2005

Modal identification using optimization approach.

Li Li
University of Windsor

MODAL IDENTIFICATION USING OPTIMIZATION APPROACH

BY

L I LI

A Thesis

Submitted to Faculty of Graduate Studies and Research through the

Department of Civil and Environmental Engineering

in Partial Fulfillment of the Requirements for the Degree of

Master of Applied Science

at the

University of Windsor

Windsor, Ontario, Canada

2005

Your file Votre référence
ISBN: 0-494-09870-8
Our file Notre référence
ISBN: 0-494-09870-8

ABSTRACT

This thesis deals with the problem of identification of modal parameters using vibration response time histories. Particular attention is given to the identification and the refinement of modal parameters through optimization approaches.

The modal identification problem, also called system identification of modal properties, has received considerable attention in recent years. The identified modal parameters can serve as input to model updating and damage identification in structures, and therefore used to analyze and monitor the operational condition and performance of structures. Furthermore, it is also the only reliable way to determine the damping in a structure.

In this thesis, the modal identification problem is pursued using two different optimization approaches. The first approach is a deterministic optimization approach that minimizes the output model error in the time domain between a direct solution using the modal model and the measured response. Examples of single-input single-output identification are used to illustrate this method; it has been shown this approach is robust against noise and can be used to fine-tune the modal parameter, especially for the damping.

The second approach is based on probabilistic optimization; the objective function is defined as the a posteriori probabilistic density of the parameters given observations/measurements. The conditional probability density is computed using the Bayesian theory of minimum-mean-square-error estimation. Examples of single-output under ambient excitation are simulated to demonstrate this approach. This methodology allows one to obtain not only the estimated parameters in the form of probabilistic mean but also the uncertainties in the form of covariance.

The optimization approaches works though the minimization of an objective function which can be calculated from given set of modal/model parameters. Since there is no gradient or Hessian available for the objective functions defined in this thesis, two direct optimization methods: Nelder-Mead simplex and the Genetic Algorithm are adopted to search the minimum of defined objective functions and thus find the structural parameters.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Faouzi Ghrib for his patient guidance, continuous support and encouragement throughout the course of this study, and for the opportunity he has given me to study system identification and model updating in the future Ph.D. study.

I would like to offer my sincere thanks to Dr. Altenhof and Dr. Das for their advice, encouragement and for their time. I would also like to recognize Dr. Guoqing Zhang at the department of industrial engineering for his excellent lectures on optimization and his suggestions in programming.

I would like to thank my parents and my wife who sacrificed their lives throughout all my life; and last but not least, I would like to thank the new born Arthur for choosing me as his father for his life in this world.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2SLS	Two Stage Least Squares (method)
ARMAX	AutoRegressive Moving Average with eXogenous excitation (model)
ARX	AutoRegressive with eXogenous excitation (model)
BIC	Bayesian Information Criterion
BMSE	Bayesian Mean Square Error
CMIF	Complex Mode Indication Functions (method)
CVA	Canonical Variate Analysis
DOF	Degree of Freedom
ERA	Eigensystem Realization Algorithm
FE	Finite Element
FFT	Fast Fourier Transform
GA	Genetic Algorithm
IV	Instrumental Variable (method)
LMS	Least Mean Square algorithm (for adaptive filtering)
LMS	Linear Multi-Stage (LMS) method
LS	Least Squares (method)
LSFD	Least Square in Frequency Domain (method)
LTI	Linear Time Invariant (system)
MAP	Maximum *a posteriori* (estimation)
MDOF	Multiple Degree-of-Freedom
MIMO	Multiple Input Multiple Output
MLE	Maximum Likelihood Estimation
MMSE	Minimum Mean Square Error (estimation)
MOESP	Multivariable Output-Error State sPace
MOESP	Multivariable Output-Error State sPace
N4SID	Numerical algorithms for Subspace State Space System Identification
NLS	Nonlinear Least-Square
OKID	Observer/Kalman Filter Identification
PDF	Probability Density Function
PEM	Prediction Error Method (for system identification)
PEM	Prediction Error Method
PSD	Power Spectral Density
rms	root mean square
SA	Simulated Annealing
SDOF	Single Degree-of-Freedom
SISO	Single Input Single Output
SVD	Singular Value Decompostion
WGN	White Gaussian Noise

CHAPTER ONE

INTRODUCTION

1.1 Description of the problem

Modal Identification, also called *System Identification of Modal Properties*, has received considerable attention in recent years. Basically, in civil engineering, the task of modal identification is to find modal parameters, including the modal frequencies, modal shape vectors and modal damping from vibration measurements to build a linear dynamic modal model describing the structure's dynamic behaviours.

This identification of structural dynamic characteristics has a wide range of applications, such as: checking the construction quality, validating or improving analytical finite element structural models, or conducting health monitoring and damage detection. Besides, modal identification is possibly the only reliable way of determining damping values for a structural model.

A structure's health or level of damage can be monitored by identifying the changes in structural or modal parameters (Salawu, 1997). These modal parameters can also serve as input to model updating and damage identification in the structure, and therefore used to analyze and monitor the operational condition and performance of structures (Jaishi, 2005). The output of identified parameters can be pursued further for the advancement and the application of the health monitoring technology.

1.2 Research objectives

Although system identification, in general, is a research field conducted mainly by electrical engineers, almost all the main technologies in system identification have their application in modal identification. According to Ljung in his classical work "System identification, theory for the user" (Ljung, 1999), a simple classification of system identification techniques is:

1) Equation solving approaches
2) Minimization approaches (have strong relation with optimization problem)
3) Correlation approaches
4) Subspace approaches (based on the state space innovation model)

The optimization/minimization approaches have their special advantage that the optimality of global solution, if found, is defined clearly with respect to some error functions, such as, in the least mean square sense, or in the minimum mean square sense. Interestingly, the techniques, tools, and algorithms for modal identification are proposed at a steady rate, however the optimization approaches seems to be overlooked by the engineering community, except the prediction error method (PEM).

Optimization is a general tool in engineering and science; it is the minimization or maximization of some cost/objective function. For modal identification problem, the application of optimization approach requires the following steps:

1) Define a cost function to be minimized.
2) The cost function must be a function of modal parameters to be identified.
3) Find or develop an optimization algorithm to solve this problem.

It is the objective of this thesis to work through the identification of modal parameters using single-output measurements with measured input or ambient excitation using optimization approaches. We define two different objective functions: the model output error in time domain, and the a posteriori probability of parameters given measurements. For both of them, we will need to find an algorithm which optimizes a non-linear function generally as follows:

$$\begin{cases} \min_{\theta} \mathrm{J} = \mathrm{J}(\theta); \;\; \theta \in \Re^n, \mathrm{J} \in \Re \\ \text{subject to } \theta_l < \theta < \theta_u, \;\; \theta_l, \theta_u \in \Re^n \end{cases}$$

where θ_l and θ_u represents the bounds of parameters that are physically meaningful, or user prior knowledge to be respected. This algorithm will, hopefully, find the value of parameter θ for which function J is the lowest and for which all the box-constraints are respected. The major difficulty associated with this optimization problem is that we do not know the derivatives of the cost function; we have access to the cost function only. Besides, there can be a limited noise on the evaluation of the cost function.

One reason that the optimization approaches are not so popular possibly lies in the fact that the derivative-based optimization methods are not applicable here; non-derivative-based direct optimization techniques are often problem-specific, and many times requires the development of special algorithms.

It is not the objective of this thesis to develop the "best" optimization program for modal identification problem. The objective is to demonstrate the capability of optimization approach in solving the modal identification problem.

1.3 Organization of the text

Chapter one introduces the thesis by situating the subject, highlighting the objective and clarifying the organization of the text.

Chapter two presents a literature review, including classification of current modal identification approaches, and a brief description of some popular methods.

Chapter three discusses the three most popular system identification methods: the peak-picking based on estimated spectrum plot, the Prediction Error Method and the N4SID (Numerical algorithms for Subspace State Space System Identification) subspace method. The simulation is done using two examples, one SDOF model and a six-storey building model, to illustrate the basic idea of modal identification.

Chapter four discusses the deterministic optimization approach minimizing the model output errors in the time domain. Genetic Algorithm and Nelder-Mead simplex method are used to search the global minimum. The optimization problem for model identification in the time domain is first formulated, followed by a description of optimization algorithms. The reasons for choosing the Nelder-Mead simplex method and Genetic Algorithm are explained and the two methods are combined to solve the nonlinear optimization problem.

Chapter five develops the probabilistic optimization approach that maximizes the posterior PDF of the parameters given observation. The system response under ambient excitation can be approximated by a Markov process of finite order, then the posterior PDF of the Markov process can be expanded as multiplications of conditional PDFs of each observation; the conditional PDFs can be computed using the Bayesian Minimum-Mean-Square-Error (MMSE) estimator along with a steepest descent iteration scheme.

Chapter six summarizes and concludes the thesis, with recommendations for future research.

CHAPTER TWO

LITERATURE REVIEW

2.1 General

Since Cooley and Tukey proposed the decimate–in-time Fast Fourier Transform (FFT) technique in 1965 (Cooley, 1965), it became more efficient to analyze linear time-invariant (LTI) systems in the frequency domain than in the time domain; modal models are intrinsic to LTI systems, therefore at beginning researchers started to formulate modal identification in the frequency domain. Peak-picking on spectrum plot and circle-fitting at frequency response function peaks are the first tools available to modal testers (Ewins, 2000).

The birth of the time domain based methods appeared at a later time, but since then they have attracted more research attention. In the 1970s, Ibrahim published the first time domain method, now called Ibrahim Time Domain method (ITD) in literature (Ibrahim 1976). This method uses the free-response data to build eigen-matrices, and thus identifies the modal parameters. ITD method offers higher accuracy than the peak-picking and circle-fitting in frequency domain, especially for modes with a large damping.

The 1970s also saw the advent of another important time domain method: the Least Square Complex Exponential Algorithm (LSCE) also called Prony method; this method requires an impact response of a structure, it applies the complex exponential relations between the impact response and the residues and poles to identify the modal parameters. This method, though obsolete nowadays, inspired the polyreference complex exponent method in 1982 which was proposed by Harvard Vold in the SDRC company (Structural Dynamics Research Company) (Vold, 1982). It extends the LSCE to multi-input and multi-output case, where impact responses of multiple input and multiple outputs are used to build an impact response matrix; modal parameters are identified based on the complex exponential relations involved.

The application of ITD, LSCE requires special experimental or testing settings to get the free vibration response, thus their applications are limited; however, they can be

applied to transient vibration responses if combined with the Random Decrement Method (RDM). The RDM transforms ambient vibration data into free decay responses, and is another important advancement made in the 1970s (Ibrahim, 1977).

Modal identification is, actually, an application of a more general subject: system identification. This subject has obtained tremendous development in the past twenty years. Much of the important contributions to this subject have been made by electrical and control engineers, rather than structural engineers.

Modal models are very important and useful models for control of structures. An important time domain method, the ERA (Eigensystem Realization Algorithm), originally applied for modal reduction in control, was published by Juang and Pappa in 1985 (Juang, 1985). This method is based on the minimum realization algorithm in control theory (this is why it is called eigensystem realization); it finds the eigenvalues and eigenvectors by the SVD (Single Value Decomposition) of a Hankel matrix composed from impulse response or free response signals. The application of ERA can also be extended to ambient vibration cases if combined with the RDM technique.

Following the idea of ERA, the control researchers worked out another class of system identification methods in the 1990s: the subspace identification method, which marks a cornerstone in the development of system identification. The subspace has a strong relation to the ERA approach; by introducing an Observable Canonical Form (OCF) (OCF is another name for innovative state space model) and the corresponding similarity transformation, it formulate a Hankel matrix using output data, or using the correlation sequences computed from the output data, and do some elegant mathematical operations on this matrix to obtain the system matrices. The subspace identification methods are very powerful methods; however they suffer from some drawbacks, notably the overestimate of damping in higher modes and the poor accuracy with short measurements.

Subspace Identification for Linear Systems is an important reference for all researchers in system theory, control theory, signal processing, automization, mechatronics, chemical, electrical, mechanical and aeronautical engineering. The subspace identification can be considered as a description of a class of related algorithms, not a specific algorithm; it is a recipe or meta-algorithm which can be used to devise particular algorithms. In the classic work of Overschee and Moor, several variations of

the general subspace algorithm are presented. (Overschee, 1997)

In the context of system identification, modal models, FRF models and innovative state space models can be regarded as white-box or grey-box models, that the model structure of the system is known or partly known, and have specific physical meanings.

