A Brief Overview on Least Square Spectral Analysis and Power Spectrum

Ingudam Gomita, Sandeep Chauhan and Raj Kumar Sagar

Abstract LSSA, i.e., least square spectral analysis method can be used to calculate the power density function of a fully populated covariance matrix. It fulfills the limitation cause by the fast Fourier transform (FFT), i.e., equally spaced and equally weighted time series. This time of experimental time series used LSSA (which used projection theorem) so that it can also analyze the time series, which is unequally spaced and unequally weighted. Signal-to noise ratio (SNR) is used to find out the probability density function.

Keywords LSSA · Power density function · Signal-to-noise ratio · Projection theorem

1 Introduction

An engineer or scientist cannot be guaranteed that they have an experimental time series with no gaps. Time series can be rendered due to machine failure or may be due to the weather or other natural factors. So the main problem is how to handle a time series when a certain time series has gaps or datum shifts, etc. The obtained time series from an experiment say weather observation is unequally spaced or to avoid aliasing, a variable sampling rate has been introduced. Besides, the instrument development advances and physical phenomenon is understood better so that data accuracy is improved.

Ingudam Gomita (✉) · Sandeep Chauhan · R.K. Sagar
Amity University, Noida, India
e-mail: Ingudamgomita007@gmail.com

Sandeep Chauhan
e-mail: chauhansandee@gmail.com

R.K. Sagar
e-mail: rksagar@amity.edu

M. Pant et al. (eds.), *Proceedings of Fifth International Conference on Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing 437, DOI 10.1007/978-981-10-0451-3_2

Well a fast Fourier transform (FFT) is normally used to determine power spectrum, but ironically FFT can handle only the time series, which are equally spaced, and does not any datum shift. Because of this limitation FFT cannot be used in astronomical time series as astronomical observations are implicitly unequally spaced and have datum shifts. Additionally, disturbances can originate from instrument failure or replacement or repair may cause datum shift in the time series. And so, instead of using FFT we use least space spectral analysis (LSSA) which was invented by Vanicek [1, 2]. And it is used as an alternative to FFT.

LSSA removes all the limitations of FFT, i.e., equally spaced, datum shift, etc. and so LSSA is more flexible than FFT. One of the important factors in LSSA is the covariance matrix, which should be fully populated, which means values may or may not be fully independent. The covariance matrix should be absolute, meaning a priori variance should be known.

2 Least Square Spectral Analysis

A time series comprises of a signal and a noise. A noise can be colored (periodic noise) [3] or nonstationary. LSSA can handle this kind of time series, which is impossible for FFT. FFT always needs a stationary time series with equal space and no datum shifts.

A noise can always distort a signal and cause datum shifts.

Advantages of using LSSA are as follows: LSSA can suppress the noise without showing any significant shift in the spectral peaks.

1. Unequally spaced time series can be used and do not need to be preprocessed.
2. Time series, which has covariance matrix, can be used.
3. Significant peaks of a time series can be found out.

3 Principles of LSSA

The principle that guides LSSA is the projection theorem in n-dimensional space. By considering the plane to be in a sub-space and then by applying projection theorem we have the shortest distance of a plane and a point in the perpendicular distance between the two, i.e., angle should be 90°.

3.1 Projection Theorem

Considering the space to be in Hilbert space (H) [4]. And point f to be a point in the time series $f(t)$, where $f \in H$ and $S \in H$, where S is the sub-space. Minimum or

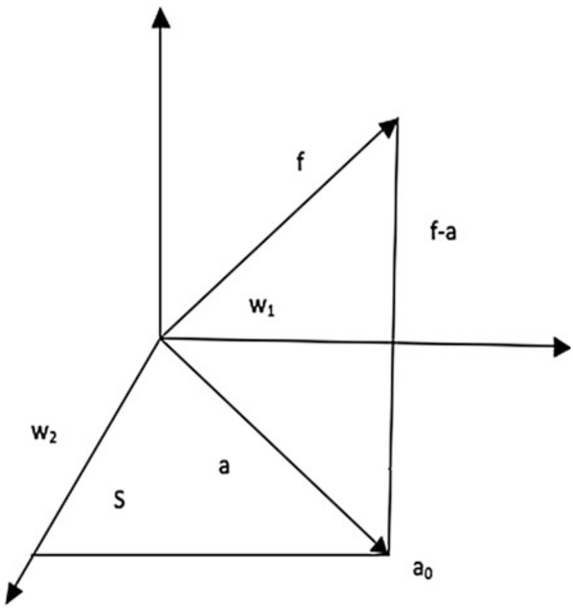

Fig. 1 Projection theorem

shortest distance(s) is achieved by using orthogonal projection as shown in Fig. 1 and a_o is the projection point on S such that $d(f, a_o) \leq d(f, a)$ and $(f - a_o) \perp S$ [4]. By assuming A consists of W_i where W_i is the basic function, then

