Research Journal of Applied Sciences, Engineering and Technology 6(5): 802-806, 2013
DOI:10.19026/rjaset.6.4123
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: September 03, 2012 Accepted: October 05, 2012 Published: June 25, 2013

Research Article

Mode Parameters Estimation of Vibration Signal Based on Aberrant Point Clustering and Elimination

Ye Qingwei, Wang Dandan and Zhou Yu
Information Science and Engineering Institute, Ningbo University, Zhejiang 315211, China

Abstract: A new mode parameter estimation method of vibration signal is put forward in this study. At first, the frequency response curve of vibration signal is fitted by Levy polynomial and the each distance between the fitted curve point and the frequency response curve point is calculated. Then the distance set is clustered by k-means algorithm into two classes. One class is clustered with smaller distance points and another class is clustered with larger distance points which are named aberrant point set. The class of larger distance points clustered will be eliminated and the new frequency response curve is obtained. At last, the new frequency response curve is fitted by Levy polynomial again and the new aberrant point set is eliminated again and so on. Finally, the fitting accuracy will be arrived according to the above algorithm. Plenty of simulation tests to vibration signals show that this algorithm can accurately extract mode parameters of the vibration frequency spectrum. It also confirms that in the different noise intensity and different distance between adjacent frequency cases, the precision of the algorithm proposed by this study is obviously superior to the existing Levy algorithm.

Keywords: Frequency response outliers, K-means clustering algorithm, levy algorithm, polynomial fitting

INTRODUCTION

Mode parameter identification is an important element in Structural Health Monitoring (SHM). Accurate mode parameters are the prerequisite for the finite element model updating, structural damage detection and evaluation of the structural performance. Modes of vibration are global properties of a structure. That is, a mode is defined by its natural frequency, damping and mode shape, which can each be measured (or estimated) from a set of FRF measurements taken from the structure. Therefore, the process of identifying parameters from the dynamic response is called curve fitting, or parameter estimation, i.e., time-frequency analysis FFT, Time-series Decomposition and so on. The Wavelets Transform (WT) (Ge *et al*., 2006; Xu and Song, 2011), one of time-frequency analysis, has good performance on the outlier detection but the choice of wavelet function and its parameters will affect the identification precision. Frequency domain method is affected by not only the FFT error but also the noise, especially in high damping ratio case. The researchers are trying new ways to improve the frequency domain method, for example, the Levy algorithm and orthogonal polynomial algorithm (Richardson, 1986) improved with the high accuracy and simple idea. However, the precision of the mode parameters is low when the noise is strong or the modes are dense, such as a strong noise near the peak spectrum will lead to a large deviation.

This study comes up with a new Levy algorithm improved by importing the k-means clustering algorithm into spectrum analysis. To improve the fitting precision, it use k-means clustering algorithm to eliminate those points far away from the fitting curve. Plenty of simulation tests of vibration signal turn out that the algorithm proposed by this study can accurately extract mode parameters of the vibration frequency spectrum and the precision is obviously superior to the existing Levy algorithm.

Levy algorithm: The vibration responses of viscous damping system (Fu and Hua, 2000) is given by:

$$y(t) = \sum_{k=1}^{n} a_k e^{-\lambda_k t} \cos(\omega_{dk} + \varphi_k) \tag{1}$$

The rational fraction of (1)' FFT is:

$$\begin{aligned} Y(\omega) &= \frac{\alpha_0 + \alpha_1 (j\omega) + \ldots + \alpha_{2n-2} (j\omega)^{2n-2}}{1 + \beta_1 (j\omega) + \ldots + \beta_{2n} (j\omega)^{2n}} \\ &= \frac{\sum_{k=0}^{2n-2} \alpha_k (j\omega)^k}{1 + \sum_{i=1}^{2n} \beta_i (j\omega)^i} \end{aligned} \tag{2}$$

Corresponding Author: Ye Qingwei, Information Science and Engineering Institute, Ningbo University, Ningbo, Zhejiang 315211, China

