中國機械工程學刊第三十九卷第一期第 113~120 頁(民國一百零七年)
Journal of the Chinese Society of Mechanical Engineers, Vol.39, No.1, pp 113~120 (2018)

Study on Relation Between Noise and Matrix Dimension of Data-driven Stochastic Subspace Identification Method

Jun-feng Xin*, Yong-bo Zhang** and Jin-lu Sheng***

Keywords: hankel matrix; denoising; data-driven subspace identification method

ABSTRACT

As a linear system identification method, the data-driven stochastic subspace method can effectively obtain modal parameters from the structural signal under ambient excitation. The noise reduction ability of the method is related with its Hankel matrix dimension. The relation between the noise and Hankel matrix of data-driven subspace identification method was introduced theoretically. And a verification procedure was proposed to justify that the noise can be eliminated properly by the data-driven subspace identification method with selected Hankel matrix. The procedure includes SVD, stability diagram and finite element result (FE) . The results of numerical study and jacket platform vibration test demonstrate that the data-driven stochastic subspace identification method with non-square Hankel matrix is of better capability of denoising and can estimate modal parameters with higher accuracy.

Effective modal identification method can obtain the relevant parameters of the structure, which can accurately get the basic health status of large structures. Many scholars at home and abroad (Ewins et al., 1984; Juang et al.;1985; Overschee et al., 1993) proposed modal parameter identification method based on time-domain response, such as time series, random decrement method, the natural excitation technique, data-driven stochastic subspace method. Among them, the data-driven stochastic subspace method is by far one of the more advanced modal parameter identification methods under environmental incentives. It can extract modal parameters of large structures such as offshore platforms more accurately. However, to determine the Hankel matrix dimension is the key to the effective application of this method based on Peeters(1999), and different Hankel matrix dimensions will result in different noise-canceling capabilities of data-driven stochastic subspace method(Maia et al., 1997; Peter et al.,1996; Li et al.,2011). How to determine the dimensions of Hankel matrices? Current research on this issue is still rarely reported.

In response to these problems, we studied the theoretical relationship between Hankel matrix dimension and the data-driven stochastic subspace noise-canceling method, then we proposed Hankel matrix dimensions selection method of data-driven stochastic subspace identification method. And we discussed the relationship between data-driven stochastic subspace noise-canceling capabilities and different Hankel matrix build ways.

Paper Received November 2015. Revised April, 2016. Accepted August, 2016. Author for Correspondence: Jin-lu Sheng

* *Department of Naval Architecture, Qingdao University of Science & Technology, Qingdao 266100, China*

** *National Ocean Graphic Center, Qingdao 266071, China*

*** *College of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, China*

MATHEMATICAL THEORY

Data-driven stochastic subspace identification method

Under the premise of considering only the random noise, the discrete state space equation of the vibration system can be expressed as:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k \end{aligned} \tag{1}$$

$x_k \in R^{\mathrm{n}}$ is the system state vector in the discrete period time of k , and n is the order of the system; $A \in R^{n\times n}$ is discrete state matrix; $C \in R^{l\times n}$ is output matrix that describes how the internal state is changed into measured values of the outside world, l is the number of measuring points; $w_k \in R^n$ is the process noise caused by interference and model error ; $v_k \in R^l$ is the measuring noise caused by sensor and other errors. Both of the two noises are non-measurable and are often assumed to be zero mean, stationary white noise in the derivation process. Two noises covariance matrix can be expressed as the following formula:

$$E\left[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \begin{pmatrix} w_q^T & v_q^T \end{pmatrix}\right] = \begin{pmatrix} Q & S \\ S^T & X \end{pmatrix}\delta_{pq} \quad (2)$$

E is the mathematical expectation operator; $Q \in R^{n\times n}$, $S \in R^{n\times l}$, $X \in R^{l\times l}$; δ_{pq} is kronecker function; $\delta_{pq} = 1\ (p = q)$, p and q represents different time points; $\delta_{pq} = 0\ (p \neq q)$, $E[w_p] = 0$, $E[v_p] = 0$. Data-driven stochastic subspace identification method is discussed in detail in reference based on Peeters(1999).