Unlike the control and structural engineers who work with mostly white-box or grey-box models, the electrical engineer's approach to system identification is more involved with the so-called black-box models, that is, the model structure of the system is not known a priori and the models they build to imitate the system's behaviour do not necessarily have any specific physical meaning. Such black-box models includes: the ARMAX or discrete filter model family, and the neural networks.

Based on the historical development of time series, Yule was the first to propose the AR (AutoRegressive) model of a single variable in 1927. However, this kind of discrete filter model for time series analysis did not gain popularity until the advent and advance of computer technology. Since the 1960s and 1970s, the AR model and more elaborated ARMA (AutoRegressive Moving Average), ARMAX (AutoRegressive Moving Average with eXogenous input) models have been studied thoroughly, different methods for system identification based on these models are proposed from different point of views, including the Least-Squares method, Two Step Least-Square (2SLS) method, instrumental variable method, maximum likelihood method, Linear Multi-Stage (LMS) method, and many more (Petsounis, 2001).

The culmination this approach for system identification is the Prediction Error Method (PEM), sometimes called Ljungian Prediction Error Method to honor the scientist Lennart Ljung for his significant contribution in system identification using ARMAX type discrete filter models. The PEM estimates model parameters through minimizing the optimally determined one-step-ahead output prediction error. Hence, a recursive optimization is necessary, and plenty of user interferences are required (Ljung, 2003). In applications which require some automation, one has to use some reduced forms of ARMAX model, including the AR model, ARX model and ARMA model, which can be identified by Least-Square method, Instrumental variable method, two-step-Least-Square method or Maximum Likelihood Method.

When attempting to identify a black-box model for a dynamic system it is common

practice to follow the procedure depicted in Figure 2.1:

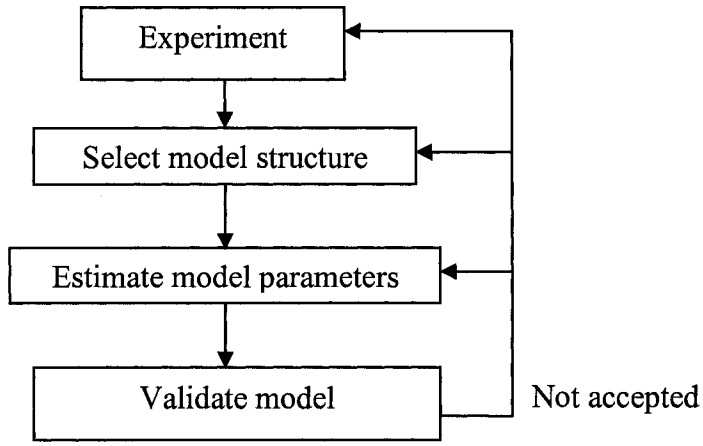


Figure 2.1: block diagram representing the system identification problem (Ljung 1999)

A detailed description of each of these steps is given in Ljung's classic book on system identification (Ljung, 1999). In this black-box model approach, usually the problem of Systems Identification starts with a time series of observed data and tries to determine the simplest model capable of exhibiting the observed behavior. Selecting a suitable model structure is usually the most difficult work; experiment should be designed to provide enough data for specific model identification, and the estimation methods.

It was not until 1972 that Gersch and his co-workers first applied the time-series model to civil engineering (Gersch, 1972). After that, almost every idea developed with the ARMAX model family has been applied to structural identifications. Besides, the structural engineers made their contributions to adapt the system identification for structural modal identification.

Adopting the concept behind Ibrahim's modal identification technique, Huang proposed a least-squares approach to determine the coefficient matrices of the AR model which is a modification of the LS solution for AR models (Huang, 2001). The modification is based on the equivalence between the correlation function matrix for the responses of a linear system subjected to white-noise input and the deterministic free vibration responses of the system. Thus the input-output LS method to identify an AR model is extended to output-only case with the assumption that the input is Gaussian

white noise.

2.2 Classification of modal identification methods

Actually, each of the methods mentioned above serves as a framework for the implementation to different specific problems. There are numerous variations of each of them to account for different experimental settings. To walk out the jungle of the prolific modal identification methods, several commonly made classifications based on different criteria are made in this section.

2.2.1 Time-domain vs. frequency-domain methods

We can group modal identification methods as frequency-domain methods or time-domain methods. Frequency-domain methods employ the FFT (Fast Fourier Transform) technique to transform experimental data into frequency domain, and identify the modal parameters from a FRF model. The peak-picking, curve-fitting, LSCE and polyreference are examples of frequency domain based methods.

Frequency-domain methods have wide applications in civil engineering; the popularity of the method is partly due to its simplistic implementation simplicity and its speed, and partly due to the fact that they have direct physical meanings. However, there exist serious drawbacks for traditional frequency-domain methods; they are very difficult for systems with close-spaced modes. They are also usually difficult for noise-corrupted data (Peeters, 2000).

On the contrast, time domain methods offer more robust performance against noise, and they are more successful at identifying closely spaced or repeated natural frequencies. Some time domain method, like the subspace identification and ERA, also offer a more systematic way of determining the approximate order of a test structure and generally able to identify a larger number of natural frequencies than frequency-domain methods do. Besides, time domain identification does not need to perform the FFT to transform data into frequency domain, and thus avoids some incidental errors.

Time-domain approaches can be divided into non-parametric state-space model methods and parametric time series model methods. Ibrahim Time Domain method (ITD), which is the first time-domain method in history, along with the ERA and subspace methods, are based upon non-parametric state space models. Identification methods using

wavelet transform, neural network are also non-parametric methods. The wavelet-based identification methods are also called time-frequency domain method (Kitada, 1998).

Parametric time-domain identification of vibrating structures refers to the process of developing finitely parameterized mathematical time-domain models of such structures based upon measured excitation and/or response signals.

Parametric (model-based) identification leads to finitely parameterized representations, usually the discrete filter system (the ARMAX model family) models, but also includes difference/differential equation and modal models, then different techniques such as instrumental variable, maximum likelihood and nonlinear-optimization can be employed to estimate the parameters of the system. Usually parametric methods can yield higher resolutions than nonparametric methods in cases when the signal length is short.

2.2.2 Input-output identification vs. output-only identification

Identification approaches can be classified by their input data sources. Typical types of field tests include: (1) ambient vibration tests, (2) forced vibration tests, (3) free vibration tests, and (4) earthquake response measurements. Difference in experiment data can have a significant influence upon the selection and implementation of identification methods.

In the first method, the structure is excited by artificial means such as shakers or drop weights. By suddenly dropping a load on the structure, a condition of free vibration can be induced. The disadvantage of artificial excitation methods is obvious that traffic has to be shut down for a rather long period of time; this could be a serious problem for the intensively used infrastructures.

In contrast, ambient vibration testing is not affected by the disturbances on the structures, because it uses the disturbances induced by traffic and wind as natural or environmental excitation. The ambient vibration testing is a kind of output-only data dynamic test method. The service condition does not have to be interrupted. The output-only dynamic testing has an advantage of being inexpensive since no equipment is needed to excite the structure. Therefore, the modal parameter identification technique through ambient vibration measurements has become a very attractive topic in the area of civil engineering structures. In the context of output-identification, the input signals is usually

modeled as stationary white noise, and in this way many output-only identification methods have been developed that circumvents the unknown inputs.

Much modal identification is accomplished using both input and output data from the structural system. Notably the methods involving the ARMAX model family, including the LS, IV, 2SLS, MLE, LMS and PEM. Other methods who develop a transfer function between the input and output, such as polyreference and LSCE also requires the input-output identification.

Output-only identification methods include the methods that incorporate the random decrement method (RDM), such as with the ITD/RDM and LSCE/RDM, in which the ambient vibration data are transformed into free decay responses using the random decrement technique.

The time-series models, AR and ARMA, can also be used for output-only identification if the input is modeled as white noise, and then attempt to estimate the parameters of that linear system.

The subspace methods can be either input-output or output-only, depending on how the Hankel matrix is formed and the mathematical operations done on the Hankel matrix.

Some special kinds of excitations: earthquake, ocean swells, wind excitation, usually require some special adjustments (Lus, 1999). One should not take for granted that one method can be applied for all problems without modification.

2.2.3 Deterministic analysis and stochastic analysis

LS, Prony, ERA, deterministic subspace are called deterministic methods because they assume the input and output as deterministic sequences without noise (although actually they may be disturbances in the data). These methods generally work well under the special noise-free case, or negligible noise, but works poor in output with non-negligible noise. For example, the LS and ERA are notoriously prone to generate false modes. (Juang, 1986 and Petsounis, 2001)

Stochastic methods (PEM, 2SLS, LMS, stochastic Subspace) are shown to lead to potential advantages in non-negligible noise cases, usually achieve lower bias errors and good overall accuracy, but at a price of increased complexity (Petsounis, 2001).

In the PEM family stochastic methods are based on the ARMAX (AutoRegressive

Moving Average with eXogenous excitation) or ARX-like (AutoRegressive with eXogenous excitation) models.

2.3 Combinatorial system identification

In his doctoral thesis, Peeters (Peeters, 2000) deemed that the identification problem of linear dynamic structures in civil engineering have been largely solved by the introduction of subspace methods and the application of the Ljungian PEM method, and not much can be improved, except the applications and implementation to specific problems, on the quality of the estimated modal parameters.

Possibly he is right, in that there were not breakthroughs like the introduction of subspace identification in recent years; however there is a trend to develop combinatorial approaches for modal identification, i.e., combining different methodologies to improve the quality of the estimated modal parameters.

A combined subspace-maximum likelihood algorithm has been developed to overcome the drawbacks of subspace identification, like overestimation of damping (Koh, 2003). Pridham and Wilson (Pridham, 2004) uses EM (Expectation-Maximization) algorithm to refine subspace identification to improve its performance in short data. For a three degree-of-freedom model it shows that the combined EM-subspace algorithm is able to circumvent some of the difficulties arising from short data sets while using the subspace identification alone, and estimation of damping ratios are also improved. Many recent research efforts belong to this class of combining subspace identification with another method.

CHAPTER THREE

SPECTRUM ESTIMATION, PEM AND SUBSPACE IDENTIFICATION

3.1 General

As a compliment to chapter 2 and a precursor to the optimization approaches in chapter 4 and chapter 5, three most popular modal identification techniques: the peak-picking from estimated spectrum plot, the Prediction Error Method (PEM), and the subspace identification, are briefly described and implemented in this chapter. MATLAB system identification toolbox is used for PEM and subspace identification. The simulation is done using two examples, the first example is a SDOF oscillator and the second example is a six-storey shear building, to demonstrate the general idea of modal identification using single measured output.

3.2 Peak-Picking from estimated Power Spectral Density (PSD) plot

3.2.1 Basic idea of peak-picking

The simplest approach to estimate the modal parameters of a structure subjected to arbitrary loading is called the peak-picking method. This is a frequency-domain spectrum driven identification method. The structural vibration model involved is the frequency-response-function (FRF). Eigenfrequencies are picked as the peaks of a spectrum plot. Probably because of its simplicity and the capability to be implemented digitally in signal processing hardware, it is the most widely used method in civil engineering.

It is based on the fact that the FRFs increases to an extreme around the natural frequencies. However, in the context of ambient vibration measurements, the FRFs are simply replaced by the auto power spectral densities of the ambient outputs without the FRF actually being computed. Due to noises in measurement and sampling effects of the Fourier transform, the damping estimates are unreliable.

If the structure is excited by a random input f(t), resulting in the response x(t), then

we have two simple relations of system transfer function equations in the frequency domain:

(3.1) $$S_{xx}(\omega) = |H(\omega)|^2 S_{ff}(\omega)$$

(3.2) $$S_{xx} = H(\omega) S_{xf}(\omega)$$

where S_{xx} is the spectrum of the response process x(t), S_{ff} is the spectrum of the input excitation, and S_{xf} is the cross-spectrum between the input and the output.