The FRF can be represented as a ratio of two polynomials, as shown in Eq. (2). So transform the rational fraction into the general polynomial form, then get Eq. (3):

$$\begin{aligned} Y(\omega) &= \sum_{k=0}^{2n-2} (j\omega)^k \alpha_k - Y(\omega)\sum_{i=0}^{2n} (j\omega)^i \beta_i \\ &= \sum_{k=0}^{2n} v_k (j\omega)^k \end{aligned} \tag{3}$$

When curve fitting this analytical form to the measurement data, the unknown coefficients of both the numerator and denominator (α_k, k = 0,…, m) and (β_i, i = 0,…, n) are determined. Then, three mode parameters (frequency, damping and complex residue) can be solved out according to the coefficients.

Now, the curve fitting can be done in the least squared error sense by solving a set of linear Eq. (3), for the coefficients. To begin the problem formulation, we need to define an error criterion. First we can write the error at a particular value of frequency (ω_k) as simply the difference between the analytical value (Y_k) and the measurement value of the FRF ($\tilde{Y}_k$), as shown in following expression:

$$\begin{aligned} e_k &= \sum_{i=0}^{2n-2} (j\omega_k)^i \alpha_i - \tilde{Y}_k \left[\sum_{i=0}^{2n} (j\omega_k)^i \beta_i + 1 \right] \\ &= p^T(j\omega_k)\vec{\alpha} - \tilde{Y}_k q^T(j\omega_k)\vec{\beta} - \tilde{Y}_k \end{aligned}$$

where,

$$\vec{\alpha} = [\alpha_0, \alpha_1, \cdots, \alpha_{2n-2}]^T \quad \vec{\beta} = [\beta_0, \beta_1, \cdots, \beta_{2n-1}]^T$$
$$p(j\omega_k) = [1 \ j\omega_k \ (j\omega_k)^2 \ \cdots \ (j\omega_k)^{2n-2}]$$
$$q(j\omega_k) = [\tilde{Y}_k \ \tilde{Y}_k(j\omega_k) \ \tilde{Y}_k(j\omega_k)^2 \ \cdots \ \tilde{Y}_k(j\omega_k)^{2n-1}]$$
$$k = 1, 2, \ldots, s$$

Furthermore, we can make up an entire vector of errors, one for each frequency value where we wish to curve fit the data, as shown in expression $\vec{e}$ = {e_1, e_2,…,e_n}′ A squared error criterion can be form from the error vector, as shown in E = $\vec{e}^H\vec{e}$ Notice that, this criterion (E) will always have a non-negative value. Therefore, we want to find values of the α_k and β_k so that the value of E is minimized, ideally zero. Using the error vector expression:

$$\vec{e} = P\vec{\alpha} - Q\vec{\beta} - \vec{w}$$

Note that it is now written as a function the two unknown coefficient vectors $\vec{\alpha}$ and $\vec{\beta}$. This criterion function has a single minimum value, so we can set its derivatives (or slope) with respect to the variables $\vec{\alpha}$ and $\vec{\beta}$ to zero to find the minimum point. The linear equations are:

$$\frac{\partial E}{\partial \alpha} = 0, \frac{\partial E}{\partial \beta} = 0$$

In principle then, the least squares estimates of $\vec{\alpha}$ and $\vec{\beta}$ can be obtained by solving the above linear equations.

From the foregoing, all the frequency points are used to fit the FRF curve by the Levy parameter identification method. When there is no noise or the noise is very low, it's fitting precision is good. If there are strong noises or larger error caused by misoperation, the fitting curve will deviate so seriously that the identification parameters are less accurate.

In previous studies (Verboven *et al*., 2005; Liu *et al*., 2005), many new methods based on the local curve fitting had been put forward to improve these disadvantages. To achieve better interference rejection, these local curve fitting methods select threshold artificially according to experience or select a part of mode peaks to fit the frequency response curve. However, it fails to select the effective frequency response point automatically, so it's difficult to practice. So for these frequency response points, this study uses clustering thought to find out and eliminate the outliers automatically. This will improve the Levy fitting accuracy and availability and estimate the natural frequency, damping and mode shape accurately.