Relationship between Hankel matrix dimension selection and noise

P , the projection matrix of data-driven stochastic subspace method, can be usually further expressed as follows:

$$P = \begin{pmatrix} p_1 & p_2 & \cdots & p_j \\ p_2 & p_3 & \cdots & p_{j+1} \\ \vdots & \vdots & \cdots & \vdots \\ p_i & p_{i+1} & \cdots & p_m \end{pmatrix} = \begin{pmatrix} s_1 & s_2 & \cdots & s_j \\ s_2 & s_3 & \cdots & s_{j+1} \\ \vdots & \vdots & \cdots & \vdots \\ s_i & s_{i+1} & \cdots & s_m \end{pmatrix} + \begin{pmatrix} n_1 & n_2 & \cdots & n_j \\ n_2 & n_3 & \cdots & n_{j+1} \\ \vdots & \vdots & \cdots & \vdots \\ n_i & n_{i+1} & \cdots & n_m \end{pmatrix}$$
$$= S + N \quad (3)$$

S represents the true signal; N represents noise that needs to be eliminated; i means 1/2 of Hankel matrix row number; j represents the columns number of the matrix; m represents the total amount of data.

After the singular value decomposition (SVD), the threshold value (the order mode) is set as n , and the equation (3) is changed into the following formular:

$$\begin{aligned} P &= U\Sigma V^T = U_S \Sigma_S V_S^T + U_N \Sigma_N V_N^T \\ &= [U_S U_N]\begin{bmatrix} \Sigma_S & 0 \\ 0 & \Sigma_N \end{bmatrix}[V_S V_N]^T = P_S + P_N \end{aligned} \quad (4)$$

$P_S = U_S \Sigma_S V_S^T$; $P_N = U_N \Sigma_N V_N^T$; $U \in R^{i\times i}$; $\Sigma \in R^{i\times j}$; $V \in R^{j\times j}$; Σ represents all the singular values of the measured signal after projection when the decomposition is finished; Σ_S is the singular values of true signal after projection, on the condition of setting a threshold value n . Σ_N is the singular values of noise signal after projection, on the condition of setting a threshold value n .

From formula (4) we can see that singular value decomposition can divide the projection matrix P into two unrelated spaces, the true signal projection P_S and noise signal projection P_N .

The following discussion is focused on the relationship between noise projection P_N and two parameters , the dimension number (rows i or columns j) as well as decomposition singular value n of all the measurement signal projection P .

According to formula (4) we can get formula (5) as follows:

$$\begin{aligned} P^T P &= (P_S + P_N)^T (P_S + P_N) \\ &= P_S^T P_S + P_S^T P_N + P_N^T P_S + P_N^T P_N \end{aligned} \quad (5)$$

Since the real signal and the noise is not relevant, we can get the following formula:

$$P_S^T P_N = P_N^T P_S = 0 \quad (6)$$

According to formula (5) and formula (6) , we can get formula (7) as follows:

$$P^T P = P_S^T P_S + P_N^T P_N \quad (7)$$

The covariance of the measurement signal, the true signal and noise projection are as follows:

$$C = \frac{1}{i}P^T P , \quad C_s = \frac{1}{i}P_S^T P_S , \quad C_N = \frac{1}{i}P_N^T P_N \quad (8)$$

So, formula (7) can be rewritten as the following formula,

$$C = C_S + C_N \quad (9)$$

Combining formula (4) and formula (8), the formula (9) can be written as follows:

$$C = \frac{1}{i}P^T P = \frac{1}{i}V\Sigma U^T U \Sigma V^T = \frac{1}{i}V\Sigma^2 V^T$$
$$C_S = \frac{1}{i}P_S^T P_S = \frac{1}{i}V_S \Sigma_S U_S^T U_S \Sigma_S V_S^T = \frac{1}{i}V_S \Sigma_S^2 V_S^T$$
$$C_N = \frac{1}{i}P_N^T P_N = \frac{1}{i}V_N \Sigma_N U_N^T U_N \Sigma_N V_N^T = \frac{1}{i}V_N \Sigma_N^2 V_N^T \quad (10)$$

From formula (8) to formula (10) , we can get the following result:

$$\frac{1}{i}V\Sigma^2V^T=\frac{1}{i}V_S\Sigma_S{}^2V_S{}^T+C_N \qquad (11)$$

Formula (11) multiplies left by iV^T and right by V simultaneously on both sides, and the results are shown as follows:

$$\Sigma^2=\Sigma_S{}^2+iV^TC_NV$$

$$C_N=\frac{V(\Sigma^2-\Sigma_S{}^2)V^T}{i} \qquad (12)$$

Formula 12 shows that under the premise of the same amount of data m, V and Σ_S are depending on the number of Hankel matrix dimension (rows number $2i$ or columns number j, $2i+j-1=m$) and the threshold value of projection matrix singular value decomposition n; Σ depending on the dimension number of Hankel matrix (rows number $2i$ or columns number j, $2i+j-1=m$) . So, if the threshold value n of the projection singular value decomposition unchanged, the dimension number of Hankel matrix is the main factor affecting noise C_N.