In case of acceleration measurement y(t), the complex modal decomposition of the acceleration spectrum can be written as (Ewins, 2000):

(3.3) $$S_y(s) = \left(\sum_{i=1}^{N} \frac{s^2}{\lambda_i^2 (s - \lambda_i)} \{v_i\} \langle l_i^T \rangle \right) R_u \left(\sum_{i=1}^{N} \frac{s^{*2}}{\lambda_i^2 (s^* - \lambda_i)} \langle l_i \rangle \{v_i^T\} \right)$$

where {v} is the mode shape vector, $<l>$ is the continuous time modal participation vector, s is the Laplace domain variable, λ_i are the complex eigenvalues, and R_u is the input covariance matrix in case of white noise inputs. This expression is the product of two summations where each term represents the contribution of a certain mode. A term of the left factor is proportional to $(s - \lambda_i)^{-1}$ and reaches a maximum if s approaches one of the complex eigenvalues

(3.4) $$\lambda_i = -\zeta_i \omega_i + j\omega_i \sqrt{1 - \zeta_i^2}$$

For low damping ratios, this is achieved around s= $j\omega_i$. If the frequencies are well separated, the spectrum at any eigenfrequency ω_i is dominated by a single mode and can be approximated by:

(3.5) $$S_y(s = j\omega_i) \cong \frac{\{v_i\} \langle l_i^T \rangle R_u \langle l_i^* \rangle \{v_i^H\}}{(\zeta_i \omega_i)^2}$$

here the superscript H stands for Hermitian transpose of a complex vector; this formula is valid for both proportional damping and non-proportional damping, in the latter case the eigenvectors can be complex.

By defining the complex scalar α_i as:

(3.6) $$\alpha_i = \frac{\langle l_i^T \rangle R_u \langle l_i^* \rangle}{(\zeta_i \omega_i)^2}$$

The approximated spectrum at resonance of a response y(t) can be rewritten as:

$$S_y(s = j\omega_i) \cong \alpha_i \{v_i\}\{v_i^H\} \tag{3.7}$$

The interpretation of this equation is that at resonance, each column (or equivalently each row) of the spectrum matrix can be considered as an estimate of the observed mode shape up to some scaling factor. Of course, a mode can not be identified if the column (or row) corresponds to a structural degree-of-freedom that is situated at a node of this mode.

Equation (3.7) is the basis of enhanced frequency-domain identification methods that are able to identify not only frequencies and damping, but also the modal vectors; one such method is the Complex Mode Indication Functions method (CMIF) which can be considered as the singular value decomposition (SVD) enhanced peak-picking. However, for the identification of modal vector multiple-output measurements are required.

Theoretically, the famous half-power bandwidth method can be used to estimate the damping ratios corresponding to each frequency. Assume that ω_1 and ω_2 are the two frequencies left and right from, and as close as possible to the eigenfrequency ω_i, where the magnitude of a certain element of the spectrum matrix is half the resonance magnitude. A damping estimate is then obtained as (Ewins, 2000):

$$\zeta_i = \frac{\omega_2 - \omega_1}{2\omega_i} \tag{3.8}$$

However this estimate is generally not accurate.

3.2.2 Spectrum estimation

Traditionally, the name Peak-picking is usually given to non-parametric estimation methods of power spectrum that use FFT to calculate periodogram as an estimate to the PSD, for this reason, they are sometimes called frequency domain estimation, or Fourier methods.

However, one should not take for granted that all spectrum estimation must be performed in the frequency domain with FFT, there are many parametric methods to estimate the PSD of time series in time domain without the use of discrete Fourier transform. The concept of Peak-Picking can be extended to all the methods of power spectrum estimation. Thus, some methods, such as the Prony method, which are

traditionally not considered as peak-picking, can be classified into the peak-picking family.

Obviously, the accuracy of this method depends on the accuracy of power spectrum estimation. There are already a handful of tools available in estimating spectrum from a given time series (Hayes, 1996), such as:

1) The Periodogram
2) The windowed periodogram
3) Bartlett's method: periodogram averaging
4) Welch's method: averaging windowed periodogram
5) Blackman-Tukey method: periodogram smoothing
6) Minimum Variance spectrum estimation (MV)
7) The Maximum Entropy Method (MEM)
8) ARMA (Autoregressive Moving average) method
9) The Pisarenko's Harmonic Decomposition method
10) The MUSIC algorithm (Multiple Signal Classification Method)
11) The Eigenvector method (EV)
12) The Principle Components Frequency Estimation using the Blackman-Tukey method
13) Neural network spectrum estimation (Zhang, 2002)

And new methods are continued to be proposed at a steady rate.

3.2.3 The periodogram

It is well known that the power spectrum is the Fourier transform of the correlation sequence. However, there is another equivalent definition of the power spectral density of a time sequence {y(n)}:

$$S_y(f) = \lim_{N\to\infty} E\left[\frac{1}{2N+1}\left|\sum_{n=-N}^{N} y(n)\cdot \exp(-i2\pi fn)\right|^2\right] \tag{3.9}$$

where E[] is the ensemble average. This definition is equivalent to the Wiener-Khinqin definition. As an approximation while the data length N is finite, we have the periodogram spectrum estimate: let n = 0,……,N−1

(3.10) $$\hat{S}_y(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} y(n)\exp(-i2\pi fn)\right|^2$$

Thus, the periodogram is proportional to the squared magnitude of the DFT of sequence y(n), and may be easily computed using FFT technique.

Similar to the cross-correlation between multiple sequences corresponding multiple-output measurements, the periodogram can be extended to periodogram matrix. For m number of discrete time series $\{y^{(p)}(n),\cdots\cdots,y^{(m)}(n)\}$,n=0,$\cdots$,N-1; $N=2^M$, the discrete Fourier transform is

(3.11) $$Y^{(p)}(k) = DFT(y^{(p)}(n)) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} y^{(p)}(n)e^{-iz\pi kn/N}\ ,\ \text{P=1,}\cdots\cdots\text{,m;}$$

k=0,$\cdots\cdots$,N-1

The periodogram for one single series is

(3.12) $$S_y(k) = \frac{1}{N}\left|\sum_{n=0}^{N-1} y(n)\exp(-i2\pi kn/N)\right|^2$$

The cross-periodogram of two series $y^{(p)}(n)$ and $y^{(q)}(n)$ is

(3.13) $$S_y^{(p,q)}(k) = \frac{1}{N}\left[\sum_{n=0}^{N-1} y^{(p)}(n)\exp(-i2\pi kn/N)\right]\cdot\left[\sum_{n=0}^{N-1} y^{(q)}(n)\exp(i2\pi kn/N)\right]$$

3.2.4 The modified periodograms

The basic idea of modifying periodogram is to apply different windows to the data sequence, instead of the original rectangle window for cutting out a finite duration sequence. Common choices are the Bartlett window, the Hamming window, etc.

1) The Bartlett's method: the averaged periodogram

Observing that the expected value of the periodogram converges to the true spectrum as the data record length goes to infinity

$$\lim_{N\to\infty} E\{\hat{S}_y(f)\} = S_y(f)$$

Bartlett proposed the sequence y(n) be partitioned into K non-overlapping sequences of length L where N = KL. The Bartlett estimation is then the averaged periodogram

(3.14) $$\hat{S}_y(f) = \frac{1}{N}\sum_{i=0}^{K-1}\left|\sum_{n=0}^{l-1} y(n+jl)\exp(-i2\pi fn)\right|^2$$

with n = 0,1,...,L-1, and j =0,1,...K-1

This new estimate is asymptotically unbiased, and the resolution is also increased.

2) Welch's method

In 1967, Welch proposed two modifications to Bartlett's method. The first is to allow the sequences to overlap, and the second is to allow a data window to be applied to each sequence, thereby produce modified periodograms that are to be averaged. This modification works well for long data series, but for short duration of measurements, it can even decrease the quality of estimate.

3.3 The Prediction Error Method (PEM)

The Prediction Error Method estimates ARMAX (Autoregressive Moving Average with exogenous excitation) models in the form:

(3.15) $$A(q)y(t) = B(q)x(t) + C(q)e(t \mid \theta)$$

With x(t), y(t) representing the measured excitation and measured noise-corrupted response, respectively, e(t) the model residual (one-step-ahead prediction error, which should, for an accurate model, form a zero-mean uncorrelated sequence), and A(q), B(q), C(q) the AR (AutoRegressive), X(eXogenous) and MA (Moving Average) polynomials, respectively, in terms of the backshift operator q, $q^n(x(t)) = x(t-nT)$. θ denotes all the polynomial coefficients (model parameters) to be identified.

The ARMAX model may be expressed in the time domain as:

(3.16) $$y(t) + \sum_{i=1}^{Na} a_i y(t-i) = \sum_{i=0}^{Nb} b_i x(t-i) + e(t \mid \theta) + \sum_{i=1}^{Nc} c_i e(t-i \mid \theta)$$

By dividing the polynomial A(q) on both sides, this model can also be written as

(3.17) $$y(t) = G(q)x(t) + H(t)e(t)$$

Where G and H are transfer functions.

Prediction Error parameter identification is based on the minimization of the quadratic least square criterion, or to minimize the Residual Sum of Squares (RSS) of the error between the measured response y(t) and the computed model response $\hat{y}(t \mid \theta)$.

(3.18) $$RSS(\theta) = \sum_{t=0}^{N-1} [y(t) - \hat{y}(t \mid \theta)]^2$$

This is a nonlinear optimization problem due to the nonlinear dependence of the model residual upon the model parameter vector. This necessitates the use of an iterative optimization such as Gauss-Newton or Levenberg-Marquardt, and requires good initial guess of parameter values. Existence of local minima is also a problem for this nonlinear optimization problem.

By dropping the noise transfer function term, this problem can be simplified to ARMA (AutoRegressive Moving Average) model identification,

(3.19) $$A(q)y(t) = B(q)x(t) + e(t \mid \theta)$$

A simplified optimization problem for minimizing the RSS is obtained, which can be solved without iterations using the Instrumental Variable (IV) method or Maximum Likelihood method (MLM). If the noise term is totally dropped, we get a linearized (noise-free) optimization problem that can be solved directly by solving a normal equation using the Least-Squares (LS) method. However, unlike the simpler ARMA or ARX (Autoregressive with exogenous excitation) model, the ARMAX model offers a noise transfer function capable of representing various types of noise characteristics, so that PEM is more robust against noise corruptions than the Least Squares or Instrumental Variable methods.

Besides, there are also some approximations of the PEM method to estimate the ARMAX model parameters without solving an iterative optimization problem. The most important of them are the Linear Multi-Stage (LMS) method and the Two Stage Least Squares (2SLS) method. The LMS estimates ARMAX models through a sequence of linear Least-Square and deconvolution operations. The 2SLS method consists of two linear Least-Squares based stages to approximately solve the original nonlinear optimization problem. In many implementations, the parameter identified by these methods can serve as an initial guess for the PEM (Ljung, 2004). PEM requires a nonlinear non-convex optimization problem to be solved, for which a good initial guess if required.

For input-output identification without noise or with negligible noise, Least Squares and Instrumental Variable methods usually give good identification; however for noise-

corrupted data, PEM is currently the more accurate one in that it optimizes the model error (Petsounis, 2001).

The determination of a suitable ARMAX model order is crucial for the success of PEM. This is basically related to the determination of the number of structural degrees of freedom. However, in reality, the two issues are decoupled, as order over-determination is generally necessary for accurate identification. Model order determination is based upon successive model fitting and examination of criteria such as the RSS (Residual Sum of Squares) or output error, the Bayesian Information Criterion (BIC), Akaike's Final Prediction Error (FPE), and the Akaike's Information Criterion (AIC).

The RSS is a measure of the quality of fit for stochastic models, but not always a good order indicator as it typically decreases for increasing model order. The BIC overcomes this difficulty by including an additional term that increases model complexity, that is:

$$BIC(\theta) = \ln\left(\frac{RSS(\theta)}{N}\right) + \dim(\theta)\frac{\ln(N)}{N} \tag{3.20}$$

here dim( ) stands for dimensionality of the indicated vector, and N for the data record length. Usually a minimum BIC indicates the optimal model order. Note that both the RSS and BIC may be applied only to those stochastic methods that lead to an uncorrelated residual (prediction error) sequence (thus not to the Instrumental Variable method).