IMPROVED LEVY ALGORITHM BASED ON OUTLIERS CLUSTERING

According to the clustering of k-means algorithm, this study constructs an iterative algorithm. It identifies and eliminates the frequency response outliers constantly and then uses the Levy algorithm to fit the rest of the effective frequency response points, finally, obtains the accurate mode parameters. The algorithm is described as follows:

According to the Eq. (3), Frequency response functions polynomial expression in frequency ω_k is:

$$\begin{aligned} &Y(\omega_k) = v_0 + v_1(j\omega_k) + \ldots + v_{2m}(j\omega_k)^{2m} \\ &k = 1, 2, \ldots, s \quad (s >> 2m) \end{aligned} \tag{4}$$

Given the measurement value of the FRF ($\tilde{Y}_k$), its discrete data point sets is:

$$\{\omega_k, \tilde{Y}(\omega_k)\} \quad k = 1, 2, \ldots, n$$

- The frequency response point sets is into the Eq. (4) to obtain the polynomial coefficients {$v_{2m}^{(1)}$, $v_{2m-1}^{(1)}$, …$v_0^{(1)}$} where the superscript (1) is the first iteration coefficients, so the polynomial expression after the first iteration is:

$$Y(\omega)^{(1)} = v_{2m}^{(1)}(j\omega)^{2m} + v_{2m-1}^{(1)}(j\omega)^{2m-1} + \ldots + v_0^{(1)} \tag{5}$$

- The analytical value of the frequency ω_k is Y $(\omega_k)^{(1)}$, the error between the analytical value and the measurement value is calculated by the below equation:

$$e_k^{(1)} = | Y(\omega_k)^{(1)} - \tilde{Y}(\omega_k) | \quad k = 1,2,...,s$$

- The error set E = $\{e_k^{(1)}\}$ is considered to be the clustering set. It divides the set E into two classes using the k-means algorithm, one is the large error $Z_1^{(1)}$ and the other $Z_2^{(1)}$ is smaller, that is:

$$E = \{e_k^{(1)}\} = Z_1^{(1)} \cup Z_2^{(1)}$$

- The mapped frequency response point of $Z_1^{(1)}$ is considered to be the outliers and to be rejected. The mapped frequency response point of $Z_2^{(1)}$ is into the Eq. (4) to obtain the polynomial coefficients$\{v_{2m}^{(2)}, v_{2m-1}^{(2)}, Lv_0^{(2)}\}$ and the polynomial is:

$$Y(\omega_k)^{(2)} = v_{2m}^{(2)}(j\omega_k)^{2m} + ... + v_0^{(2)} \quad \{\omega_k \in Z_2^{(1)}\}$$

  Likewise, the fitting error can be calculated by the similar expression:

$$e_k^{(2)} = | Y(\omega_k)^{(2)} - \tilde{Y}(\omega_k) | \qquad \{\omega_k \in Z_2^{(1)}\}$$

- The K-means is used to classify the error vector $\{e_k^{(2)}\}$ and rejected the outliers set $Z_1^{(1)}$. The above process is repeated until the fitting parameters are stable.

The outliers clustering algorithm eventually cluster the error set E into two classes E = $\{E_k\} = Z_1 \cup Z_2$ where the large error Z_1 is mapped into the frequency response outliers and must be rejected, the small error $Z_2^{(1)}$ is mapped into the reasonable frequency response points and to be kept.

In engineering analysis, the parameter estimation method using Levy algorithm based on outliers clustering of frequency response can not only effectively detect and eliminate the outliers but also fit the reasonable frequency response points to identify the parameters. It reduces the influence of noise and dense modes and makes the system analysis more accurate and reliable. The next, this algorithm is applied to different modes of vibration simulation signal to verify the effectiveness of the algorithm. Experiments also show that this algorithm can get good fitting effect under strong noise, dense modes and high ratio of peak.