Stabilization diagram

In this paper, the Fourier transform is used as the background of the stabilization diagram. The stabilization diagram indicates the modal identification results under the condition that the number of modes order is between 0 to 30 Hz. For the same modal, if two adjacent modal identification result in a 1% frequency error, while the response damping recognition results in a 5% error, that is,

$$\frac{\left|f^{(n)}-f^{(n+1)}\right|}{f^{(n)}}<1\% \quad , \quad \frac{\left|\xi^{(n)}-\xi^{(n+1)}\right|}{\xi^{(n)}}<5\% \, ,$$

Then, the recognition result is identified as a stable, otherwise unstable based on Li et al.(2011).

DATA-DRIVEN STOCHASTIC SUBSPACE METHOD---THE METHOD OF SETTING AND EVALUATING MATRIX DIMENSION NUMBER

Normalizing singular value (SVD), stabilization diagram and finite element modal recognition results (FE), we construct a three-dimensional integrated assessment method. And each step of the operation flow and function are shown in Figure 1.

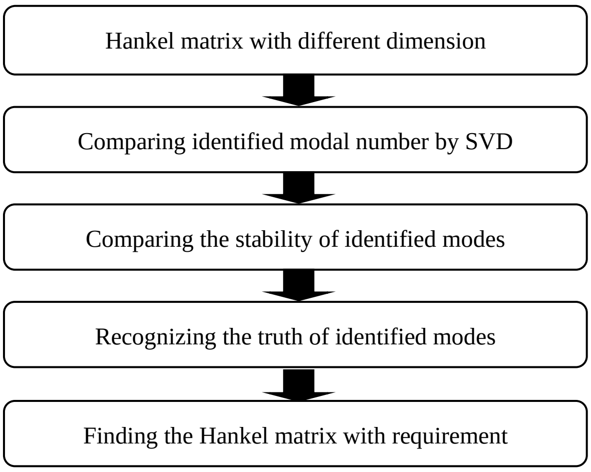

Fig.1. Operation process of evaluation method

HANKEL MATRIX CONSTRUCTION PLAN

Set the total of data is m ($2i+j-1=m$), which studies four Hankel matrix dimension number cases (2i/m=2/10， 3/10， 4/10， 5/10) respectively . Hankel matrix H is a relatively flat non-square one in Case 1 (2i/m=2/10), whereas H is a square in Case 4 (2i/m=5/10). So the plan shows the process of Hankel matrix change from a flat non-square into a square.

NUMERICAL EXAMPLES

This section will go through five degrees of freedom mass - spring - damper system white noise excitation test (model shown in Figure 2) to study the impact on data-driven stochastic subspace noise cancellation capabilities with different dimensions of Hankel matrix. The effectiveness of combination of SVD, Stabilization diagram and FE will be evaluated at the same time.

Five degrees of freedom mass - spring - damper system

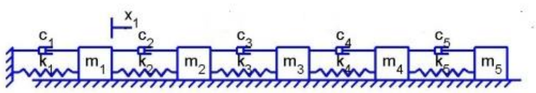

Fig.2. A 5-DOF mass-spring-dashpot system

Each unit has the same mass, stiffness and damping , and is respectively set to : m_n= 50kg, k_n= 2.9 *10^7 N/m , c_n= 1000N˙ s/m . x_n is displacement, and n = 1, ..., 5. By eigenvalue analysis, theoretical values of modal frequencies of 5 were obtained: 34.499, 100.700, 158.730, 203.880, 232.520 Hz; theoretical values of modal damping ratio of 5 were: 0.003737, 0.010909, 0.017197, 0.022092, 0.025198 based on Li et al.(2011).

White noise excitation loading and data acquisition

Gaussian white noise (Figure 3) was used as an input, which was generated by the Matlab randn function, with mean zero and standard deviation of 1.006. The white noise was loaded against the right-side of the first fixed mass point (Figure 2). Vibration response was generated from lsim function. 9000 response data were extracted from the first fixed mass with a sampling frequency of 500Hz, as is shown in Figure 4.