A measure of quality of fit that focuses on the input-output dynamics, and is thus applicable to all methods, is the Output Error (OE),

$$OE(\theta) = \frac{1}{N}\sum_{t=0}^{N-1}[y(t) - \hat{y}(t \mid \theta)]^2 \tag{3.21}$$

Once an estimated model has been validated, its global modal parameters may be extracted as follows:

$$\text{damping ratio:} \quad \zeta_k = \frac{\ln\left(\frac{1}{|p_k|}\right)}{\left[\left(\tan^{-1}\left(\frac{p_k^I}{P_k^R}\right)\right)^2 + \ln^2\left(\frac{1}{|p_k|}\right)\right]^{1/2}}$$

and frequency: $\omega_k = \frac{\ln\left(\frac{1}{|p_k|}\right)}{\zeta_k \Delta t}$

where Δt is the sample time step, p_k^I and p_k^R are the imaginary and real parts of the k-th pole respectively.

3.4 Subspace identification

The subspace identification identifies a full state space model; the most generalized state space model is (Overschee, 1997)

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + w(k) \\ y(k) = Cx(k) + Du(k) + v(k) \end{cases}$$

with

$$E\left[\begin{pmatrix} w_p \\ v_p \end{pmatrix}\begin{pmatrix} w_q^T & v_q^T \end{pmatrix}\right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}\delta_{pq} \geq 0$$

where vectors u_k and y_k are the observations at time instant k of respectively the input and output of the process. The vector $x_k \in \Re^n$ is the state vector of the process at discrete time instant k and contains the numerical values of n states. v and w are unobserved vector signals, usually called the measurement noise and input noise, respectively. It is assumed here that they are zero mean, stationary, white Gaussian noise vector sequences. δ_{pq} is the Kronecker delta that equals 0 if $p \neq q$, and equals 1 if $p = q$.

One has to note that the role played by the input noise w is different from that of measurement noise v: w as an input will have a dynamic effect on the system as well as the deterministic input u, while v only affects the output y directly and therefore is called a measurement noise.

The matrix A is called the (dynamical) system matrix. It describes the dynamics of the system (as characterized by its eigenvalues). B is the input matrix, which represents the linear transformation by which the deterministic inputs influence the next state. C is the output matrix, which describes how the internal state is transferred to the outside world in the observations y. Matrix D is called the direct feed-forward term, generally this term is absent in civil engineering structures. The matrices Q, S and R are the covariance

matrices of the noise sequences w and v. They are assumed to be positive definite, as is indicated by the inequality sign. In structural dynamics, usually the measurement noise process and the input noise process are assumed to be independent, therefore S is all zeros. Besides, the matrix pair {A; C} is assumed to be observable, which implies that all *modes* in the system can be observed in the output y and can thus be identified. The matrix pair {A; [B Q]} is assumed to be controllable, which in its turn implies that all *modes* of the system can be excited by either the deterministic input u and/or the stochastic input w.

This model is the one most widely used linear system model by control engineers. Many industrial processes can be described very accurately by this type of model. However, not all of the terms are necessary, some of the terms, such as B or D, can be ignored in civil engineering applications to get a simpler reduced model.

If the unmeasured input and output noise processes w and v are identically zero, we obtain the deterministic subspace identification. The goal is to identify the A, B, C, and D system matrices of a deterministic state space model, by solving an input-output identification problem

If there is no measured external input u, we get purely stochastic subspace identification. Here the goal is to identify the A, C, Q, S, R from given output data only; this is an output-only identification problem, or the identification of modal parameters under ambient vibration.

However in stochastic identification, "output-only" does not means we know absolutely nothing about the system input; we need to make some assumptions, such as white noise ambient excitation, about the excitations to the system.

Another variant of subspace algorithms is the so-called combined deterministic – stochastic subspace identification algorithm, which identifies A; B; C; D; Q; R; S system matrices. This variant can solve an input-output identification problem with noises in both the input side and the output/measurement side.

Table 3.1: variants of subspace identifications (Overschee, 1997)

u is measured, w and v are zeros	Deterministic input-output identification
u is measured, w and v are not zeros	Combined deterministic-stochastic input-output identification
w is the sole input to the system	Stochastic output-only identification or ambient vibration identification

Subspace identification algorithms are based on concepts from system theory, control theory, numerical linear algebra and statistics. The most important achievement in the development of subspace identification is that the Kalman filter states (an estimation of the system states) can be obtained directly from input-output data using linear algebra tools (QR-decomposition and singular value decomposition) without knowing the mathematical model, while the original Kalman filter estimates system states from given system matrices.

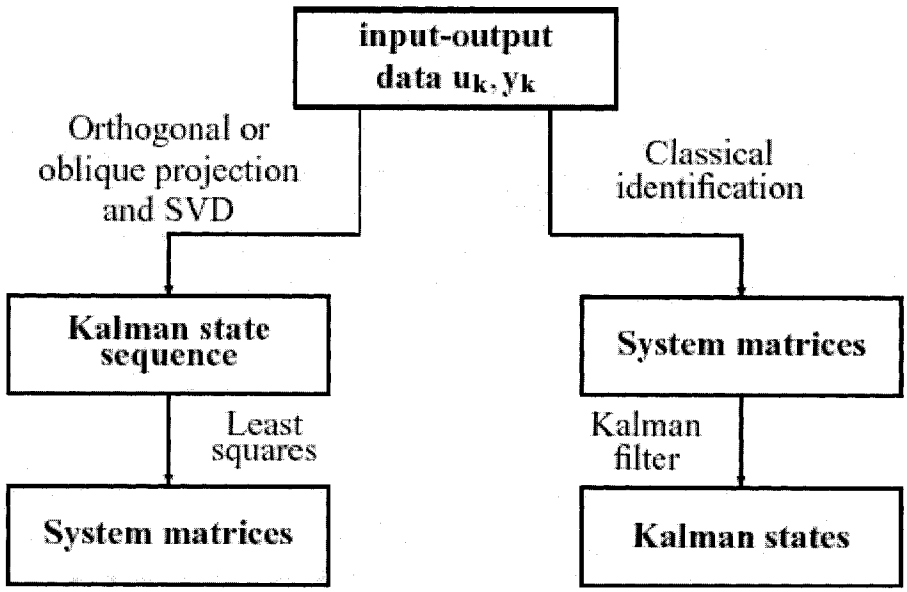


Figure 3.1: Comparing the subspace identification and classical control approach using Kalman filters. (Overschee, 1997)

Figure 3.1 shows the difference between subspace identification and classical control approaches using Kalman filter. Subspace identification aims at constructing state space

models from input-output data. The left hand side shows the subspace identification approach: first the (Kalman filter) states are estimated directly (either implicitly or explicitly) from input-output data or from output data only, then the system matrices can be obtained. The right hand side is the classical approach in control: first obtain the system matrices, and then estimate the states.

The state sequence estimates $\{ \hat{x}_i \}$ of a deterministic system can be found by computing the intersection of the past input and output and the future input and output spaces. For a stochastic model, the state sequence estimate can be obtained in two steps: first, the future output space is projected orthogonally into the past output space and next, singular value decomposition is carried out. The operation is performed through an orthogonal or oblique projection of the row spaces of certain block Hankel matrices constructed by measured data, or by covariance computed from data, into the row spaces of other block Hankel matrices, followed by singular value decomposition (SVD) to determine the order, the observability matrix and /or the state sequence.

Based on the how the projections are performed on the input and output row spaces, different algorithms have been proposed. The most popular ones are the N4SID (Numerical algorithms for Subspace State Space System Identification) algorithm, which performs the projection without any weight, and the CVA (Canonical Variate Analysis), MOESP (Multivariable Output-Error State sPace), which use some different weights in the projections.

Once these states are known, the identification problem becomes a linear least squares problem in the unknown system matrices.

$$\underbrace{\begin{pmatrix} x_{i+1} & x_{i+2} & \cdots & x_{i+j} \\ y_i & y_{i+1} & \cdots & y_{i+j-1} \end{pmatrix}}_{\text{known}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \underbrace{\begin{pmatrix} x_i & x_{i+1} & \cdots & x_{i+j-1} \\ u_i & u_{i+1} & \cdots & u_{i+j-1} \end{pmatrix}}_{\text{known}} + \begin{pmatrix} w_i & w_{i+1} & \cdots & w_{i+j-1} \\ v_i & v_{i+1} & \cdots & v_{i+j-1} \end{pmatrix}$$

State matrices estimates can be computed form the least-square problem.

$$\begin{pmatrix} \widehat{A} & \widehat{B} \\ \widehat{C} & \widehat{D} \end{pmatrix} = \min_{A,B,C,D} \left\| \begin{pmatrix} \hat{x}_{i+1} & \hat{x}_{i+2} & \cdots & \hat{x}_{i+j} \\ y_i & y_{i+1} & \cdots & y_{i+j-1} \end{pmatrix} - \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \hat{x}_i & \hat{x}_{i+1} & \cdots & \hat{x}_{i+j-1} \\ u_i & u_{i+1} & \cdots & u_{i+j-1} \end{pmatrix} \right\|^2$$

And the estimates of the noise covariance matrices follow from the least squares residuals.

The above procedures are implemented implicitly using several elegant numerical linear algebra tools, for example, the oblique projections can be done using QR algorithm.

Subspace methods for identification of linear multivariable dynamical systems have been very successful. The subspace system identification algorithms make full use of the well developed body of concepts and algorithms from numerical linear algebra. Numerical robustness is guaranteed because of the well-understood algorithms, such as the QR-decomposition, the singular value decomposition and its generalizations. Therefore, they are very well suited for large data sets and large scale systems. Moreover, subspace algorithms are not iterative. Hence, contrary to the PEM, there are no convergence problems. Unlike some other methods like PEM and polyreference, which either require the user choices on model orders, or need specific implementation for each different problem, one significant advantage of the subspace method is that it does not require much human interference in the identification process, i.e., it can be automated, which makes it very popular in control applications.

One of the disadvantages of subspace methods is the fact that it does not optimize a certain cost function. The reason for this is that, contrary to input-output models, we can not (as of this moment) formulate a likelihood function for the identification of the state space model, that also leads to an amenable optimization problem. So, in a certain sense, subspace identification algorithms provide (though often surprisingly good) "approximations" of the linear model.

3.5 Examples

3.5.1 Example 1: SDOF oscillator

In this example, we consider the identification of a single degree-of-freedom (SDOF) spring-mass system shown in figure 3.2 using simulated noisy transient acceleration response data. Here the excitation is a white-noise with spectral intensity S_0.

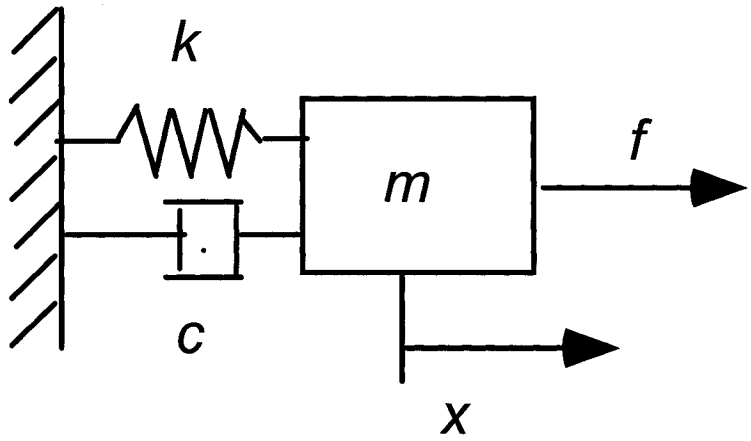


Figure 3.2: Single DOF spring mass system

The stiffness is 10 kN/cm, and damping ratio is 0.05, the mass is taken as 1 $kN \cdot s^2/cm$; for this SDOF system, the modal properties are equivalent to the stiffness and damping ratio.

The parameter values used to generate the simulated data are $S_0 = 1.0\ cm^2/sec^3$, and the noise variance $\sigma_v^2 = 0.251$(for the acceleration response). The chosen value of the noise variance corresponds to a 10 per cent root-mean-square (rms) prediction-error level, i.e., the noise is 10 per cent of the rms of the noise-free response. The time step used to generate the data is 0.01 sec, which is much smaller than the structure period 2 sec. However, a larger sampling time step was chosen ($\Delta t = 0.02$ sec) and the total time interval is T =50 sec, so that the number of data points is N =2500. This system is simulated using MATLAB programs.