ILLUSTRATIVE EXAMPLES

In order to test the algorithm this study proposed, the following viscous damper system simulation signals are used:

$$R = segam * Randn(256)$$

$$S_1(t) = 0.1e^{-\delta_1 t}\cos(2\pi f_1 t + 0.1) + 0.2e^{-\delta_2 t}\cos(2\pi f_2 t + 0.2) + 0.15e^{-\delta_3 t}\cos(2\pi f_3 t + 0.2)$$

$$S_2(t) = 0.25e^{-\delta_1 t}\cos(2\pi f_1 t + 0.1) + 0.5e^{-\delta_2 t}\cos(2\pi f t + 0.2)$$

$$t = 0 - 5s$$

where, R is the white Gaussian noise and segam is its intensity. The noise is stacked on the signals S_1 S_2 to evaluate the resistance performance of noise. δ_i, i = 1, 2, 3 is damping coefficients. To turn out the effectiveness of the algorithm in this study, the fixed-two mode frequency signal S_1 is used. Likely, f in S_2 is a variable to assess the influence of different mode frequency space. Sampling number is N = 512.

In order to evaluate the extraction accuracy of mode parameters, this study uses the evaluation function defined in literature (Ye and Wang, 2009). Suppose that the Theoretical mode parameters are $\{r_k, \xi_k, \omega_{dk} | k = 1,..., M\}$. The Fitting mode parameters are $\{r_k^*, \xi_k^*, \omega_{dk}^* | k = 1,..., M\}$. So the average relative error is ε, that is:

$$\varepsilon = \frac{1}{2M}\sum_{k=1}^{M}\left(\left|\frac{\xi_k - \xi_k^*}{\xi_k}\right| + \left|\frac{\omega_{dk} - \omega_{dk}^*}{\omega_{dk}}\right|\right) \times 100\%$$

Several examples of using the outliers clustering fitting algorithm are given here.

- For the bridge signal's characteristic (rapid decreases, low natural frequency, dense mode), the frequencies in S_1 are given f_1 = 5 Hz, f_2 = 8 Hz, f_3 = 12 Hz so the number of sampling can be obtained by $n_k = \frac{f_k}{f_s}$ gN respectively, $n_1 = 25$, $n_2 = 40$, $n_3 = 60$. And damping coefficients are $\delta_1 = 1.5$, $\delta_2 = \delta_3 = 1.8$ the damping ratio is $\xi_1 = 0.0477$, $\xi_2 = 0.0358$, $\xi_3 = 0.0239$ by $\xi_k = \frac{\delta_k}{\omega_{nd}}$ where ω_{nd} is natural frequency. All of these mode parameters are ideal and noise free. When the simulation experiment is done, the noise must be added. In this experiment, two algorithms are compared by using MATLAB and the fitting figure is as Fig. 1.
  Figure 1 shows that in different noise intensity the fitting result of the improved algorithm and the Levy algorithm respectively. The Blue curve (H + R) stands for an adding noise signal spectrum curve (H is the simulation signal, R is noise signal), black curve for the algorithm this study proposed and the red dashed line for Levy fitting result. Table 1 is the fitting value of parameters.

Table 1: Two methods fitting parameters under different noise intensity

	SNR (dB)	Natural frequency (Hz)	Damping ratio	Fitting error (%)
Improved algorithm	4.8159	$F_1 = 4.9757$, $F_2 = 8.0356$, $F_3 = 12.0293$	$\xi_1 = 0.0475$, $\xi_2 = 0.0356$, $\xi_3 = 0.0220$	0.0455
	1.8785	$F_1 = 4.9635$, $F_2 = 8.0755$, $F_3 = 11.9973$	$\xi_1 = 0.0481$, $\xi_2 = 0.0364$, $\xi_3 = 0.0240$	0.1010
Levy algorithm	4.8159	$F_1 = 4.9401$, $F_2 = 7.8962$, $F_3 = 11.9962$	$\xi_1 = 0.0385$, $\xi_2 = 0.0368$, $\xi_3 = 0.0261$	0.3381
	1.8785	$F_1 = 36.3484$, $F_2 = 8.0138$, $F_3 = 12.1775$	$\xi_1 = 0.0116$, $\xi_2 = 0.0286$, $\xi_3 = 0.0360$	7.0846

