Determination of the Confidence Level of PSD Estimation with Given D.O.F. Based on WELCH Algorithm

Xue-wang Zhu[1, *], Si-jian Zhang[1] and Qing-lin Liu [1]

[1]Institute of Systems Engineering, CAEP, Mianyang Sichuan, China

Keywords: Confidence level, PSD estimation, WELCH algorithm, Degree of freedom, Random error

Abstract. Approach is proposed for determining the confidence level of the power spectral density (PSD) estimation with give Degree-of-freedom (D.O.F.) based on WELCH algorithm, a modified periodogram estimator for estimating PSD. A relationship between the random error, defined as a ratio of variance and mean value, and the D.O.F. is achieved firstly due to the analysis results of the statistic characteristic quantities. Then a quantitative relation between the confidence level, described as confidence probability and confidence interval, and the D.O.F. is established. Some examples are provided finally to demonstrate the proposed method. It is showed from the results that the PSD WELCH estimation with 128 D.O.F. have probability 68.4% only to ensure the true PSD within 0.88~1.2 times the estimation value.

Introduction

In engineering field it is customary to estimate power spectral density (PSD) of random vibration signal by WELCH algorithm[1,2] , one of modified periodogram PSD estimation methods which reduces effectively variance segment smoothing and data overlapping etc. It is proved that PSD estimation by WELCH algorithm is consistent result and its mean tends toward the exact solution and its variance to null if segment number and data number in every subsection are large enough.

In practice the variance of estimation is never equal to zero because of finite limited numbers of segment and data length per segment. A valuable concept, random error[3], which is defined as a ratio of the variance and the mean value, has been proposed to describe the relative magnitude of the variance of estimation. Higher the random error, higher is the probability of the estimation far from the mean value. In other word lower the random error, higher is the probability of the estimation near the mean value.

In order to obtain the random error for a PSD WELCH estimation, a special random variable in mathematics, c^2 variable[3], is reviewed. Random error can be expressed by a simple function of the D.O.F. listed in equation 1 for A c^2 variable with N D.O.F. which is defined as a square sum of N normal Gaussian variables with zero mean values and unity variance.

$$e = s/m = \sqrt{2/N} \tag{1}$$

Where s， m are the variance and the mean respectively.

For a given random error e requirement the number of D.O.F. is educed easily as equ.2 for c^2 variable.

$$N = 2e^{-2} \tag{2}$$

It is pity that PSD WELCH estimator, as a random variable, cannot be cognized as a c^2 variable blindly because of non-unity variance for every term used to be summed. A PSD WELCH estimation variable with M segments data can be expressed as a square sum of 2M random variables which are real part and imaginary part of Fourier spectrum with non-unity variances for each segment data weighted and overlapped. After standardizing those 2M random variables the PSD estimation variable can be rewritten as a linear transformation of a c^2 variable with 2M D.O.F. and the random error can be gained by computing variance and mean value. Coincidentally, the expressions of random errors have the same formula for c^2 variable and PSD WELCH estimation variable thus equation (1) and (2) can

© 2015. The authors - Published by Atlantis Press

be also used to determine the number of D.O.F. of PSD WELCH estimation with a given random error requirement.

Confidence level[4] is another description on requiring precision to PSD WELCH estimation which is defined by confidence probability and confidence interval for the certain D.O.F.. The relationship between confidence level and D.O.F. need to be established. Two kinds of description for the relationships are discussed in the paper. (a) To find the confidence interval to a given confidence probability for a PSD WELCH estimation with the certain D.O.F. ; and (b) to find the confidence probability to a given confidence interval for a PSD WELCH estimation with the certain D.O.F. .

The rest parts of the paper is organized as follows. In Sect.2, WELCH algorithm for PSD estimation is reviewed firstly for providing a summed formulation to evaluate PSD. In Sect. 3 the random error expression of the PSD WELCH estimation with given D.O.F. is proved. Confidence level analysis is discussed in Sect.4 for the PSD WELCH estimation with given D.O.F. and the relationship is developed between D.O.F. and confidence level. The procedures in detail for confidence level analysis and examples are provided in Sect.5. The conclusion is derived in Sect.6.