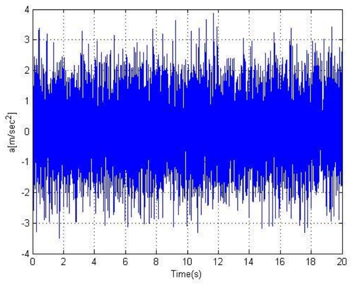

Fig.3. White noise loading

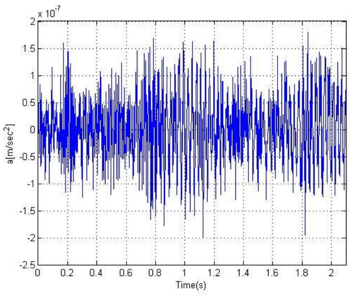

Fig. 4. Response of mass close to fixed point

Comparison of different cases

First, normalized singular value was used to compare the number of modes that was identified by different cases.

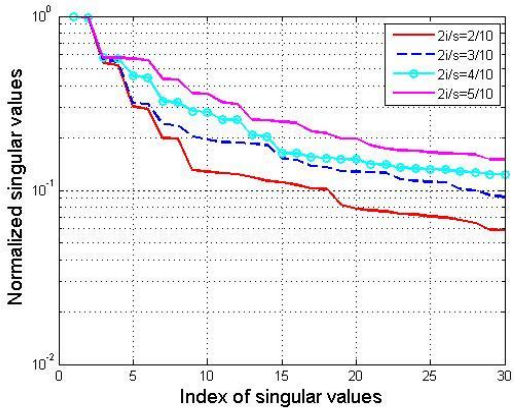

Fig. 5. Normalized singular values in four cases

Normalized singular values (Figure 5) clearly shows: When the Hankel matrix transformed gradually from flat (Case 1, 2i / m = 2/10) to square (Case 4, 2i / m = 5/10), the number of model that can be clearly identified by data-driven stochastic subspace modal method reduced from 4 (Case 1, the last larger gap of the red line appeared when the horizontal axis is 8) to 1 (Case 4, the last larger gap of the purple line appeared when the horizontal axis is 2). This shows that during the process Hankel matrix tends to square, data-driven stochastic subspace noise-canceling capability weakened gradually, which results in less modal can be identified.

Next, stability diagram analysis was used to compare the stability of modals generated from different cases, shown in Figure 6.

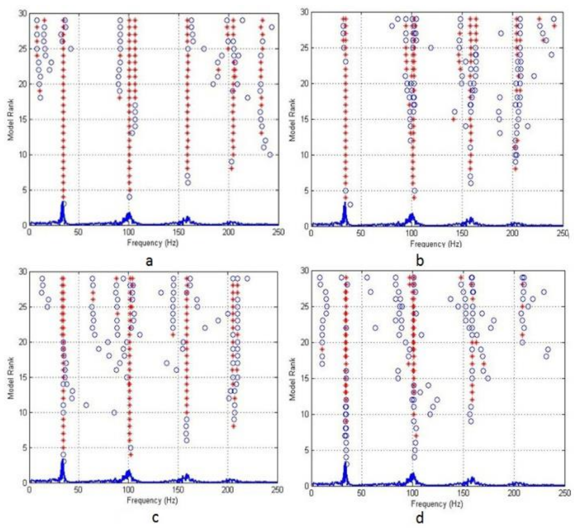

Fig.6. Stability Diagrams for SSI-data in four cases (a-d responding to case 1 (2i/m=2/10) to case 4, (2i/m=5/10), *stable, ° not stable)

Stability diagram clearly shows the process that the Hankel matrix tends to be square (Figure 6, a-d):

a. Data-driven stochastic subspace method could not identify steadily the fourth modal around 200Hz and the fifth modal around 230Hz gradually (The number of circles that indicates instability gradually increased; Peak corresponding linear disappeared or was in chaos). This shows that the data-driven stochastic subspace noise-canceling capability gradually weakened, causing it to become less sensitive to weaker modal.

b. Using data-driven stochastic subspace method to estimate the strongest two modes (The first modal peak is at around 30Hz range and the second modal peak around 100Hz range) gradually become unstable (Circle that represents unstable gradually replaces asterisks that represents stable). This proves that the data-driven stochastic subspace method’s ability of de-noising gradually weakened, leading to its recognition of the strong modal also weakened.