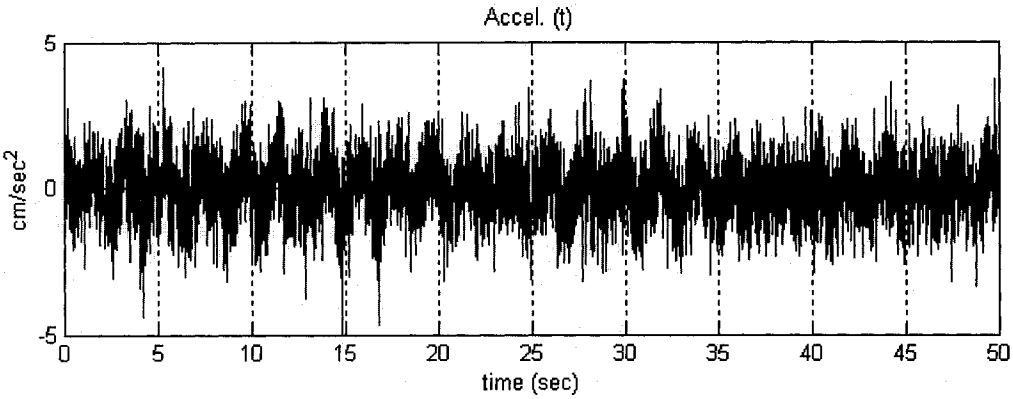


Figure 3.3: Simulated measurement of acceleration response

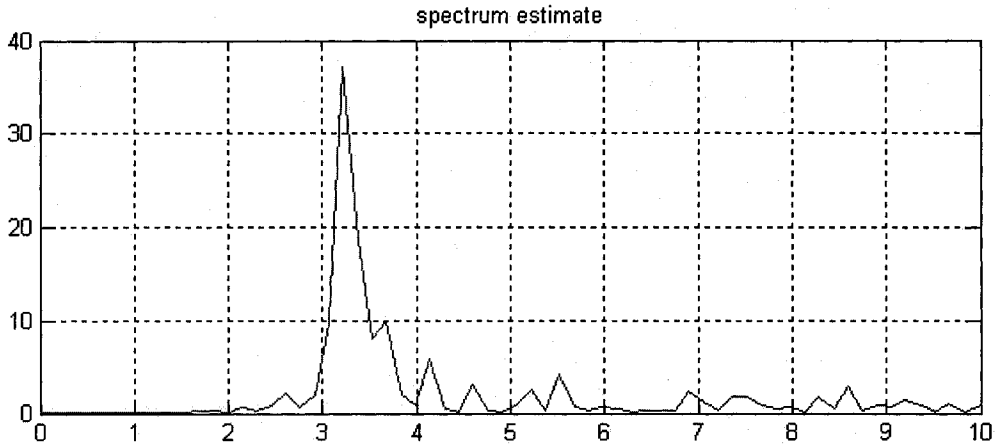


Figure 3.4: Periodogram spectrum estimate

Four spectrum estimation methods are used to estimate the spectrum plot from the acceleration output for the application of peak-picking method.

Table 3.2: Frequency estimated from PSD plot

Expected frequency (rad/sec)	periodogram	Modified periodogram with Hamming window	Bartlett	Welch
3.16	3.22	3.37	3.37	3.52

Interestingly, here the modified periodograms do not offer any benefits; this should

be because the measured time series length is short. For short series, the different average and smoothing techniques can not improve the quality, albeit an increase of the bias at some occasions.

3.5.2 Example 2: six-storey building

The second example refers to the six-story shear building (see Figure 3.5) subjected to a white-noise excitation at the base. The acceleration responses are measured at top floor and the third floor. The expected natural frequencies of this system are 6.78, 17.74, 26.96, 36.74, 45.15 and 54.99 rad/sec (computed using MATLAB). The simulation was conducted for 5% of all the modal damping ratios. The measurement error level is assumed to be 20% of the root-mean-square (rms) of the noise-free acceleration response at corresponding channels.

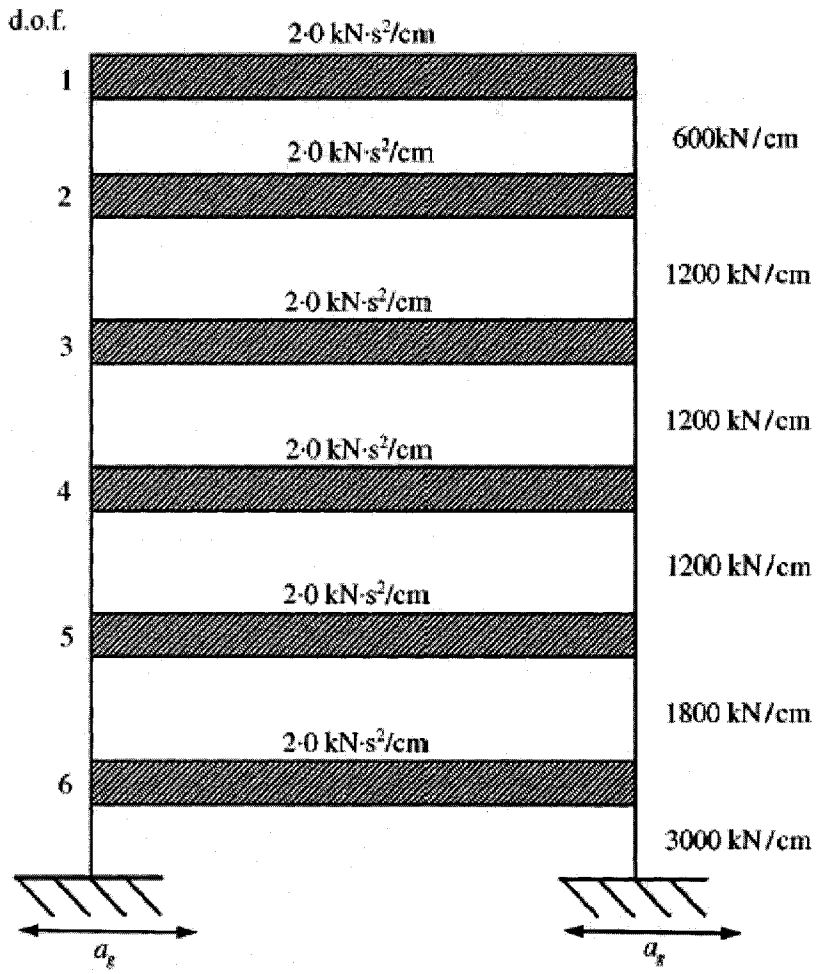


Figure 3.5: Six storey building model

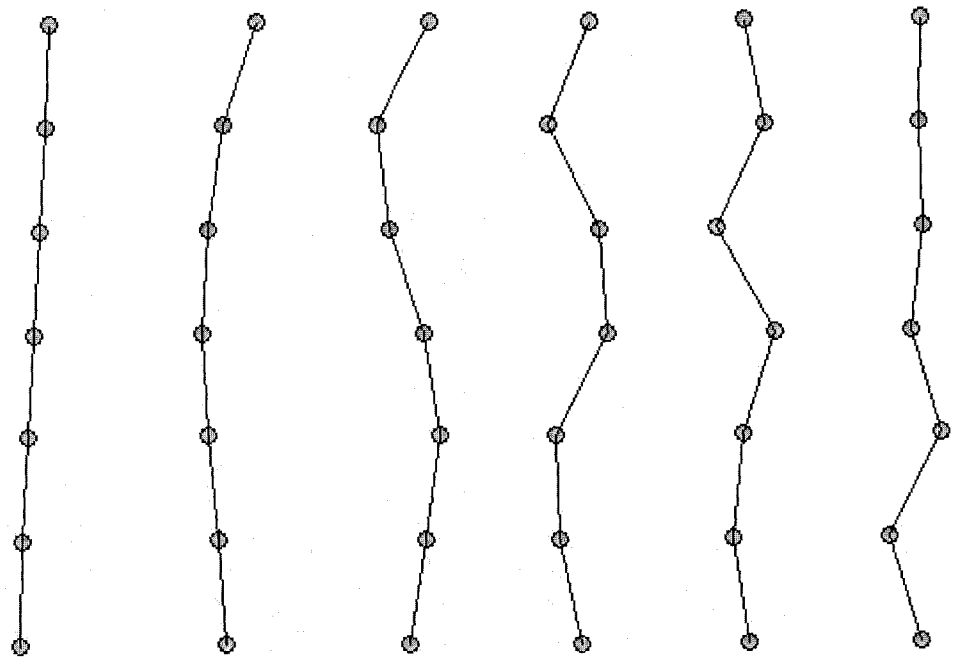


Figure 3.6: The six mode shapes (simulated in MATLAB)

The time step used to generate the data is 0.01 sec; however, the sampling time step was chosen (Δt =0.02 sec), much smaller than the structure's fundamental period 0.93 sec. The total time interval is t =50 sec, so that the number of measured data points is N =2500.

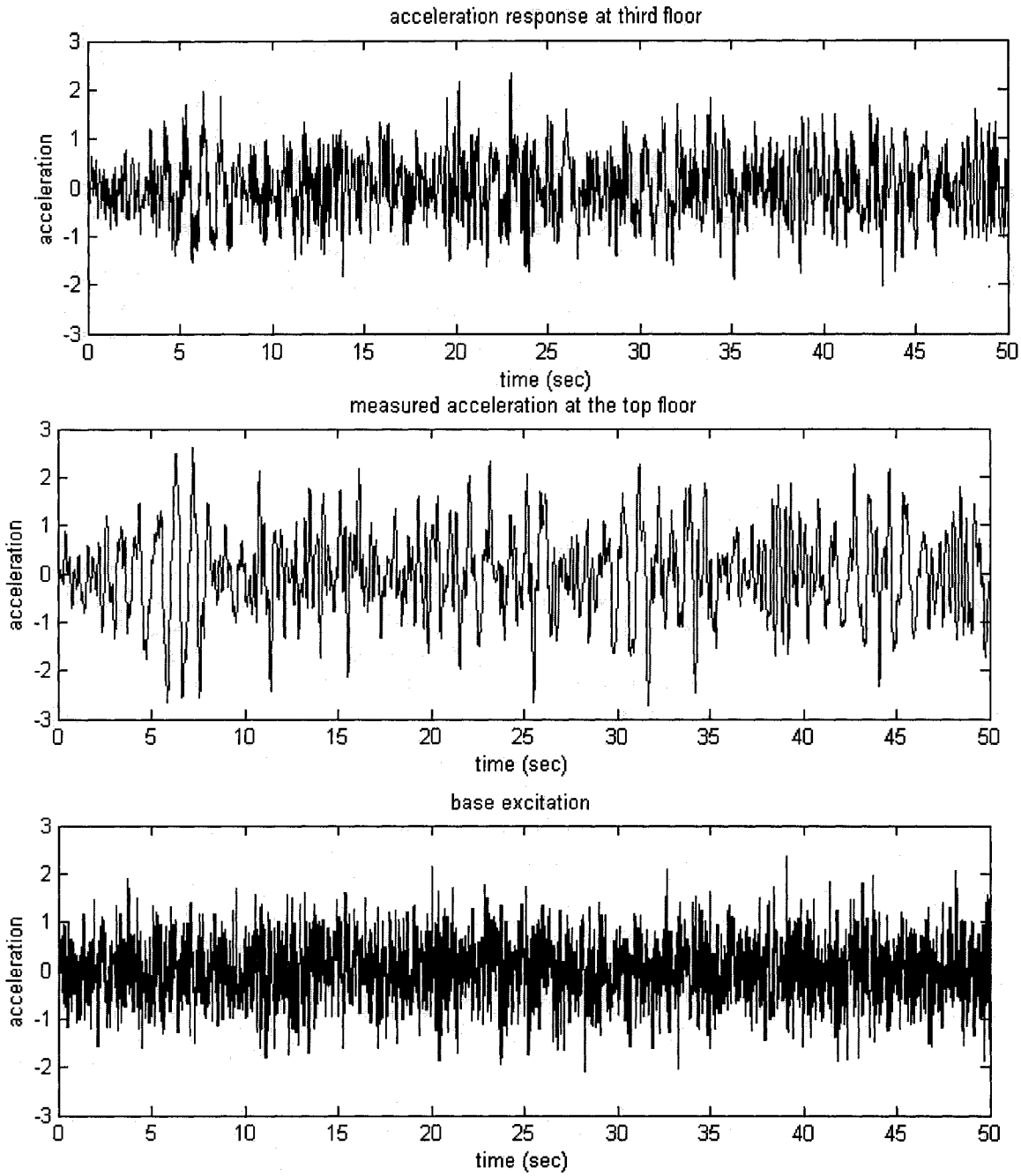


Figure 3.7: Measured acceleration response at the third floor, top floor and the measured base input

The peak-picking methods are tried first. The results are tabulated in table 3.3 and table 3.4. Although simple, these methods can give fairly good estimate for the frequencies. This is why they are not obsolete while so many advanced methods having been developed.

