PARAMETER ESTIMATION OF THE EXPONENTIALLY DAMPED SINUSOID FOR NOISY SIGNALS*

T. Pelaia II, Oak Ridge National Lab, Oak Ridge, TN 37831, USA

Abstract

The damped sinusoid equation is a common model for many scientific processes involving damped periodic signals. Here we present two methods for estimating the damped sinusoid parameters for noisy signals. Both methods are based upon an exact, closed form solution to fit the parameters for signals without noise, and they estimate the parameters for the noisy signals by the statistical maximum likelihood criterion. The first method relies on an optimizer to minimize the mean square signal error. The second method estimates the parameters by direct calculation and is suitable when the signal noise is small and the frequencies are sufficiently far from the integer and half integer values.

INTRODUCTION

The damped sinusoid equation with regularly spaced signals is often used to model physical phenomena involving damped periodic signals. For example, at the Spallation Neutron Source (SNS) in Oak Ridge, the beam position monitors in the accumulator ring [1] provide waveforms of turn by turn beam position signals. These waveforms provide information about the beam dynamics for directly determining the tune and closed orbit and indirectly for measuring the Courant-Snyder beta function [2] and the momentum spread. Since the damping of the signals is dominated by decoherence due to chromaticity and the synchrotron period is long compared to the storage time, the actual signals are better modeled using a gaussian damped sinusoid equation. However, for fitting the tune, phase, amplitude and closed orbit, the exponentially damped sinusoid is a sufficient approximation under suitable conditions, and can be used for priming a gaussian damped sinusoid fit.

Equation 1 shows the general form of the exponentially damped sinusoid with amplitude α, growth rate λ, angular frequency μ, phase φ and offset b. The waveform signal q_n is evaluated at the index (often associated with time) given by the integer, n.

$$q_n = \alpha\, e^{\lambda n} \sin\left(\mu n + \phi \right) + b \qquad (1)$$

Common approaches to fitting waveform data to a damped sinusoid include adaptations of Prony's method [3] and differential algebra [4], but neither of these approaches solve the problem as stated. A search of the literature indicates a lack of a closed form method to directly fit a waveform of discrete, evenly spaced signals to a damped sinusoid regardless of whether noise is present. This paper presents a derivation of a closed form solution to the problem that is used as the foundation of two parameter estimation methods for noisy signals.

EXACT PARAMETER FIT

The exact, closed form solution fitting waveform signals to the damped sinusoid when there is no noise is derived here. The damped sinusoid is recast in equation 2, and the parameters of equation 1 are recovered later. For consistency with code, zero based indices are used.

$$q_n = \left(A \cos \mu n + B \sin \mu n \right) e^{\lambda n} + b \qquad (2)$$

After some algebra, one obtains the equation:

$$\cos \mu = \frac{\left(q_n - b\right) e^{\lambda} + \left(q_{n+2} - b\right) e^{-\lambda}}{2\left(q_{n+1} - b\right)} \qquad (3)$$

Since this equation for the frequency is true for any three consecutive signals, equating the solutions for two consecutive sets of three signals and solving for the growth factor yields:

$$e^{2\lambda} = \frac{\left(q_{n+3} - b\right)\left(q_{n+1} - b\right) - \left(q_{n+2} - b\right)^2}{\left(q_n - b\right)\left(q_{n+2} - b\right) - \left(q_{n+1} - b\right)^2} \qquad (4)$$

Since the growth factor is independent of n, we equate it for two consecutive sets using the following definitions.

$$\begin{aligned} u_n &\equiv 2q_{n+1} - q_n - q_{n+2} \\ v_n &\equiv q_n q_{n+2} - q_{n+1}^2 \end{aligned} \qquad (5)$$

This substitution yields the following relation.

$$\begin{aligned} 0 = &\left(u_{n+1}^2 - u_n\, u_{n+2}\right) b^2 \\ &+ \left(2u_{n+1}\, v_{n+1} - u_n\, v_{n+2} - u_{n+2}\, v_n\right) b \\ &+ v_{n+1}^2 - v_n\, v_{n+2} \end{aligned} \qquad (6)$$

To simplify the expressions, the following definitions are made for the coefficients of the quadratic equation.

$$\begin{aligned} &R_n\, b^2 + S_n\, b + T_n = 0 \\ &R_n \equiv u_{n+1}^2 - u_n\, u_{n+2} \\ &S_n \equiv 2\, u_{n+1}\, v_{n+1} - u_n\, v_{n+2} - u_{n+2}\, v_n \\ &T_n \equiv v_{n+1}^2 - v_n\, v_{n+2} \end{aligned} \qquad (7)$$

* This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

ISBN 978-3-95450-122-9

Combining two quadratic equations (for n and n+1) and eliminating the quadratic term yields the following equation for the constant offset, b.

$$b = \frac{R_n T_{n+1} - R_{n+1} T_n}{R_{n+1} S_n - R_n S_{n+1}} \quad (8)$$

This offset can be fed to equation 4 to recover the growth factor which subsequently can be used in equation 3 to recover the frequency.

