

Statistical Signal Processing

Debasis Kundu ¹

Signal processing may broadly be considered to involve the recovery of information from physical observations. The received signals is usually disturbed by thermal, electrical, atmospheric or intentional interferences. Due to the random nature of the signal, statistical techniques play an important role in signal processing. Statistics is used in the formulation of appropriate models to describe the behavior of the system, the development of appropriate techniques for estimation of model parameters, and the assessment of model performances. Statistical Signal Processing basically refers to the analysis of random signals using appropriate statistical techniques. The main purpose of this article is to introduce different signal processing models and different statistical and computational issues involved in solving them.

THE MULTIPLE SINUSOIDS MODEL:

The multiple sinusoids model may be expressed as

$$y(t) = \sum_{k=1}^M \{A_k \cos(\omega_k t) + B_k \sin \omega_k t\} + n(t); \quad t = 1, \dots, N. \quad (1)$$

Here A_k 's and B_k 's represent the amplitudes of the signal, ω_k 's represent the real radian frequencies of the signals, $n(t)$'s are error random variables with mean zero and finite variance. The assumption of independence of the error random variables is not that critical to the development of the inferential procedures. The problem of interest is to estimate the unknown parameters $\{A_k, B_k, \omega_k\}$ for $k = 1, \dots, M$, given a sample of size N . In practical applications often M is also unknown. Usually, when M is unknown, first estimate M using some model selection criterion, and then it is assumed that M is known, and estimate the amplitudes and frequencies.

¹Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. Part of this work has been supported by a grant from the Department of Science and Technology, Government of India.

The sum of sinusoidal model (1) plays the most important role in the Statistical Signal Processing literature. Most of the periodic signals can be well approximated by the model (1) with the proper choice of M and with the amplitudes and frequencies. For several applications of this model in different fields see Brillinger (1987).

The problem is an extremely challenging problem both from the theoretical and computational points of view. As a statistician Fisher (1929) first considered this problem. It seems that the standard least squares estimators will be the natural choice in this case, but finding the least squares estimators, and establishing their properties are far from trivial issues. Although, the model (1) is a non-linear regression model, but the standard sufficient conditions needed for the least squares estimators to be consistent and asymptotically normal do not hold true in this case. Special care is needed in establishing the consistency and asymptotic normality properties of the least squares estimators, see for example Hannan (1973) and Kundu (1991) in this respect. Moreover, for computing the least squares estimators, most of the standard techniques like Newton-Raphson or its variants do not often converge even from good starting values. Even if it converges, it may converge to a local minimum rather than a global minimum due to highly non-linear nature of the least squares surface. Special purpose algorithms have been developed to solve this problem.

Several approximate solutions have been suggested in the literature. Among several approximate estimators, Forward Backward Linear Prediction (FBLP) and modified EquiVariance Linear Prediction (EVLV) work very well. But it should be mentioned that none of these methods behaves uniformly better than the other. More than 200 references on this topic can be found in Stoica (1993), and see also Quinn and Hannan (2001), the only monograph written by statisticians in this topic.

TWO-DIMENSIONAL SINUSOIDAL MODEL

Two dimensional periodic signals are often being analyzed by the two-dimensional sinusoidal model, which can be written as follows:

$$y(s, t) = \sum_{k=1}^M \{A_k \cos(\omega_k s + \mu_k t) + B_k \cos(\omega_k s + \mu_k t)\} + n(s, t), \quad s = 1, \dots, S, \quad t = \dots, T. \quad (2)$$

Here A_k 's and B_k 's are amplitudes and ω_k 's and μ_k 's are frequencies. The problem once again involves the estimation of the signal parameters namely A_k 's, B_k 's, ω_k 's and μ_k 's from the data $\{y(s, t)\}$.