Table 3.3: Peak picking estimate from measurement at top floor

Expected frequencies (rad/sec)	periodogram	Hanning windowed periodogram	Bartlett method	Welch's method
6.78	6.90	6.90	7.21	6.90
17.74	18.41	18.41	18.1	17.33
26.96	26.38	25	25.92	25.46
36.74	30.37	29.91	29.91	N/A
45.15	N/A	N/A	41.88	N/A
54.99	N/A	N/A	51.85	N/A

Table 3.4: Peak picking estimate from measurement at third floor

True frequencies	periodogram	Hanning windowed periodogram	Bartlett method	Welch's method
6.78	6.75	7.36	7.52	3.66
17.74	17.79	17.64	18.41	17.95
26.96	27.30	27.77	27.00	24.08
36.74	35.90	35.44	36.66	N/A
45.15	46.48	45.56	46.91	N/A
54.99	N/A	52.15	53.05	N/A

The identified frequencies are good estimates, no damping is identified. It seems that the good results for higher modes of Bartlett method comes by a coincidence but not rule. The spectrum plot from measurement at third floor gives one more identified mode than from the measurement at top floor; this can be explained by viewing the mode shapes shown in Figure 3.6, the fifth mode is weak at the top floor, i.e., having a smaller relative amplitude, and is strong, i.e., having a relatively larger amplitude at the third floor.

Next we try the PEM method. The PEM function in MATLAB system identification toolbox is used for the identification. The PEM in the system identification toolbox adopts a Levenberg-Marquardt algorithm to solve the nonlinear optimization problem (Ljung, 2004).

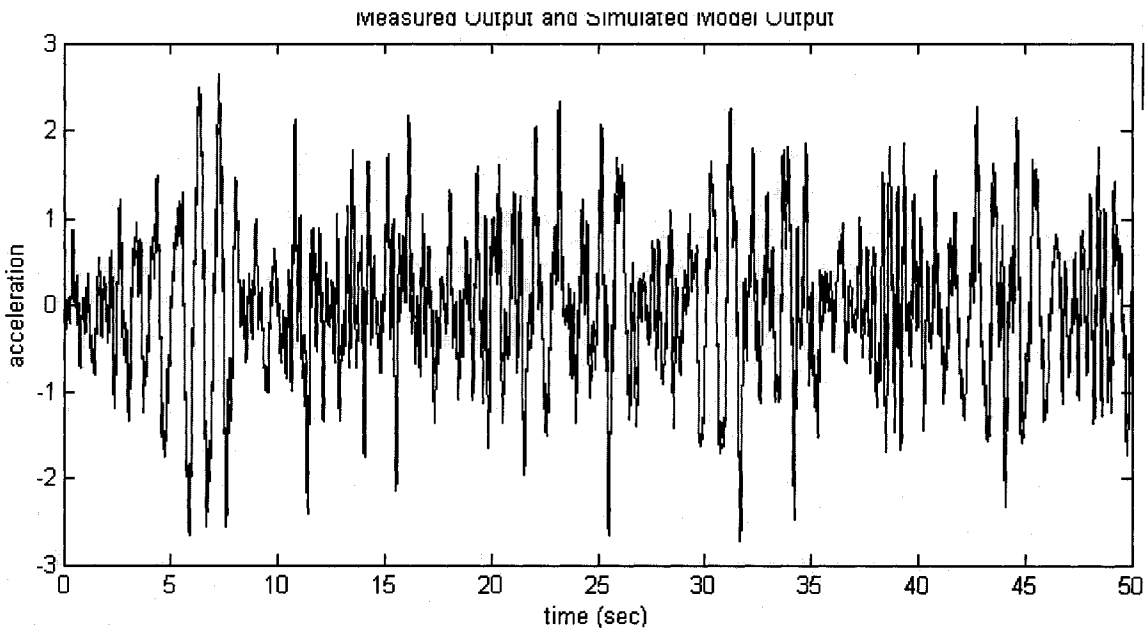


Figure 3.8: Output fitting of the PEM with measurement at the top floor

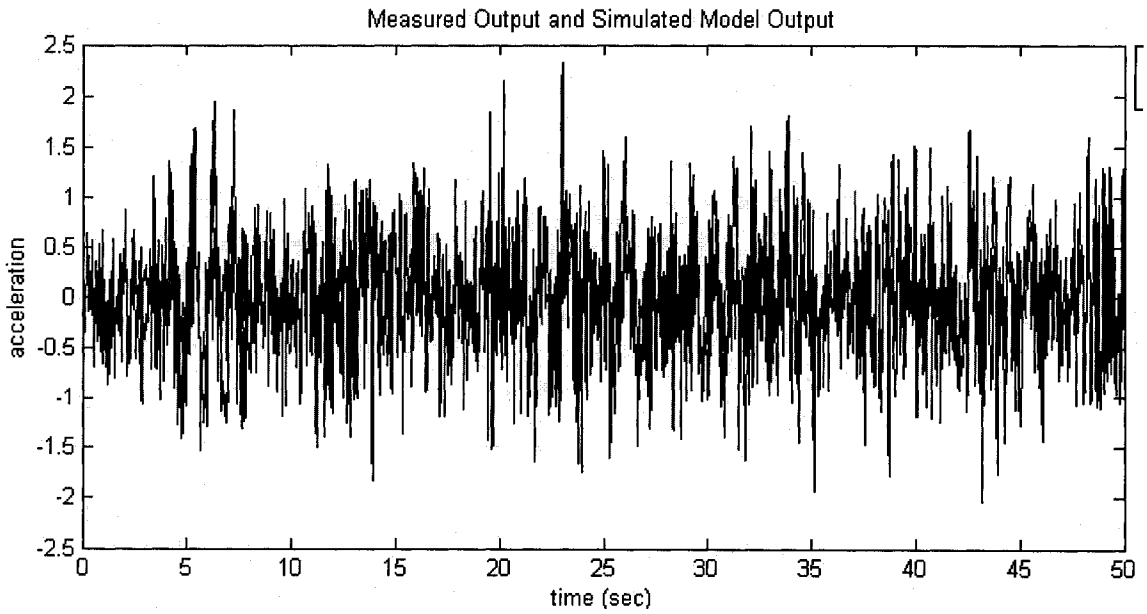


Figure 3.9: Output fitting of the PEM with measurement at the third floor

In using PEM, there is no direct method for finding the optimal model order, instead one must solve by trial and error. The model order chosen here is ARMAX(12,8,3), where 12 is the order for AR (AutoRegressive) polynomial, 8 is for X(eXogenous) polynomial and 3 is for MA (Moving Average) polynomial, respectively. This model choice gives a small value of FPE and OE; however, we are not going to say it is the best choice. Figure 3.8 illustrates the fitting between the identified model output and the measured output at the top floor, and Figure 3.8 shows the fitting at the third floor. It is observed that the output fitting is better at the top floor than at the third floor, thus indicating this choice of model order may be good for data measured at top floor.

Table 3.5: Identified modal parameters using PEM

Parameter	Reference value	Using measurement at third floor	Using measurement at top floor
ω_1	6.78	N/A	6.79
ω_2	17.74	12.32	18.41
ω_3	26.97	31.73	30.51
ω_4	36.74	31.80	31.368
ζ_1	0.05	0.077	0.048
ζ_2	0.05	0.061	0.060
ζ_3	0.05	0.058	0.077
ζ_4	0.05	0.063	0.065

Next, the N4SID function in MATLAB system identification toolbox is used here to illustrate the subspace identification. The identification from the measured 2500 points of data can not give any meaningful results, since subspace identification requires a Hankel matrix with output blocks, or a Hankel matrix with correlations calculated from output data, thus requires very long data for identification. Therefore, a 500 sec, 25000 points of measurement of acceleration response is simulated for the system identification using subspace method. In reality, it could be difficult to measure a structure's response for this long time unless in some controlled circumstances; this is one difficulty in applying the subspace methods. The identified frequencies are good, but there is a trend to overestimate the damping. This overestimation of damping is long been criticized as one weak point of subspace identification (Pridham, 2004).

Table 3.6: Identified modal parameters using N4SID subspace identification

Parameter	Reference value	Using measurement at third floor	Using measurement at top floor
ω_1	6.78	6.79	6.79
ω_2	17.74	18.41	18.41
ω_3	26.96	Miss	21.45
ω_4	36.74	33.55	33.41
ζ_1	0.05	0.048	0.048
ζ_2	0.05	0.060	0.060
ζ_3	0.05	Miss	0.116
ζ_4	0.05	0.170	0.080

3.6 Conclusions

The peak-picking, PEM and subspace identification methods are commonly considered as the simplest, the most accurate, and the most powerful tools in modal identification. The peak-picking is fast, but its accuracy is poor; and it seems no much improvement in its resolution and ability using "advanced" spectrum estimates. The PEM is the best tool for input-output measurements, due to the fact it minimized the OE in the least squares sense. The selection of ARMAX model order should be done with much care. The subspace identification requires long data to run, and it has a trend to overestimate the damping.

CHAPTER FOUR

DETERMINISTIC OPTIMIZATION OF MODEL OUTPUT IN TIME DOMAIN

4.1 General

This chapter presents a deterministic optimization based procedure to perform modal identification from given vibration data. The problem is formulated as a nonlinear least square fit problem to minimize the difference between measured and simulated outputs in the time domain. This cost function is simple and straightforward, but the resultant optimization problem is a nonconvex, nonlinear one, and no gradient is available in explicit form.

Following the exploitation/exploration paradigm, genetic and Nelder-Mead simplex algorithms are integrated to locate the optimum modal parameters in the search domain: when a promising area is detected by the Genetic Algorithm, the Nelder-Mead algorithm can be started with initial simplex in the localized domain to exploit this area and obtain the optimum. These examples demonstrate that the proposed procedure performs well, even the damping ratios can be accurately estimated which is a remarkable feature in the identification of civil engineering structures.

4.2 Formulation of the deterministic optimization problem

The equation of motion for a damped dynamic system using the finite element method can be formulated as:

(4.1) $$[M]\{\ddot{x}\}+[C]\{\dot{x}\}+[K]\{x\}=\{f(t)\}$$

Here [M], [C], [K] $\in R^{n\times n}$ denote the mass, damping and stiffness matrix, whereas n is the degree of freedom, {f(t)} denotes the load vector. The modal matrix $\Phi = \{\Phi_k\}$ (k=1...n) consists of the eigenvectors of the system. Due to the orthogonality property of the eigenvectors, the modal transformation using matrix Φ transforms the mass and the stiffness matrix into diagonal matrices. In most cases, the proportional damping is used,

and then the damping matrix C is also transformed into a diagonal matrix by modal transformation. Thus we obtain a system of uncoupled equations in the modal space.

$$\ddot{q}_k(t) + 2\zeta_k \omega_k \dot{q}_k(t) + \omega_k^2 q_k(t) = \frac{g_k(t)}{m_k}, \qquad \text{k=1,...,n} \tag{4.2}$$

where $\{g_k(t)\} = \Phi^T \{f(t)\}$ and $\{q(t)\} = \Phi\{x(t)\}$.