Finally, true and false of modes identified by different cases were determined according to the results of the modal finite element analysis (FE). Table 1 and Table 2 list the modal frequency and damping ratio of different cases.

Table 1. Frequencies estimated in four cases using SSI/data

Mode	FE(Hz)	Case 1	Case 2	Case 3	Case 4
1	34.499	34.499	34.183	34.242	34.148
2	100.700	100.593	101.259	101.295	101.277
3	158.730	159.108	158.359	158.181	158.633
4	203.880	204.563	202.828	205.889	-
5	232.520	232.453	-	-	-

Table 2. Damping ratios estimated in four cases using SSI/data

Mode	FE	Case 1	Case 2	Case 3	Case 4
1	0.003737	0.004131	0.008159	0.006852	0.006032
2	0.010909	0.011083	0.011901	0.008772	0.008942
3	0.017197	0.017849	0.012354	0.011258	0.006311
4	0.022092	0.020996	0.041446	0.016069	-
5	0.025198	0.024856	-	-	-

Table 1 and Table 2 show the authenticity of the five modes that Case 1 identified. And we can also observe from the two tables that: During the process that the Hankel matrix changes from square (Case 4) to non-square (Case 1) data driven stochastic subspace modal parameter estimation method gradually accurate estimate especially for modal damping ratio.

In addition, data with different noise levels were used and the conclusions obtained were consistent with the analysis carried out above.

In short, numerical examples proved :(1) Hankel matrix of data-driven stochastic subspace method should be set to a non-square form; (2) the SVD, stability diagram and FE combination method can effectively evaluate the impact that different Hankel matrixes have on the data-driven stochastic subspace modal identification capabilities.

JACKET PLATFORM SHAKING TABLE TEST

This section will be steel jacket platform physics model white noise excitation test to further validate the above conclusions. Test program: using hydraulic shaker produces white noise excitation, at the bottom of the steel jacket platforms physical model is loaded (with the x-axis and y-axis angle of 45 degrees), to obtain the vibration response of the structure data, and then compare different data-driven program stochastic subspace noise-canceling capability.

Jacket platform physical model and test

The physical model used in the steel jacket made in the Y to set up three main beams, X to set up two, three layers horizontal cross brace, which are located -88cm, 142cm and 188cm at the design level crossbars at 88cm to -142cm located at the vertical diagonal brace (Figure 7), the size of the main components are: the main leg 25 × 2.5mm, horizontal struts 16 × 1.5mm, the diagonal brace 16 × 1.5mm.

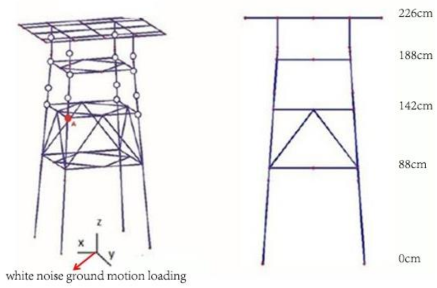

Fig.7. Physical model of jacket platform, white noise ground motion loading position and the sensor A location

The first 9 modal steel jacket platform finite element model are shown in Table 3,

Table 3. The first nine modal frequencies of FE

Component	Mode	Frequecy(Hz)
1	x-1	17.91
2	y-1	22.22
3	ɵ-1	27.31
4	x-2	43.69
5	y-2	46.32
6	ɵ-2	59.71
7	x-3	135.07
8	y-3	137.73
9	ɵ-3	161.84

Hydraulic shaker generate white noise excitation, shown in Figure 8. With a sampling frequency of 500Hz, A y direction sensor 1024 data points are extracted (Figure 8).

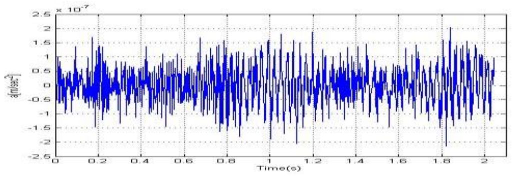

Fig.8. Response of sensor A from white noise ground motion

Comparison of different cases

With numerical examples considered in the analysis process consistent with the normalized first singular value comparison of different solutions.