The model (2) has been used very successfully for analyzing two dimensional gray texture data, see for example Zhang and Mandrekar (2001). A three dimensional version of it can be used for analyzing color texture data also, see Prasad (2009). Some of the estimation procedures available for the one-dimensional problem may be extended quite easily to two or three dimensions. However, several difficulties arise when dealing with high dimensional data. There are several open problems in multidimensional frequency estimation, and this continues to be the active area of research.

ARRAY MODEL

The area of array processing has received a considerable attention in the past several decades. The signals recorded at the sensors contain information about the structure of the generating signals including the frequency and amplitude of the underlying sources. Consider an array of P sensors receiving signals from M sources ($P > M$). The array geometry is specified by the applications of interest. In array processing, the signals received at the i -th sensor is given by

$$y_i(t) = \sum_{j=1}^M a_i(\theta_j) x_j(t) + n_i(t), \quad i = 1, \dots, P. \quad (3)$$

Here $x_j(t)$ represents the signal emitted by the j -th source, and $n_i(t)$ represents additive

noise. The model (3) may be written in the matrix form as;

$$y(t) = [a(\theta_1) : \cdots : a(\theta_M)] x(t) + n(t) = A(\theta)x(t) + n(t), \quad t = 1, \cdots, N. \quad (4)$$

The matrix $A(\theta)$ has a Vandermonde structure if the underlying array is assumed to be uniform linear array. The signal vector $x(t)$ and the noise vector $n(t)$ are assumed to be independent and zero mean random processes with covariance matrices Γ and $\sigma^2 I$ respectively. The main problem here is to estimate the signal vector θ , based on the sample $y(1), \cdots, y(N)$, when the structure of A is known.

Interestingly, instead of using the traditional maximum likelihood method, different subspace fitting methods, like Multiple Signal Classification (MUSIC) and Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) and their variants are being used more successfully, see for example the text by Pillai (1989) for detailed descriptions of the different methods.

For basic introduction of the subject the readers are referred to Kay (1987) and Srinath *et al.* (1996) and for advanced materials see Bose and Rao (1993) and Quinn and Hannan (2001).

References

- [1] Bose, N.K. and Rao, C.R. (1993), *Signal Processing and its Applications*, Handbook of Statistics, vol. 10, North-Holland, Amsterdam.
- [2] Brillinger, D. (1987), "Fitting cosines: some procedures and some physical examples", in *Applied Statistics, Stochastic Processes and Sampling Theory*, eds. MacNeill, I.B. and Umphrey, G.J., Reidel, Dordrecht.

- [3] Fisher, R. A. (1929), "Tests of significance in Harmonic analysis", *Proceedings of Royal Society of London: Series A*, vol. 125, 54 - 59.
- [4] Hannan, E.J. (1973), "The estimation of frequencies", *Journal of Applied Probability*, vol. 10, 510 - 519.
- [5] Kay, S. M. (1987), *Modern Spectral Estimation*, Prentice Hall, New York, NY.
- [6] Kundu, D. (1997), "Asymptotic theory of the least squares estimators of sinusoidal signal", *Statistics*, vol. 30, 221 - 238.
- [7] Pillai, S.U. (1989), *Array Signal Processing*, Springer-Verlag, New York, NY.
- [8] Prasad, A. (2009), *Some non-linear regression models and their applications in statistical signal processing*, Ph.D. thesis, Indian Institute of Technology Kanpur, India.
- [9] Quinn, B.G. and Hannan, E.J. (2001), *The estimation and tracking of frequency*, Cambridge University Press, Cambridge, UK.
- [10] Srinath, M.D., Rajasekaran, P.K. and Viswanathan, R. (1996), *Introduction to Statistical Processing with Applications*, Prentice-Hall, Englewood Cliffs, New Jersey.
- [11] Stoica, P. (1993), "List of references on spectral estimation", *Signal Processing*, vol. 31, 329-340.
- [12] Zhang, H. and Mandrekar, V.(2001), "Estimation of hidden frequencies for 2-D stationary processes", *Journal of Time Series Analysis*, 613 - 629.