With the estimated modal parameters, we can calculate the response in the time domain using numerical integration. The objective function to be minimized is defined in the time domain as:

$$J(\theta) = \frac{1}{2}\sum_{j=1}^{M} (\ddot{x}(t_j, \theta) - \ddot{x}_j)^2 \tag{4.3}$$

Where θ is the vector of unknown modal parameters, including ω_k and ζ_k, if an element in the modal vectors is involved in the computation of the response, the modal vector element should also be included. t_j (j=1...M) is the sampling time, M is the number of observations, usually a large number, and θ is the vector of modal parameters to be identified. Vector $\{\ddot{x}_j\}$ is the measured accelerations and $\{\ddot{x}(t_j, \theta)\}$ are the calculated response using the estimated parameters; here we write $\{\ddot{x}(t_j, \theta)\}$ to indicate the explicit dependence on the modal parameters θ. The measured and the simulated responses are taken as accelerations, since the acceleration is the most accessible measured response. Accelerometers are the cheapest instruments and easier to manipulate.

The objective function expressed in (4.3) can be written as:

$$J(\theta) = \frac{1}{2}\sum_{j=1}^{M} (\ddot{x}(t_j, \theta) - \ddot{x}_j)^2 = \frac{1}{2}\sum_{j=1}^{M} \left\| r_i(\theta) \right\|_2^2 = \frac{1}{2} R(\theta)^T R(\theta) \tag{4.4}$$

The vector R = $\{r_1, \ldots r_M\}$ is called the residual. We see that:

$$\nabla J(\theta) = R'(\theta)^T R(\theta) \in R^M \tag{4.5}$$

The necessary conditions for optimality require that

$$\nabla J(\theta^*) = R'(\theta^*)^T R(\theta^*) = 0 \tag{4.6}$$

This objective function has been used by Dunn (1997) and Koh (2003) in their structural identification that the unknown parameters are the structural mass and stiffness, instead of modal parameters. Dunn (1997) used a modified Genetic Algorithm to identify a model of an aircraft structure with eight unknowns. Koh (2003) proposed a hybrid

computational strategy combining Genetic Algorithm with a compatible local search operator to identify a 3-DOF and a 10-DOF structural system in conditions including noise polluted input/output signals and no prior knowledge of mass, damping or stiffness of the systems.

The objective function can also be defined in the frequency-domain; the LSFD (Least-Square in Frequency Domain) method implemented in the "Structural dynamics toolbox for use with MATLAB" (Balmes, 1997) identifies the modal model properties through fitting the frequency-response-function (FRF). Bernstein and Richter (Bernstein, 2003) used the RSS between the measured and computed FRF as the objective function in their genetic algorithm optimization of finite element model identification. In this work we will consider a formulation in the time domain.

There are two major difficulties associated with this type of formulation: no direct derivative information is available, therefore we must have a code which evaluates the cost function only; and that there can be noise in the evaluation of the cost function. Therefore solving the above-stated optimization problem is not a trivial task; nonetheless this type of optimization problem is not rare in engineering. As an example is the optimization of blade shape which requires the computation from CFD (Computational Fluid Dynamics) software to form the objective function for optimization of multiple parameters (Berghen 20003, Berghen 2004).

In the next section we will review some general optimization techniques including global optimization approaches, derivative-based methods and direct search methods. The Nelder-Mead and Genetic Algorithm, which are chosen to in our approach will be described in detail.

4.3 Selection of optimization algorithms

4.3.1 derivative-based methods

The iterative optimization procedures solving the nonlinear least-square problem fall into two general classes, the derivative-based methods and the derivative-free direct search methods (Chong, 1996).

Most of the former approaches belong to the quasi-Newton hill-climbing and trust-

region family and are best for unimodal functions. They are usually fast convergent algorithms, but they are not suitable for our problem because of the fact they need the gradient or Hessian information in determination of search directions.

4.3.2 Direct search methods

Direct search methods do not require the gradient or Hessian to be computed, they use values of the objective function from a set of sample points and use that information to continue the sampling and search for the optimum. These kinds of methods were first suggested in the 1950s, and continuously improved in the 1960s ad 1970s (Kelley, 1999). Initially, these methods were typically motivated by low-dimensional geometric intuition rather than mathematical theory. There are two important classes of direct search methods: the geometry-based and the model-based methods.

1) Geometry based methods:

In geometry-based methods, the function values are used to create and maintain a geometric figure, most commonly a simplex. These methods make minimal assumptions about the cost function and do not create any mathematical model for it.

Examples of geometry-based algorithms are the coordinate search method with fixed step sizes (Davidon, 1991), the pattern search method based on automata theory (Hooke, 1961), the multi-directional Parallel Direct Search (PDS) method of Dennis and Torczon (Dennis, 1991). Usually, these methods need only the function values.

A simplex-based method constructs an evolving pattern of n + 1 points in $\Re^n$ that are viewed as the vertices of a simplex. A new simplex is formed at each iteration by reflecting the vertex with the largest value of objective function, over the centre of the opposite face of the simplex, or by contracting toward the vertex with the smallest value of objective function. Among all the geometry based methods adopting the simplex concept, the Nelder-Mead method is the most popular one; its robustness has been tested in many applications. We will discuss the Nelder-Mead simplex method in detail in section 4.4.

2) Model Based Methods

Model-based methods use the function values to build a convenient model, such as a quadratic function, through interpolation or approximation. The underlying assumption in defining a model is that the objective function J is, in some sense, "nice". Recent survey of such methods is given in (Conn, 1997). One successful application of this kind of method is the CONDOR algorithm (COnstrained, Non-linear, Direct, parallel Optimization using trust Region method for high-computing load function), which was developed for a CFD related optimization problem (Berghen, 2004).

The main difficulty in practice to use this method is that one must retain the interpolation point set with certain geometric properties at each iteration. Different interpolation functions, mostly polynomials have been used in many development of this kind of methods.

There are two essential ingredients of derivative-free methods. The first is to pick better points. In geometry-based methods, the algorithm should be designed to exploit the next place to sample. In model-based methods, the expectation is that the minimum of the surrogate model will predict suitable points. The second important ingredient is to determine an appropriate search subspace. Different ways to determine such search subspaces result in different algorithms.

4.3.3 Global search methods to avoid local minimum

Another difficulty that arises from the optimization approach defined in this chapter is the problem of local minima. Either the geometric or the model-based methods are prone to trap into a local minimum and not able to escape. Different strategies have been suggested for this problem, the most famous are the Simulated Annealing and continuation approach.

1) Simulated Annealing

As its name implies, the Simulated Annealing (SA) exploits an analogy between the way in which a metal cools and freezes into a minimum energy crystalline structure (the annealing process) and the search for a minimum in a general system. If a physical system is melted and then cooled slowly, the entire system can be made to produce the most

stable (crystalline) arrangement, and not get trapped in a local minimum.

SA's major advantage over other methods is its ability to avoid being trapped at a local minimum. The algorithm employs a random search, which not only accepts changes that decrease the objective function f, but also some changes that increase it. The latter are accepted with a probability given by:

$$p = \exp\left(-\frac{\Delta J}{T}\right) \tag{4.7}$$

where ΔJ is the increase in objective function J, T is a control parameter, which by analogy, is known as the system "temperature".

Let us describe briefly how SA works. Simulated Annealing starts at a high artificial temperature; while cooling the temperature slowly, it repeatedly chooses a subset of the variables and changes them randomly in a certain neighbourhood of the current point. If the objective function has a lower function value at the new iterate, the new values are chosen to be the initial values for the next iteration. If the function value is higher, the new values are chosen to be the initial values for the next iteration with a certain probability, depending on the change in the value of the objective function and the temperature. The higher is the temperature and the lower is the change, the more probable the new values are chosen to be the initial variables for the next iteration. Throughout this process, the temperature is decreased gradually, until eventually the values do not change anymore. Then, the function is presumably at its global minimum. Since we can always choose a higher temperature to start, the temperature is never increased.

2) Continuation Approach

Continuation Approach is another useful strategy in searching of a global minimal. In this approach, the original function is gradually transformed into a smoother function with fewer local minima. This is illustrated in Figure 4.1, for a one-dimensional non-convex objective function.

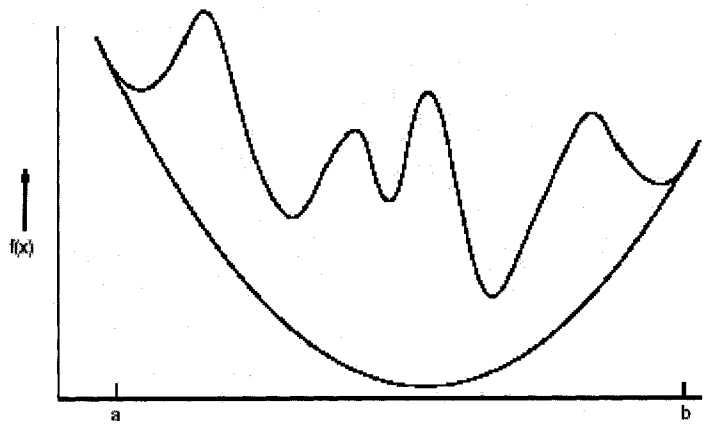


Figure 4.1: Continuation approach using transformation

Consider for example the problem:

$$\min J(x)$$

where $J(x) : \Re^n \rightarrow \Re$ is a non-convex objective function with many local minima. The problem is transformed into

$$\min h(x) = J(x) + \mu\, q(x)$$

where q(x) is strictly convex, and μ is a positive constant. Thus h(x) is a convex function for a sufficiently large value of μ. Let $x(\mu) = \arg\min_x \{h(x)\}$, be the minimizer of h(x) for a fixed value of μ, then x approaches $x^0 = \text{argmin}\ \{J(x)\}$ as $\mu \rightarrow 0$. Here, the objective is transformed to be convex by adding a strictly convex term $\mu\, q(x)$, and then gradually changed back to the original when $\mu \rightarrow 0$. This is not simulated annealing, but μ might be considered as a similar control parameter as the temperature T in the simulated annealing. Though there are many reports on successful application the continuation approach, it cannot succeed for any arbitrary function. The type of local minima and noise in the objective function are different for different applications and therefore the success is not always guaranteed.

3) Genetic algorithm

Besides those local search techniques, one very powerful derivative-free direct search method is the Genetic Algorithm. Genetic algorithms are optimization techniques derived from the principles of evolutionary theory. They contain a population of

individuals; each of them has a known fitness. The population is evolved through successive generations until a stopping criterion is satisfied. A genetic algorithm represents points in the search space by a vector of discrete (typically) bit values. A new child is produced by crossover and mutation parts of the bit vector from its parent. This is analogous to the way that chromosomes of DNA (which contains the inherited genetic material) are passed to children in nature.

Genetic algorithm (GA) is one of the most widely applied meta-heuristic search methods that have the power to well explore the whole feasible search domain. It had been applied to Finite Element (FE) update (Jaishi, 2005) and FE structural identifications (Dunn, 1997); however, the convergence becomes slow while approaching the solution.

The GA is almost a panacea to solve any direct optimization problem, if coded in the correct way and suitable control parameters are chosen. However, it is notoriously slow in convergence; the power of Genetic Algorithm is usually associated with parallel processing..

4.3.4 Genetic and Nelder-Mead

We propose in our work to combine the Nelder-Mead simplex method and the Genetic Algorithm for the solution of an optimization problem minimizing the objective function (4.3). The Nelder-Mead direct simplex method is adopted to search the local minimum because it is a derivative-free direct method. However, examples shows that the Nelder-Mead usually converges to local minima that are not necessarily the true solution to the nonlinear least square problem; to overcome this difficulty, a global method, Genetic Algorithm, is adopted to explore the whole feasible search domain and localize a promising area likely to contain a global minimum. When a promising area is detected, the Nelder-Mead algorithm can be started with initial simplex in the localized domain to exploit this area and obtain the optimum as accurately and quickly as possible.

The idea of hybridizing global search method and local search method is not new. Various combinations of Genetic Algorithm and some hill-climbing algorithm have been proposed in the literature. Chelouah et al. proposed a genetic and Nelder-Mead hybridized algorithm as one ideal pair for global optimization of many difficult continuous multi-

minima functions (Chelouah, 2003). A hybrid GA and Least-Square algorithm to identify structural parameters has been proposed by Koh et al. (Koh, 2003).