Review of WELCH algorithm for estimating PSD

WELCH algorithm[1][2] for estimating the PSD of a random vibration is a periodgram spectral estimator in which the data with given length is divided into segments with length including overlapped data and the PSD is averaged to the periodgrams of every segment window weighted. For example, discrete random vibration signals with length N are noted as $\{x(0),x(1),\cdots\cdots,x(N-1)\}$.Then K segments with length L can be formed where L-D data is overlapped, $N=L+D(K-1)$.For segment i the L data are $x_i(n)=x(n+iD)$,

$n=0,1,\cdots\cdots,L-1,i=0,1,\cdots\cdots,K-1$.After weighting with window function $w(n)$ the periodgram spectral can be estimated as:

$$\hat{S}_M^{(i)}(f)=\frac{1}{LU}\sum_{n=0}^{L-1}\left|w(n)x(n+iD)e^{-j2p\,fn}\right|^2 \tag{3}$$

Where $U=\frac{1}{L}\sum_{n=0}^{L-1}\left|w(n)\right|^2$ is power parameter factor of window function $w(n)$ in order to ensure the estimation no bias intimately.

PSD WELCH estimation is the average of K period gram spectral above, that is:

$$\hat{S}_w(f)=\frac{1}{K}\sum_{i=0}^{K-1}\hat{S}_M^{(i)}(f) \tag{4}$$

It is showed from equation (1) that periodgram spectral for every segment windowed is proportional to the square of the Fourier transform of the signals. Hence equation (4) can be rewritten as:

$$\hat{S}_w(f_i)=g\sum_{m=0}^{K-1}\left\{\left[X_a^m(f_i)\right]^2+\left[X_b^m(f_i)\right]^2\right\} \tag{5}$$

Where $g=\frac{1}{KLU}$ is constant and $X_a^m(f_i)$、$X_b^m(f_i)$ are real part and imaginary part respectively of Fourier transform to the m_th segment signal windowed.

Then the PSD WELCH estimation is described as a square sum of 2 K variables with respect to Fourier transform.

Random error of a PSD WELCH estimation

It is known that the random vibration signal in engineering can be simplified as Gaussian variable. And the two induced complex variables by Fourier transformation are also Gaussian probability because

Fourier transformation is linear operator. Hence, PSD WELCH estimation described by equation (5) is proportional to square sum of 2K Gaussian variables .The induced Gaussian variables have zero mean values and non-unity variances if the raw vibration signal is Gaussian with zero mean value. So the random error of the estimation can't be obtained by equation (1) directly. Further it is proven that the two induced Gaussian variables, $X_a^m(f_i)$ 、 $X_b^m(f_i)$, have the same variance, notated as l . Then a standardizing process can be applied to equation (5) and obtain equation (6):

$$\hat{S}_w(f_i)=q\hat{Q}_x(f_i)=q\sum_{m=0}^{K-1}\left\{\left[\frac{X_a^m(f_i)}{l}\right]^2+\left[\frac{X_b^m(f_i)}{l}\right]^2\right\} \tag{6}$$

Where $q=gl^2$ is constant, and:

$$\hat{Q}_x(f_i)=\sum_{m=0}^{K-1}\left\{\left[\frac{X_a^m(f_i)}{l}\right]^2+\left[\frac{X_b^m(f_i)}{l}\right]^2\right\} \tag{7}$$

Variable $\hat{Q}_x(f_i)$ is c^2 variable with 2K D.O.F. if all the variables to be summed, $X_a^m(f_i)$, $X_b^m(f_i)$ are independent. For this reason the PSD WELCH estimation using K segments data is defined as the PSD estimation with 2K D.O.F.. On the other word, the number of D.O.F. for a PSD WELCH estimation is twice of the number of segments used to estimate the PSD. Following the main statistic parameters mean value and variance will be calculated for a PSD WELCH estimation with 2K D.O.F..

$$E\left[\hat{S}_w(f_i)\right]=qE\left[\frac{\hat{S}_w(f_i)}{q}\right]=qE\left[\hat{Q}_x(f_i)\right]=2Kq \tag{8}$$

$$\begin{aligned} s_{\hat{S}}^2 &= E\left[\left(\hat{S}_w(f_i)\right)^2\right]-\left(E\left[\left(\hat{S}_w(f_i)\right)\right]\right)^2 \\ &= q^2\left(s_{\hat{Q}}^2+\left(E\left[\left(\hat{Q}_x(f_i)\right)\right]\right)^2\right)-4K^2q^2=4Kq^2 \end{aligned} \tag{9}$$

Then the random error can obtained easily and the result is coincident with equation (1). And equation (2) can be used to determine the number of D.O.F. of a PSD WELCH estimation with given random error requirement in reason.