$$a_o = \sum\nolimits_i C_i W_i$$

3.2 Second Projection Theorem

In second projection as shown in Fig. 2 corresponding with Fig. 1 we can project a back onto f, such that length is given by

$$\frac{\langle f,\ a \rangle}{\|f\|}$$

3.3 Mathematical Calculation

Let $f(t)$ be a time series and belong to Hilbert space H [4]. The observed time series are equally spaced and its value is observed at t_i where $i = 1, 2, 3, 4,\ldots, m$. Let C_f be fully populated covariance matrix and C_f matrices H.

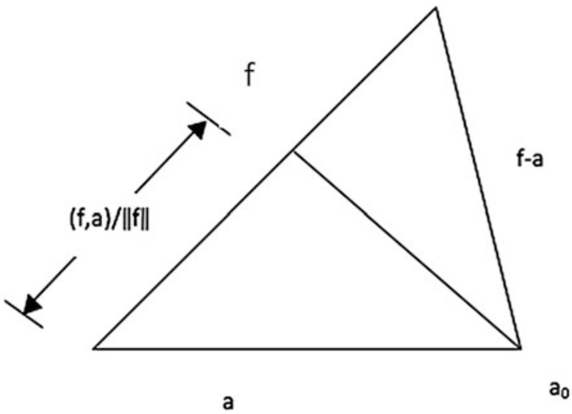

Fig. 2 Second projection theorem

The time series *f*(*t*) contains both signal and noise and so the time series *f* can be represented as

$$g = Wx$$

W is Vander mode matrix [4] and *x* is the vector of unknown parameters where $W = W_s/W_n$, where W_s and W_n are functional forms and *W* and *x* are arbitrary.

And so we need to find the difference between g and f by using least square method. This difference is the residual.

$$\hat{r} = f - \hat{g}$$

where *g* is the model time series. Figure 2 represents the relationship between *f*, *a*, *S* and $\{W_i, \forall i\}$. any point $s \in S$ [4] can be represented as

$$s = \sum_i C_i W_i \tag{1}$$

And so far n tuple $\{C_i, \forall_i\}$

$$a = \sum_i \hat{C}_i W_i \tag{2}$$

Now we have the condition [4],

$$(f - a) \perp s \equiv (f - a) \perp W_j \forall_j \tag{3}$$

In terms of $\{\hat{C}_i, \forall I\}$ corresponding to *a*,

$$\forall_j : \langle (x - \sum_i \hat{C}_i W_i),\ W_j \rangle = 0$$

This can be written as

$$\sum\nolimits_i \hat{C}_i\langle W_i,\, W_j\rangle = \langle f,\, W_j\rangle \qquad j = 1, 2, 3\ldots n \tag{4}$$

We can define

$$N = \begin{bmatrix} \langle W_1,\, \mathrm{W}_1\rangle\langle W_2,\, W_1\rangle\ldots\langle W_n,\, W_1\rangle \\ \langle W_1,\, W_2\rangle\langle W_2,\, W_2\rangle\ldots\langle W_n,\, W_1\rangle \\ \cdots \\ \cdots \\ \langle W_1,\, W_n\rangle\langle W_2,\, W_n\rangle\ldots\langle W_n,\, W_n\rangle \end{bmatrix}$$

$$U = \begin{bmatrix} \langle f,\, W_1\rangle \\ \langle f,\, W_2\rangle \\ \vdots \\ \langle f,\, W_n\rangle \end{bmatrix}$$

$$\hat{C} = \begin{bmatrix} \hat{C}_1 \\ \hat{C}_2 \\ \vdots \\ \vdots \\ \vdots \\ \hat{C}_n \end{bmatrix}$$

Then Eq. 4 becomes

$$N\hat{C} = \mathrm{u} \tag{5}$$

Since $f - a \perp S$, then normal equation

$$\begin{aligned} N &= W^T W \text{ and} \\ U &= W^T f \end{aligned}$$

Equation 5 becomes

$$\begin{aligned} \hat{C} &= N^{-1}u \\ &= \left(W^T W\right)^{-1} W^T f \end{aligned} \tag{6}$$

And from Eq. 2

$$\mathrm{A} = W\hat{C} = \sum_i W_i\hat{C}_i$$

Therefore, residual

$$\begin{aligned} r &= f - a \\ &= f - W\hat{C} \\ &= f - W\left(W^T W\right)^{-1} W^T f \end{aligned} \tag{7}$$

So by using the projection theorem,

$$\begin{aligned} & r = f - a \\ & \text{so } \hat{r} \perp a \text{ and } r = f - a \perp S \end{aligned}$$

so we can say that $r \perp g$ [4], i.e., they are orthogonal (from Fig. 1). This means the signal x is divided into noise, i.e., residual r and signal g.