The cosine and sine amplitudes of equation 2 can be computed by applying the equation twice (once for n and once for n+1) and solving for the unknown amplitudes. Performing algebra and applying some trigonometry yields the solution for the sine amplitude, B.

$$B = \frac{(q_{n+1} - b) e^{-\lambda(n+1)} \cos \mu n - (q_n - b) e^{-\lambda n} \cos \mu (n+1)}{\sin \mu} \quad (9)$$

The solution for the cosine amplitude, A is as follows.

$$A = \frac{(q_n - b) e^{-\lambda n} - B \sin \mu n}{\cos \mu n} \quad (10)$$

The amplitude in equation 1 is recovered as:

$$\alpha = \sqrt{A^2 + B^2} \quad (11)$$

The tangent of the phase is A over B, and the signs of these two amplitudes determines the quadrant with the phase determined uniquely in the range -π to π.

$$\tan \phi = A / B \quad (12)$$

MAXIMUM LIKELIHOOD ESTIMATION

Equation 1 can be recast to include the signal errors.

$$q_n + \varepsilon_n = (A \cos \mu n + B \sin \mu n) e^{\lambda n} + b \quad (13)$$

Updating equation 3 and rearranging terms leads to the following recursion for signal errors.

$$\begin{aligned} \varepsilon_{n+2} &= 2e^{\lambda} \cos \mu \,(q_{n+1} + \varepsilon_{n+1} - b) \\ &\quad - (q_n + \varepsilon_n - b) e^{2\lambda} + b - q_{n+2} \end{aligned} \quad (14)$$

For independent, random and normally distributed signal error with a common variance (conditions consistent with the application to beam position waveforms in our accumulator ring), the fit of maximum likelihood minimizes the penalty function which is the sum of the square of all the signal errors [5].

Since the recursion for each error is linear in the two previous errors, the error for any waveform element can be expressed as a linear combination of the first two errors.

$$\varepsilon_n = f_n \varepsilon_0 + g_n \varepsilon_1 + h_n \quad (15)$$

Substituting this equation into the recursion equation for errors, yields a recursion for the coefficients.

$$\begin{aligned} f_n &= 2e^{\lambda} \cos \mu \, f_{n-1} - e^{2\lambda} f_{n-2} \\ g_n &= 2e^{\lambda} \cos \mu \, g_{n-1} - e^{2\lambda} g_{n-2} \\ h_n &= 2e^{\lambda} \cos \mu \, h_{n-1} - e^{2\lambda} h_{n-2} + b - q_n \\ &\quad + 2(q_{n-1} - b) e^{\lambda} \cos \mu + (b - q_n) e^{2\lambda} \\ f_0 &= 1, f_1 = 0, g_0 = 0, g_1 = 1, h_0 = h_1 = 0 \end{aligned} \quad (16)$$

Using this recursion equation for coefficients, the penalty function (mean square signal error) can be expressed over the N signals as:

$$P = \frac{1}{N} \sum_{n=0}^{N-1} \varepsilon_n^2 = \frac{1}{N} \sum_{n=0}^{N-1} (f_n \varepsilon_0 + g_n \varepsilon_1 + h_n)^2 \quad (17)$$

The penalty function is minimized in the standard way. The first two signal errors can be solved explicitly.

$$\varepsilon_0 = \frac{\sum_{n=0}^{N-1} h_n g_n \sum_{n=0}^{N-1} f_n g_n - \sum_{n=0}^{N-1} g_n^2 \sum_{n=0}^{N-1} f_n h_n}{\sum_{n=0}^{N-1} f_n^2 \sum_{n=0}^{N-1} g_n^2 - \sum_{n=0}^{N-1} f_n g_n} \quad (18)$$

$$\varepsilon_1 = \frac{\sum_{n=0}^{N-1} h_n f_n \sum_{n=0}^{N-1} f_n g_n - \sum_{n=0}^{N-1} f_n^2 \sum_{n=0}^{N-1} g_n h_n}{\sum_{n=0}^{N-1} f_n^2 \sum_{n=0}^{N-1} g_n^2 - \sum_{n=0}^{N-1} f_n g_n} \quad (19)$$

The five parameter search space of the damped sinusoid has been reduced to three (frequency, growth factor and offset). An optimizer can be used to search just this parameter space to minimize the penalty function.