Confidence level of a PSD WELCH estimation

Confidence level is another representation for describing the precision of a PSD WELCH estimation which is defined as confidence probability under the certain confidence interval and confidence interval under the certain confidence probability. In Sect. 3 we have stated that PSD WELCH estimation is a linear transformation of c^2 variable whose confidence level can be obtained easily from textbooks or EXCEL function. Following the description of confidence level for a PSD WELCH estimation will be discussed based on the confidence level analysis method of c^2 variable.

According to equation (6), following relation exists:

$$P\left\{a\le\hat{S}_w(f_i)\le b\right\}=P\left\{\frac{a}{q}\le\hat{Q}_x(f_i)\le\frac{b}{q}\right\} \tag{10}$$

Where $P\{\ \}$ is the confidence probability under the confidence interval$\{\ \}$.

Notes that $\frac{a}{q}\le\hat{Q}_x(f_i)\le\frac{b}{q}$ when $a\le\hat{S}_w(f_i)\le b$, this means that:

$$n_1 = \frac{a/q}{E\left[\hat{Q}_x(f_i)\right]} \le \frac{\hat{S}_w(f_i)}{E\left[\hat{S}_w(f_i)\right]} \le \frac{b/q}{E\left[\hat{Q}_x(f_i)\right]} = n_2 \qquad (11)$$

Equation (11) gives the confidence interval with given probability for PSD WELCH estimation with given D.O.F. by using the interval of c^2 variable with the same probability and the same D.O.F.. Rewrite Equation (11) to obtain:

$$n_1 \times E\left[\hat{S}_w(f_i)\right] \le \hat{S}_w(f_i) \le n_2 \times E\left[\hat{S}_w(f_i)\right] \qquad (11a)$$

$$1/n_2 \times \hat{S}_w(f_i) \le E\left[\hat{S}_w(f_i)\right] \le 1/n_1 \times \hat{S}_w(f_i) \qquad (11b)$$

Equation (11a) is the confidence interval of the estimation and (11b) the confidence interval of the mean value of the estimation.

Procedure of confidence level analysis and exercises

The procedure for determining confidence interval (n_1, n_2) to certain probability a of the PSD WELCH estimation with given D.O.F. (N) is like this:

According to the probability a, to find the corresponding one-side probabilities b_1, b_2 for c^2 variable with the same D.O.F., that is:

$$P\left\{\hat{Q}_x(f_i) \ge \frac{a}{q}\right\} = b_1 \quad P\left\{\hat{Q}_x(f_i) \ge \frac{b}{q}\right\} = b_2 \text{,while } P\left\{\frac{a}{q} \le \hat{Q}_x(f_i) \le \frac{b}{q}\right\} = a \qquad (12)$$

Let symmetry exist, $P\left\{\hat{Q}_x(f_i) \ge \frac{b}{q}\right\} = P\left\{\hat{Q}_x(f_i) \le \frac{a}{q}\right\} = 1 - P\left\{\hat{Q}_x(f_i) \ge \frac{a}{q}\right\}$, then:

$$b_1 = \frac{1+a}{2}, \; b_2 = \frac{1-a}{2} \qquad (13)$$

Determine the confidence interval $[a/q, b/q]$ of the c^2 variable with the same probability a and the same D.O.F. N. Either the textbooks on probability or EXCEL tool be can used for this aim. Simply excel function gives the results:

$$\frac{a}{q} = \text{CHIINV}(b_1,\text{N}) \text{ and } \frac{b}{q} = \text{CHIINV}(b_2,\text{N}) \qquad (14)$$

Applying equation (11) in Sect.4 the confidence interval can be obtained:

$$n_1 = \text{CHIINV}(b_1,\text{N})/\text{N}, \; n_2 = \text{CHIINV}(b_2,\text{N})/\text{N} \qquad (15)$$

Applying equation (11a) or (11b) the interval of estimation value or mean value of the PSD WELCH estimation with N D.O.F. is obtained.