Consider Fig. 2, we project a_o to x to know that how much of a_o is present in x [3].

$$\begin{aligned} \mathrm{S} &= \langle f,\, \hat{g}\rangle / |f| / |f| \\ &= \langle f,\, \hat{g}\rangle / |f|^2 \\ &= f^T C_f^{-1} \hat{g} / f^T\, C_f^{-1} f \quad \epsilon\, (0, 1) \end{aligned} \tag{8}$$

Signal can be expressed in the form of sine and cosine in spectral analysis, i.e., W can be expressed as

$$W = [\cos \theta_i t,\ \sin\theta_i t], I = 1, 2, 3\ldots k \tag{9}$$

For each angle there is different spectral value and so LSSA can be represented as

$$S(\theta_i) = f^T C_f^{-1} \hat{g}(\theta_i) / f^T C_f^{-1} f, \quad I = 1, 2, 3, \ldots k \tag{10}$$

This above equation shows it depends on the trigonometric function θ.

Reexamining Eq. (1) and matrix W. W is partitioned into W_S and W_N where W_S is for signal and W_N is for noise. From Eq. (4) we know that this θ is used to describe the components of the series, which is periodic. In least square solution which is used in least square spectral method depends on the parameters. These parameters are driven by noise and can change their phase and amplitude due to the noise. And so, we need to find out the pdf (partial distribution function) of the least square spectrum.

4 Revisiting Least Square Spectrums

$$\hat{u} = (I - K) f \tag{11}$$

where I is the identity matrix and

$$K = W\left(W^T C_f^{-1} W\right)^{-1} W^T C_f^{-1}$$

Substituting Eq. (11) in (10)

$$S = f^T C_f^{-1} K_f / f^T C_f^{-1} f \tag{12}$$

To derive the pdf equation, the two equation should be statistically independent but in the Eq. (12) the two equations are not statistically independent as $C_f{}^{-1} J C_f C_f{}^{-1} \neq 0$ (Lemma 3) [3].

By rearranging the denominator of the equation we have

$$f^T C_f^{-1} f = f^T C_f^{-1} Jf + f^T C_f^{-1} (I - K) f \tag{13}$$

The denominator is decomposed into two terms. The first term is the signal and the second term is the noise. Now Eq. 13 is statistically independent.

Using Eq. (13) in (12) and then rearranging it we get

$$\begin{aligned} S &= \left[1 + f^T C_f^{-1} (I - K) f / f^T C_f^{-1} Jf\right]^{-1} \\ &= [1 = q_n / q_s] \end{aligned} \tag{14}$$

Equation (14) is a function of two quadratic equations q_s and q_n which are statistically independent and the ratio gives the least square spectrum. It is inverse of signal to noise ratio (SNR). The above Eq. (14) can be used to calculate the PSD (power spectral density) by applying logarithm function. The PSD is given by the equation

$$\mathrm{PSD} = 10 \log[S/1 - S] \tag{15}$$

This PSD that is given by Eq. (15) is equivalent to the PSD of the FFT method only when the time series is equally weighted and equally spaced. So FFT, PSD is not useful for time series with datum shift.

But the above PSD given by Eq. (15) can be used to calculate any time series even if it is not equally spaced.

5 Conclusion

So by using least square method we can eliminate the limitation of FFT. So for analysis time series with datum biased and unequally spaced value least square spectral method is used. The spectral density is same as the spectral density of the FFT but the only difference is that it can be used for time series with datum shifts and unequally spaced.

References

1. Vanicek, P.: Approximate spectral analysis by least-square fit. Astrophys. Space Sci. **4**, 387–391 (1969)
2. Vanicek, P.: Further development and properties of the spectral analysis by least-squares. Astrophys. Space Sci. **12**, 10–33 (1971)
3. Pagiatakis, S.D.: Application of the least-squares spectral analysis to superconducting gravimeter data treatment and analysis. Cahiers du Centre Europeen de Geodynamique et Seismologie **17**, 103–113 (2000)
4. Wells, D., Vanicek, P., Pagiatakis, S.: least square spectral analysis revisited (1985)

http://www.springer.com/978-981-10-0450-6

Proceedings of Fifth International Conference on Soft Computing for Problem Solving
SocProS 2015, Volume 2
Pant, M.; Deep, K.; Bansal, J.C.; Nagar, A.; Das, K.N. (Eds.)
2016, XIX, 1045 p. 413 illus., Softcover
ISBN: 978-981-10-0450-6