \sigma^2(m) & \text{if } |t_i - t_j| \leq D\Delta t/2 \\ 0 & \text{otherwise} \end{cases},$$

where D is an estimate of the number of consecutive data points which are correlated. Instead of the above, we choose the following form for the correlation function of the magnitude fluctuations:

$$\langle \Delta m_i \Delta m_j \rangle = \sigma^2(m)\; e^{-\frac{(t_i - t_j)^2}{(\Delta t D/2)^2}}$$

If we use this expression for the correlation to compute $\sigma^2(\omega) \equiv \langle \omega^2 \rangle$, then, after some manipulations and making the same approximations as before, we find that

$$\sigma^2(\omega) = \sigma_0^2(\omega) \cdot A(\omega, D),$$

where $\sigma_0^2(\omega)$ is the variance of ω in the uncorrelated case as given by equation 9, and the function A is

$$A(\omega, D) = D\,\frac{\sqrt{\pi}}{2}\; e^{-\left(\frac{D\Delta t \omega}{4}\right)^2}. \qquad (12)$$

If we take these correlations into account for the errors in phase and amplitude, we find that the variance of these quantities is also multiplied by the factor $A(\omega, D)$.

For correlations in time which are much shorter than the period of the signal, i.e., $D\omega\Delta t \ll 1$, we have $A \approx D$. This is the result found by Schwarzenberg-Czerny (1991), in his equation 25, for example (this is sometimes referred to as "red noise"). In the other limit, $D\omega\Delta t \gg 1$, the function A approaches zero and the errors vanish. This occurs because the errors in magnitude have become completely correlated in time and now just represent a constant offset to all the data points, which produces no errors in the least squares fits. Finally, we note that A is a maximum when $D\Delta t = 2\sqrt{2}/\omega \approx 0.5\,P$, i.e., when the correlation time is of order the period of the signal, which, of course, is to be expected. In this case, the value of A is $A_{\rm max} \approx 0.24\,P/\Delta t$. For observations of pulsating white dwarfs, the periods may be as high as 1000 sec and the sampling may be every 10 sec, so that $A_{\rm max} \approx 24$. Thus, the error estimates derived in the second section could be too low by a factor of $\sqrt{A} \approx 5$ if the errors are correlated in the way we have assumed.

To summarize, the effect of correlations in the noise is in general to increase the magnitude of the errors. For correlations in time which are short compared to the period of the signal, the result is to multiply the previously obtained variances by D, the number of consecutive data points which are correlated. Equivalently, this means multiplying the equations for the errors themselves, equations 4, 8, and 9, by a factor of $\sqrt{D}$. As

Schwarzenberg-Czerny (1991) points out, it is possible to obtain an estimate for D from a fit to the data, either visually, or by calculating the autocorrelation function of the residuals and fitting an appropriate function (such as a Gaussian) of width D to it.

Discussion

We hope that the results of the previous section have convinced the reader that the naive least squares formulae provide only a lower limit to the errors, and that the true errors may be much higher. In practice, we have found that the actual errors in frequency can in some cases be a factor of 10 or more larger than those indicated by the formal least squares fits.

The effect of correlations in the data is to increase the size of the errors. When the correlation time of the noise is short compared to the period of the signal, the least squares error estimates should be multiplied by the square root of the number of consecutive points which are correlated, $\sqrt{D}$. For noise with a correlation time of order the signal period, a more precise model for the correlations would be required. Of course in real data, we expect the noise to be comprised of several components, each with its own correlation time, as well as a component due to white noise.

In conclusion, we wish to re-iterate that the naive formulae of the second section represent a lower bound only to the errors, and we urge strong skepticism in the application of these equations. If reliable error estimates are needed, then the correlations in the residuals must be examined in more detail before any statements can be made about the true sizes of the errors.

Acknowledgements. We would like to thank M. Breger and D. Kurtz for useful discussions.

References

Breger et al. 1999, A&A, in press
Koen, C. 1999, MNRAS, in press
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in Fortran 77 (Cambridge, New York, Melbourne: Cambridge University Press)
Schwarzenberg-Czerny, A. 1991, MNRAS, 253, 198
Schwarzenberg-Czerny, A. 1999, APJ, 516, 315

The new period determination of the δ Scuti star HR 5437

Liu Zong-Li, Jiang Shi-Yang, Zhou Ai-Ying, Liu Yan-Ying, Li Zhi-Ping
Beijing Astronomical Observatory, Chinese Academy of Sciences, 100012 Beijing, China

The photoelectric observations of HR 5437 were carried out between 1986 and 1988 at the Piszkéstetö Station of Konkoly Observatory of Hungary and the Xinglong Station of Beijing Observatory of China. Using the data obtained at the Xinglong Station in 1988 Jiang and Li (1988) discovered HR 5437 to be a δ Scuti variable star. They did not derive frequencies. From the 1988 data set obtained at the Piszkésteto Station and the Xinglong Station, Paparó et al. (1990) determined two frequencies. The first frequency is 8.437 cd^{-1}, the second is 12.362 cd^{-1}. Joining the 1988 data set and the data obtained at the Piszkéstetö Station in 1986-1987, Paparó et al. (1990) gave two other frequencies. The first is 8.5367 cd^{-1}, the second is 11.3759 cd^{-1}. Paparó et al. pointed out that due to the poor quality of the data distribution and the complex structure of the spectral window the precise values of frequencies were not found. Li and Jiang (1992) obtained only one frequency of 11.38993 cd^{-1} from their data obtained in 1989. However, comparing with the results given by Paparó et al. (1990), Li and Jiang (1992) thought that the frequency of 8.5367 cd^{-1} given by Paparó et al. (1990) was the fundamental frequency, the frequency of 11.38993 cd^{-1} was the first overtone frequency. The situation is quite confused.

In order to obtain the correct pulsation frequencies of HR 5437 and study its character we observed this star from 8 to 17 April 1997, and from 6 to 8 March 1998 using the 85 cm telescope at the Xinglong Station of Beijing Astronomical Observatory. The 4-channel photometer (Michel and Chevreton 1991) with Strömgren v filter was used. The star SAO 16408 was used as comparison, and SAO 16394 was used as check star. The three stars and sky background were observed simultaneously. The integration time was one minute. No evidence for any variability of SAO 16408 was found. Seven tracks of data were obtained for HR 5437 from 1997 to 1998 covering a period of 334 days. The light curves are shown in Fig. 9, where the ordinate is the v magnitude difference normalized to zero. The abscissa