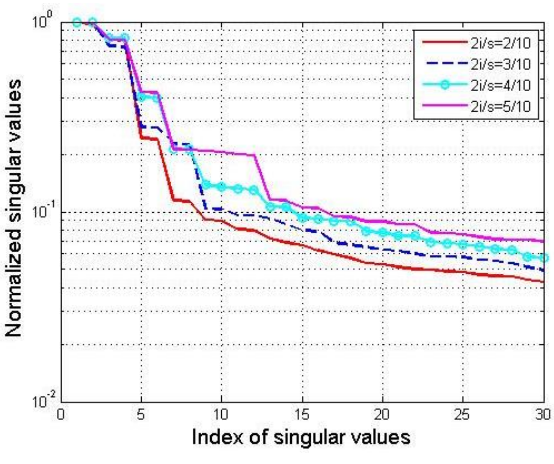

Fig.9. Normalized singular values in four cases in jacket platform model test

Normalized singular values (Figure 9) clearly shows: When Hankel matrix H tends to square the process, stochastic subspace method can identify the modal number of data-driven increase. From 4 (Case 1 (2i / m = 2/10), a large gap in the horizontal red line for the final eight when) gradually increased to 6 (Case 4 (2i / m = 5/10), Purple Line final large gap in the horizontal axis is the time 12). This result seems to indicate that the square of the Hankel matrix to make data-driven stochastic subspace identification more modes, in order to confirm this, combined with stability diagram below, to continue the analysis.

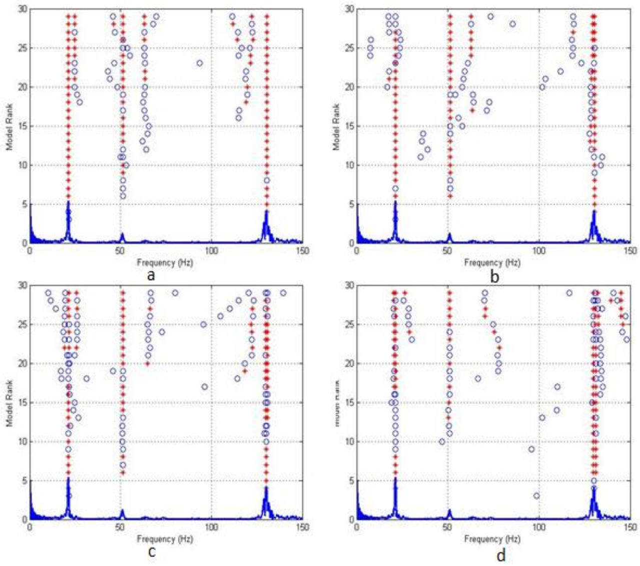

Fig.10. Stability Diagrams for SSI-data in four cases in jacket platform model test.
(a-d responding to case 1, 2i/s=2/10 to case 4，2i/s=5/10, *stable,º not stable)

With tends square Hankel matrix (Figure 10, a to d), from the stability diagram can be observed:

a. At 30Hz, 50Hz, 130Hz three largest peak, data-driven stochastic subspace identification results more and more unstable (on behalf of the unstable circle gradually increased).

b. Case 1 (2i / m = 2/10, Fig. `10 (a)) and Case 2 (2i / m = 3/10, Figure 10 (b)) can identify the four stable modes than other multi-modal recognition program, but they differ in the order of the fourth mode of recognition: a program identified in the fourth-order mode of about 60Hz, and the second solution is about 130Hz.

c. Only. Case 1 (2i / m = 2/10, Figure 10 (a)) Analysis and Case 2 (2i / m = 3/10, Figure 10 (b)) stability diagram results with normalized singular value analysis is consistent: to identify the four modes, this phenomenon shows that the need for the proposed second step assessment methods (stability diagram analysis).

In the analysis of the stability diagram, Case 1 and Case 2 identify a modal too more, but there have been inconsistencies in the fourth-order modal identification. Thus, we need to identify the FE as a reference to identify the authenticity of the modes. Table 4 and 5 list the results of the assessment of different schemes (modes of order is set as 20).