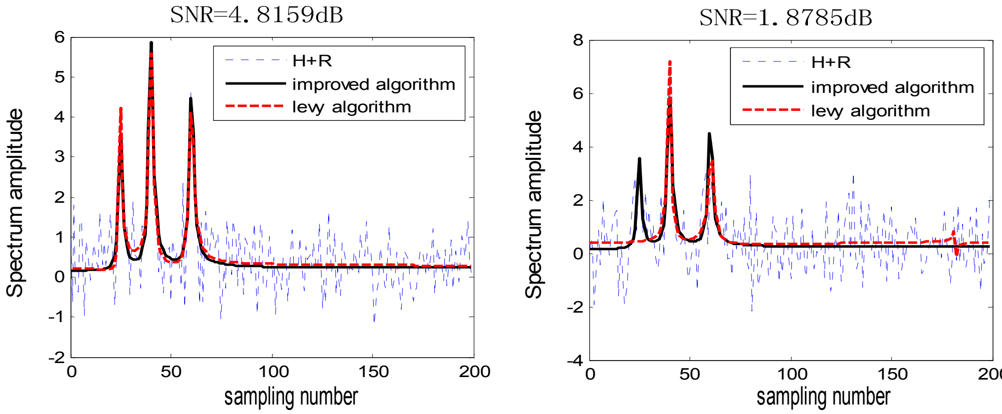

Fig. 1: The two algorithms in different noise intensity

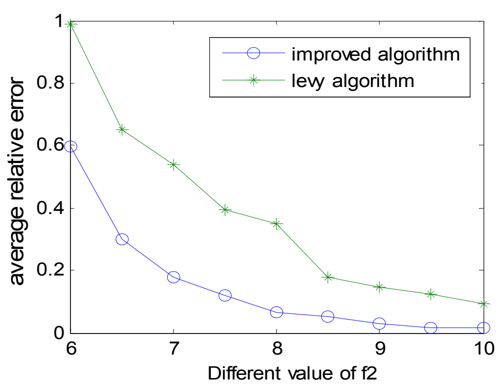

Fig. 2: Two algorithm's recognition error in different frequency space

From the Fig. 1, the algorithm in this study and Levy algorithm can both identify the two modes when the simulation signals SNR = 4.8159 dB and fitting errors are 0.0455 and 0.3381, respectively. But when SNR = 1.8785 dB, the noise is so strong that the frequency f_1 = 5 Hz is almost covered by noise, levy algorithm's fitting error is absolutely larger than the improved algorithm's. To be Specific, using the outliers cluster fitting method can effectively identify two mode parameters and the fitting error is 0.1010. Furthermore, this algorithm's fitting error keeps within 0.2 when signal strength rises. This experiment results show the algorithm can still precisely identify mode parameters in strong noise cases.

- The influence of Different frequency space is verified between the improved algorithm and Levy method. Like S_1, the signal S_2 is given that $f_1 = 5$

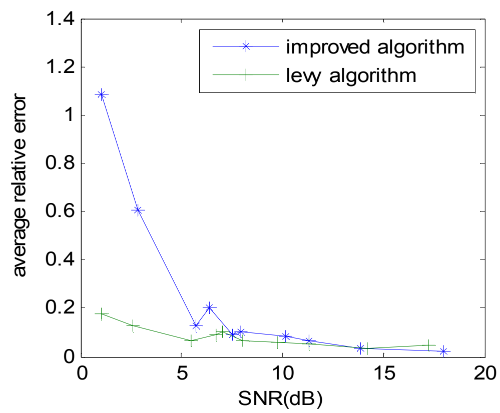

Fig. 3: Average relative error comparison between two algorithms in different SNR

Hz $\delta_1 = 1.5$, $\delta_2 = 1.8$ segam = 0.02 and the variable f_2 'range is 6-10. For each f_2 we calculate the parameters identification of relative error (ε) by using the improved algorithm and levy algorithm.