DIRECT PARAMETER ESTIMATION

A direct method is sought to efficiently provide fits without the need for an optimizer. Generally, the value for a k indexed parameter, ξ_k, can be expressed as a function, θ_k of the sample, s, and the noisy signals.

$$\xi_k = \theta_k (s, q_0 + \varepsilon_0, \ldots q_{N-1} + \varepsilon_{N-1}) \quad (20)$$

If the signal errors are small and the function is well behaved (continuously differentiable with respect to the signals over the region of interest), the function can be expanded to first order in the signal errors resulting in an approximation of the parameter.

$$\xi_k \approx \theta_k(s, q_0, \ldots q_{N-1}) + \sum_{n=0}^{N-1} \frac{\partial \theta_k(s, q_0, \ldots q_{N-1})}{\partial q_n} \varepsilon_n \quad (21)$$

Assuming purely statistical signal errors, we can take the expectation value of the square error of the fit and solve for the variance of the signal error to first order over all samples. In the vicinity of the best parameter fit, the following penalty function approximates the signal error variance where N_s is the number of samples.

$$P_k = \frac{1}{N_s} \sum_{s=0}^{N_s-1} \frac{\left(\xi_k - \theta_k(s, q_0, \ldots q_{N-1})\right)^2}{\rho_{ks}} \quad (22)$$

Here, the following definition has been made and referred to as the "estimation sensitivity."

$$\rho_{ks} = \sum_{n=0}^{N-1} \left(\frac{\partial \theta_k(s, q_0, \ldots q_{N-1})}{\partial q_n} \right)^2 \quad (23)$$

The penalty function is as expected from standard error analysis. The best estimate of the parameter minimizes this penalty function and is calculated to be the weighted mean over the samples where the weight is the estimation sensitivity.

$$\xi_k \approx \frac{\sum_{s=0}^{N_s-1} \left(\frac{\theta_k(s, q_0, \ldots q_{N-1})}{\rho_{ks}} \right)}{\sum_{s=0}^{N_s-1} \frac{1}{\rho_{ks}}} \quad (24)$$

This parameter estimation formula is quite general and is useful if the signal errors are sufficiently small as quantified by standard error propagation.

The exact parameter fit equations derived earlier can be used to explicitly calculate the parameters of the damped sinusoid for a given sample and are suitable as parameter functions for the parameter estimation equation. This method requires the first order differentials with respect to the waveform signals to be calculated. Due to the complexity of the parameter estimator for an individual sample, a Java class in XAL [6, 7] was developed to calculate the exact first order differentials. It has methods to efficiently propagate the first derivatives for several mathematical operations including all operations required for this problem.

Care must be taken to avoid estimation failure for sensitive conditions. The growth factor must be positive, and any sample resulting in a negative growth factor is removed. If the best fit for the cosine of the angular frequency is less than negative one, the angular frequency is set to π, and if it is greater than one, it is set to zero. The presence of a sine of the angular frequency appears in the denominator of the sine amplitude, B, which is catastrophic at the bounds. To avoid catastrophe, B is set to zero near these frequencies which has the effect of ignoring the sine like contributions which approach zero.

As the frequency approaches an integer or half integer, the estimation fails since the cosine term approaches its limits. At zero frequency there is a degeneracy since the phase and amplitude become indeterminate.

This direct method was coded into a Java class in XAL to provide fast estimation and to prime the least squares optimizer with initial values if better accuracy is desired. The direct method calculates the exact fit when the noise level is zero.

CONCLUSION

An exact, closed form solution to compute the parameters of a damped sinusoid from error free waveform data was derived. This solution was adapted for an efficient method to directly estimate the frequency, offset, growth rate, amplitude and phase from noisy signals. This direct method can be used alone when realtime analysis is appropriate and the noise is low and the frequency is sufficiently far away from integer and half integer values. For greater accuracy, the direct parameter estimations can be used to prime an optimizer that minimizes the sum of the square of signal errors using an efficient recursion to calculate these errors.

ACKNOWLEDGMENT

The author would like to thank Dr. Willem Blokland and Dr. Andrei Shishlo for references and discussions on the topic of parameter estimation for the damped sinusoid. The author would also like to thank Dr. Maolian Chen for her feedback on this paper.

REFERENCES

[1] M.A. Plum, "Commissioning Of The Spallation Neutron Source Accelerator Systems," Proceedings of PAC07, Albuquerque, New Mexico, 2007.

[2] S. Henderson et al., "Techniques For Measurement And Correction Of The SNS Accumulator Ring Optics," Proceedings of 2005 Particle Accelerator Conference, Knoxville, Tennessee, 2005

[3] Ramdaus Kumaresand et al., "A Prony method for noisy data: Choosing the signal components and selecting the order in exponential signal models," Proceedings of the IEEE, vol. 72 issue 2, June 2005, pp. 230 – 233.

[4] M. Mboup, "Parameter Estimation via Differential Algebra and Operational Calculus," inria-00138294, version 1, March 2007

[5] H. W. Sorenson, "Least-squares estimation: from Gauss to Kalman," IEEE Spectrum, vol. 7, July 1970, pp. 63 – 68.

[6] J. Galambos et al., "XAL Application Programming Structure," Proceedings of PAC 2005, Knoxville, TN 2005

[7] Open XAL Project; http://xaldev.sourceforge.net/

ISBN 978-3-95450-122-9