Table 4. Frequencies estimated in four cases using SSI/data

Mode	Direction	FE	Case 1	Case 2	Case 3	Case 4
1	y	22.221	21.432	21.410	21.416	21.413
2	y	46.326	51.269	51.237	51.034	46.929
3	ɵ	59.719	64.533	--	--	--
4	y	135.070	130.219	130.305	130.190	131.058

Table 5. Damping ratios estimated in four cases using SSI/data

Mode	Direction	Case 1	Case 2	Case 3	Case 4
1	y	0.002634	0.002765	0.002245	0.001256
2	y	0.006254	0.005556	0.011845	0.666656
3	ɵ	0.009534	--	--	--
4	y	0.004356	0.003956	0.001423	0.001545

In Tables 4 and 5, reference result of FE can be observed:

a. Fourth order mode should be around torsion mode 63 Hz. The recognition result of Case 1 (2i / m = 2/10) is correct.

b. Only Case 1 (2i / m = 2/10) can identify about 63Hz fourth modal that proved the advantages of a program, and the results show that the proposed evaluation method third step (FE verification) of necessity.

In summary, consistent with the analysis results jacket platform white noise test and numerical tests: raised above assessment method is effective and non-square matrix of Hankel make data-driven stochastic subspace have more capacity and higher noise-canceling modes recognition accuracy.

CONCLUSIONS

A key requirement to the success of accurately estimating modal parameters from noisy measurements is a proper separation of the system information (signal) and noise from the measured data for data-driven stochastic subspace of Hankel matrix dimension. In this paper, SVD, stability diagrams and FE were combined to choose the Hankel matrix dimension number of data-driven stochastic subspace method. First SVD was used to compare modal numbers of different Hankel matrix identified by data-driven stochastic subspace method. And the Hankel matrix that can identify more modal is selected. Then stabilization diagram was used to compare modal stability of different Hankel matrix and choose the most stable one. Finally, the result of FE was used as a reference to determine the true stable model and then further filter out the Hankel matrix with the most genuine modal number.

Based on this method, numerical analysis and shaking table tests jacket platform, systematically explores the relationship between data-driven stochastic subspace of Hankel matrix dimension between its noise-canceling capability. That proved: To maximize the data-driven stochastic subspace noise-canceling capability and sensitivity to weak modal, Hankel matrix should be used in the form of a non-square (ie, the number of rows is greater than the number of columns), that is the number of block-rows and the number of block-columns are not chosen to be closest to each other. For future work this paper can be data-driven stochastic subspace method provides an effective promotion and application parameters.

ACKNOWLEDGMENTS

The study has been supported by Natural Science Foundation of China (Grant No.51609120), Science and Technology Plan for Shandong University (Grant No. J16LB7) and Qingdao Urban and Rural Construction Committee (Grant No. JK2015-15).

REFERENCES

Ewins,D.J., *Modal Testing: Theory and Practice*， Research Studies Press Ltd(1984).

Juang,J.N.,Pappa,R.S.,"An Eigensystem Realization Algorithm for Modal Paremeter Identification and Modal Reduction," *Control and Dynamics*, Vol. 8, No.5, pp. 620-627(1985).

Overschee, P. V., Moor, A. B. D., "Subspace Algorithms for The Stochastic Identification Problem，"*Automatica*,Vol.29,No.3,pp.649-

660(1993).

Peeters, B., Roeck, G. D. "Reference-based Stochastic Subspace Identification for Output-only Modal Analysis," *Mechanical Systems & Signal Processing*, Vol.13, No.6, pp.855-878 (1999).

Maia,N.M.M. "Theoretical and Experimental Modal Analysis," *Bmc Public Health*, Vol.13, No.6, pp.1-5(1997).

Peter, V.O., Bart.,D.E.M., *Subspace Identification for Linear System: Theory-Implementation-Applications*, Kluwer Academic Publishers ,Boston (1996).

Li, H., Li, P., Hu, S. L. J., "Modal Parameter Estimation for Jacket-type Platforms Using Free-vibration Data," *Journal of Waterway Port Coastal & Ocean Engineering*, Vol. 37, No.5, pp.234-245(2011).

NOMENCLATURE

A discrete state matrix

C output matrix

E the mathematical expectation operator

N noise that needs to be eliminated

P the projection matrix of data-driven stochastic subspace method

P_N noise signal projection

P_S true signal projection

S the true signal

c damping of each unit

i 1/2 of Hankel matrix row number

j the columns number of the matrix

k stiffness of each unit

l the number of measuring points;

m mass of each unit

n the order of the system

v_k the measuring noise caused by sensor and other errors

w_k the process noise caused by interference and model error

x_k the system state vector in the discrete period time

x_n displacement of each unit

δ_{pq} kronecker function

$\sum$ all the singular values of the measured signal after projection

$\sum_N$ singular values of noise signal after projection, on the condition of setting a threshold value n.

$\sum_S$ the singular values of true signal after projection