Figure 2 shows that the algorithm this study put forward performs better than Levy algorithm and its mode parameter identification average relative error is less than 0.6. The Fig. 2 also suggests that the parameter identification average relative error when dealing with two intensive modes is much bigger than two far apart modes. So this experiment turns out that this algorithm performs better in identifying the dense mode parameters than Levy algorithm. Next, the accuracy of the algorithm under different SNR is evaluated.

We still use S_2 to do the test. Suppose that the frequency $f_2 = 8$ Hz is regular but the noise intensity changes. Then the average relative error is calculated in different SNR and the result is in Fig. 3 shows:

In Fig. 3, the noise intensity (segam) changes and for each (segam) value, we randomly generated 500 different noise signal to add to the signal S_2 and calculate the mode parameters average relative error. From the graph, when the SNR >10 dB, both of the two algorithm's parameters identification average relative error is less than 0.2. However, When the SNR is between 5 and 10 dB, the improved algorithm's parameters identification average relative error is more accurate than the Levy algorithm's. When SNR <5 dB, Levy algorithm identification error is larger, but the algorithm does not change obviously, still keeping in the range 0.2.

In a word, these above simulation experiments have proved that the improved Levy algorithm based on outliers clustering in this study is more accurate to identify dense mode parameters than the existing Levy algorithm and has stronger ability to resist the noise.

CONCLUSION

This study has presented here a new mode parameters identification algorithm using the spectrum of clustering fitting method. The outliers are eliminated by the K-means clustering algorithm and the rest of the available frequency response points are used to estimate the mode parameters. Plenty of simulation experiments show this multi-mode fitting method cannot only obtain accurate mode parameters value but also has good interference rejection. It also turns out that this algorithm performs better in different frequency space than Levy algorithm. It can effectively identify two main frequencies being close. And in different noise intensity, the algorithm also performs well. This study proves that this algorithm has the advantage of high stability, low noise sensitivity.

ACKNOWLEDGMENT

This study has been financially supported by state natural sciences funds of China, grant No. 61141015. And this study also has been financially supported by Nature Science Funds of Zhejiang Province, China, grant No. Y1110161. This study has been financially supported by Nature Science Funds of Ningbo city, China, grant No. 2011A610181.

REFERENCES

Fu, Z. and H. Hua, 2000. The Theory and Application of Mode Analysis [M]. Shanghai Jiao Tong University Press, Shanghai, pp: 53-109.

Ge, Y., Q. Meng and Y. An, 2006. A method of removing the outliers in data processing of miss-distance based on wavelet transform. J. Spacecraft TTC Technol., 03: 64-67.

Liu, X., Q. Li and W. Guo, 2000. Research on a new method of detecting abrupt information in mechanical vibration signal. Chin. J. Mech. Eng., 05: 69-77.

Richardson, M.H., 1986. Global frequency and damping estimates from frequency response measurements. Proceedings of the 4th International Mode Analysis Conference, Los Angeles, CA, pp: 465-470.

Verboven, P., P. Guillaume and B. Cauberghe, 2005. Multivariable frequency-response curve fitting with application to mode parameter estimation. Automatica, 41(10): 1773-1782.

Xu, H. and X. Song, 2011. The application of new developed mode parameter identification to structure engineering. J. Trans. Sci. Eng., 04: 24-30.

Ye, Q. and T. Wang, 2009. Optimization mode analysis with PSO based on spectrum segmentation. Chin. J. Sci. Instr., 30(8): 1584-1590.