The procedure for determining confidence probability a to certain interval $[n_1, n_2]$ of the PSD WELCH estimation with given D.O.F. (N) is listed following:

a) Construct the confidence intervals of the c^2 variable with the same probability and same D.O.F..

$$\frac{a}{q} = n_1 \times N, \; \frac{b}{q} = n_2 \times N \qquad (16)$$

b) Determine the confidence probability b_1 and b_2 of the c^2 variable with D.O.F. N to confidence intervals $\left[a/q, \infty\right]$ and $\left[b/q, \infty\right]$.Excel functions are:

$$b_1 = \text{CHIDIST}(a/q,\text{N}), \; b_2 = \text{CHIDIST}(b/q,\text{N}) \qquad (17)$$

c) Determine the confidence probability a of the c^2 variable with N D.O.F to the interval $\left[a/q, b/q\right]$ which is the same as the confidence probability of the PSD WELCH estimation.

$$a = b_1 - b_2 \tag{18}$$

Example 1: Determine the confidence intervals to various confidence probabilities for the PSD WELCH estimation with 128 D.O.F.. Such a PSD Welch estimation is satisfy with the precision requirement of the standard MIL-STD-810F[5] MIL-STD-810G[6]. Table 1 lists the results for several typical probabilities. Notes that the random error of 128 D.O.F. WELCH PSD is 12.5% according to equation (1), the confidence probability is about 68% with respect to the estimation which fluctuates within 1 times random error around the mean value. Similarly the probabilities are about 95% and 99.9% to 2 times and 3 times random error fluctuation for the estimation. Further Table 2 and Table 3 listed the confidence levels of the PSD WELCH estimations of 256 D.O.F. and 512 D.O.F.. The probabilities reach about 85% and 95% to the estimations within 1 times random error fluctuation.

Table 1 the confidence level of the PSD WELCH estimation with 128 D.O.F.

a	b_1	b_2	n_1	n_2
60.0%	0.800	0.200	0.894	1.103
68.0%	0.840	0.160	0.876	1.124
80.0%	0.900	0.100	0.844	1.163
95.0%	0.975	0.025	0.770	1.259
99.0%	0.995	0.005	0.707	1.351
99.9%	0.9995	0.0005	0.639	1.463

Table2 the confidence level of the PSD WELCH estimation with 256 D.O.F.

a	b_1	b_2	n_1	n_2
60.0%	0.800	0.200	0.925	1.074
68.0%	0.840	0.160	0.912	1.088
80.0%	0.900	0.100	0.888	1.115
85.0%	0.925	0.075	0.876	1.130
95.0%	0.975	0.025	0.834	1.181
99.0%	0.995	0.005	0.787	1.242
99.9%	0.9995	0.0005	0.734	1.317

Table 3 the confidence level of the PSD WELCH estimation with 512 D.O.F.

a	b_1	b_2	n_1	n_2
60.0%	0.800	0.200	0.947	1.052
68.0%	0.840	0.160	0.938	1.062
80.0%	0.900	0.100	0.921	1.081
95.0%	0.975	0.025	0.881	1.126
99.0%	0.995	0.005	0.846	1.168
99.9%	0.9995	0.0005	0.807	1.218

Conclusion and discussion

A approach and procedure is proposed for determination of the confidence level to the PSE WELCH estimation based on a special random variable, c^2 variable. The relationship is established between the PSD estimation and c^2 variable. The procedures are suggested for two descriptions of confidence level and the examples are provided. It is showed that under 95% confidence probability the confidence interval of 128 D.O.F PSD WELCH estimation is [0.77,1.26].On other words such an estimation fluctuates around the mean value between [0.77,1.26] times. On same conditions the range is [0.83,1.18] time for the PSD estimation with 256 D.O.F.

The limitations need to be illuminate for the independence and symmetry. In this paper PSD WELCH estimation is described as a combination of c^2 variable on the condition of independence for all date segments summed. In fact dependence exists certainly because of overlapping. In the calculation of one-side probabilities b_1 and b_2 of the c^2 variable symmetry has been applied which can also effect the quantity of the results.

Reference

[1] P.D. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Trans.Audio Electro acoustics 15 (2) (1967) 70–73.

[2] Zhiqiang Sun, Hongchu Chen, Yanping Chen, Application of periodogram and Welch based spectral estimation to vortex frequency extraction, 2011 Second International Conference on Intelligent System Design and Engineering Application

[3] Newland DE , An introduction to random vibrations and spectral analysis. Wiley, New York. 1984

[4] Politis, Romano, J.P., Lai, T.L.,.Bootstrap confidence bands for spectra and cross-spectra. IEEE Transactions on Signal Processing 40(1992):1206–1215

[5] MIL-STD-810F Department of Defense Test Method Standard For Environmental Engineering Considerations And Laboratory Tests, Method 514.5 vibration

[6] MIL-STD-810G Department of Defense Test Method Standard For Environmental Engineering Considerations And Laboratory Tests, Method 514.6 vibration

Xue-wang Zhu, tel: 13981112652, E-mail: zhuxw@caep.cn