It is noted here that we are not assuming that they are the best choice. The best optimization algorithm for our modal identification problem, if exist, should be one model-based method that use some interpolation function close to the point of local minimum; they are faster than simplex methods, and usually are more powerful in multi-variate optimization. If hybridized with the continuation approach, this method can localize the global minimum. However, to apply this method, the choice for interpolation functions and the decision for updating direction are usually problem-specific; i.e., the model-based optimizer developed for a CFD optimization problem may not work well for modal identification problem. One must develop such an algorithm specifically for the modal identification problem. The development of such an algorithm requires testing of different interpolation functions and search directions as they are usually problem-dependent in model-based method. Since the purpose of this thesis is to demonstrate the usefulness of optimization approach for modal identification, we would like to leave the development of good algorithms to future research.

4.4 Nelder-Mead Simplex Search

The Nelder-Mead simplex search method is possibly the most famous geometry-based direct search method. It requires only function values. It is also applicable for non-smooth problems, where the function is not given explicitly.

A "simplex" is a geometrical figure consisting of (n+1) points $\{x^0, x^1, \ldots, x^n\}$, each in $\Re^n$, i.e. the triangle for n = 2 or tetrahedron for n = 3. The shape of a simplex is convex with positive volume.

The worst vertex is also called the highest vertex that corresponds to the highest objective function value, with an analogous definition of the best point or lowest point. Through a sequence of elementary geometric transformations, the initial simplex moves, expands or contracts. The method uses only the values of the objective function at the considered points.

After each transformation, the current worst point is replaced by a better one. The basic operations on the simplex are:

Reflection: $x^r = (1+\alpha)\bar{x} - \alpha x^n$

Expansion: $x^e = (1-\gamma)\bar{x} + \gamma x^r$

Contraction: $x^c = (1-\beta)\bar{x} + \beta x^n$

where $\bar{x} = \frac{1}{n}\sum_{j=0}^{n-1} x^j$, and α , β , γ are constants: $0 < \alpha < 1$ reflection factor; $0 < \beta < 1$ contraction factor; $1 < \gamma$ expansion factor.

At the beginning the highest vertex which has the maximum objective function value is replaced by another point image of the highest vertex. This operation is the reflection. If the reflected point is better than all the other points, the method expands the simplex in this direction; otherwise, if it is at least better than the worst, the reflection is performed again with the new highest point. While the worst point is at least as good as the reflected point, the contraction is performed. If the worst point is better than the contracted point, the multi-contraction is performed. At each step, the new point should be checked not outside the feasible solution space. Through these operations the simplex finds and adapts its way on the function "landscape", and finally surrounds the optimum.

One Nelder-Mead iteration has two possible outcomes: (1) a single new point replaces the worst vertex; or (2) if a shrink is performed, the new simplex contains the best point from the previous iteration and n new points closer to the best point than the previous ones. A typical iteration of the Nelder-Mead algorithm is outlined by Kelley (Kelley, 1999) as follows:

1. **Initialization**: Let a starting point x be given.
   Set parameters $0.5 \le \alpha \le 1$, $1 < \gamma$, $0 < \beta < 1$, and $0.25 \le \sigma < 1$ as factors for reflection, expansion, contraction and shrink, respectively.
2. **Ordering**: Order the n + 1 simplex vertices to satisfy $f(x^1) \leqslant f(x^2) \leqslant \ldots \leqslant f(x^{n+1})$, using a consistent tie-breaking rule for the equality cases.
3. **Reflection**: Compute a reflection point x^r from

   $$x^r = \bar{x} + \alpha(\bar{x} - x^{n-1}),$$

   where $\bar{x}$ is the centroid of the n best vertices (all except x^{n+1}), i.e.,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x^i$$

and α is the reflection factor.

Evaluate $f_r = f(x^r)$.

If $f_1 \leqslant f_r < f_n$, accept the reflected point x_r, remove x_{n+1}, terminate the iteration and goto Step 2.

4. **Expansion**: If $f_r < f_1$, calculate an expansion point x^e from

$$x^e = \bar{x} + \gamma(x^r - \bar{x})$$

where γ is the expansion factor.

Evaluate $f_e = f(x^e)$. If $f_e < f_r$, accept x^e, remove x^{n+1}, terminate the iteration and goto step 2.

Otherwise (if $f_e \geqslant f_r$), accept x^r, remove x^{n+1}, terminate the iteration and goto Step 2.

5. **Contraction**: If $f_r \geqslant f_n$, perform a contraction between $\bar{x}$ and the better one in x^{n+1} and x^r. Here the program has two branches.

   a) Outside Contraction: If $f_n \leqslant f_r < f_{n+1}$ (i.e., x^r is strictly better than x^{n+1}), perform an outside contraction. Calculate

$$x^{oc} = \bar{x} + \beta(x^r - \bar{x})$$

   where β is the contraction factor, and evaluate $f_{oc} = f(x^{oc})$.

   If $f_{oc} \leqslant f_r$, accept x^{oc}, remove x^{n+1}, terminate the iteration and goto Step 2. Otherwise, goto step 6 (perform a shrink).

   b) Inside Contraction: If $f_r \geqslant f_{n+1}$ (i.e., x^{n+1} is better than x^r), perform an inside contraction. Calculate

$$x^{ic} = \bar{x} - \beta(\bar{x} - x^{n+1})$$

   where β is the contraction factor, and evaluate $f_{ic} = f(x^{ic})$.

   If $f_{ic} < f_{n+1}$, accept x^{ic}, remove x^{n+1}, terminate the iteration and goto step 2. Otherwise goto step 6 (perform a shrink).

6. **Shrink**: Define n new vertices by

$$x^i = x^1 + \sigma(x^i - x^1),\ i = 2, \ldots, n+1,$$

where σ is the shrink factor, and evaluate function f at these points.

If stopping criteria is satisfied, then terminate the algorithm. Otherwise, terminate the iteration and goto Step 2.

The formulation implemented in this thesis allows two criteria for termination: the first is checking the tolerance of simplex diameter, which is defined as f_{n+1}-f_1 or $f(x^{n+1})$-$f(x^1)$, is sufficiently small or not; the second is whether a user-specified number of function evaluations has been expended.

Unlike the Newton family of search methods, the development of Nelder-Mead simplex method is based on geometrical intuitions, rather than mathematical deductions. Despite major efforts, only very weak convergence results (Kelley, 1999) have been established by mathematicians, and only for the original Nelder-Mead method in one and two dimensions (Wright, 1995).

4.5 Examples using the Nelder-Mead method:

4.5.1 SDOF example

As a first example, we consider the single degree-of-freedom (SDOF) spring-mass system shown in figure 3.2. A fourth-order Runge-Kutta method is used to integrate the system's response under an impulsive load. The simulated system is treated as the "true system", and the simulated acceleration outputs are taken as the measurements. Noise/disturbance is added to the simulated outputs to mimic the measurement noise which can not be avoided in real engineering practice. The exact value of stiffness to the "true system" is assumed 10 kN/cm, and damping ratio is 0.05, the mass is taken as 1 $kN \cdot s^2/cm$; thus for this SDOF system, the modal properties are equivalent to the stiffness and damping ratio.

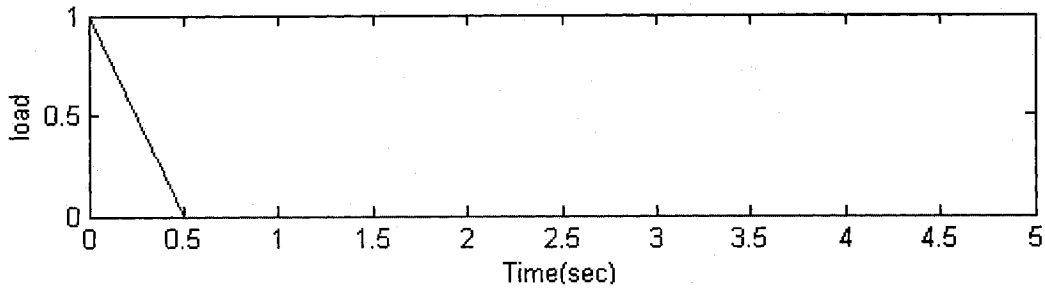


Figure 4.2: The loading applied to the structure

In this example, there are only two parameters to be estimated, the stiffness and the damping ratio. Then each simplex consists of three vertices, each is a pair of stiffness and damping ratio. The iteration of Nelder-Mead search terminates when the diameter of the simplex is smaller than the tolerance. While the tolerance for simplex diameter is set to be 0.001, we obtain the following experimental results:

Table 4.1: Identified parameters with different initial guess and noise level in the measurement (the reference parameters are (10, 0.05))

Case	Initial guess		Identified parameters		Noise level
	k	ζ	k	ζ	
1	1, 2, 3	0.04, 0.06, 0.07	fail	fail	50%
2	1, 2, 3	0.04, 0.06, 0.07	0.05	0.03	0 %
3	10, 2, 3	0.04, 0.06, 0.07	9.9767	0.053	50%
4	8, 9, 11	0.04, 0.06, 0.07	9.9752	0.053	50%
5	8, 9, 11	0.04, 0.06, 0.07	10.00	0.05	0 %
6	8, 9, 11	0.04, 0.06, 0.07	10.00	0.049	100%
7	8, 9, 11	0.04, 0.06, 0.07	9.87	0.061	200%
8	8, 9, 11	0.04, 0.06, 0.07	9.70	0.082	300%
9	8, 9, 11	0.0, 0.01, 0.02	10.01	0.05	0%
10	8, 9, 11	0.0, 0.01, 0.02	9.97	0.07	50%

In table 4.1 the last column, level of noise is the percent of noise variance against the rms of the noise-free response, i.e., a 50% noise level corresponds to noise with a

variance that is 50% prediction-error level.

In the first and second row of the table, we see that the identification results are poor when the initial guesses for the stiffness are not close to the true value. The 4^{th} and 5^{th} row show that the identification procedure leads to good results with good initial guess close to the reference solution. In the 6^{th} case, although the noise level is high, the obtained results are very close to reference values of the stiffness and damping ratio. In the 7^{th} and 8^{th} cases, the deteriorating effect of high level of noise on the identification of damping is clearly demonstrated, and we see that the identification of damping is more sensitive to the noise than the identification of the stiffness. However, this high level of noise, 200% (case 7) and 300% (case 8), will definitely make all the conventional system identification methods fail, but this deterministic optimization approach can still give some reasonable, though not accurate, results. In the last two cases, with good initial guess for stiffness but poor initial guess for damping, results demonstrate the additive effect of poor initial guess and noise; without noise, the optimizer is able to find the optimal even with a poor initial guess of damping, but if there is a relatively high level of noise, the damping identification becomes poor. In theory of direct search optimization, this is called the non-smoothness effect: that the noise in the data adds some non-smoothness to the objective function, and makes the optimizer difficult to find the valley.

The following observations can be drawn from the shown results:

1) The initial guess is crucial in the performance of the nonlinear LS identification.
2) With a good initial guess, this method performs well, even under a high level of noise.
3) Damping is more susceptible to noise. If the noise level is high, the estimation of damping is poor, but the stiffness estimation can be acceptable.
4) Good initial guess of damping is more crucial for high level of noise.
5) The proposed method, if successful, is surprisingly robust against measurement noise.

In the next step we check whether the biased estimations are due to trapped local minimum or poor convergence of global minimum. We use this initial simplex: 10 2 3 for

stiffness and 0.04 0.06 0.07 for damping, and test the performance of our program using different tolerance for Nelder-Mead search algorithm.

Table 4.2: Identified parameters with different tolerance of the Nelder-Mead search algorithm

tolerance	Identified parameters		Noise level
	k	ζ	
0.1	9.76	0.07	50%
0.001	9.834	0.063	50%
0.00001	9.77	0.067	50%

We observe that for poor initial guess, the reduction of tolerance does not guarantee improvement of the solution. This clearly indicates the bias in estimation is attributed to the trap into a local minimum, not unconverged solution.

Figure 4.3 shows the convergence history of the 4th case. Since the objective function defined is the RSS of the output error, the function value should approach zero after convergence. The simplex diameter, which is one criterion for terminating program, should also be close to zero. The simplex gradient is computed as the ratio of the maximum Euclidean distance between simplex vertices and the simplex diameter $f(x^{n+1})$-$f(x^1)$; its norm is also a measure of the performance of the Nelder-Mead algorithm, that a large number of the norm indicates possibly there is stagnation at a non-optimal point (Kelley, 1999). This stagnation is possible to happen if there is a flat region of the objective function.

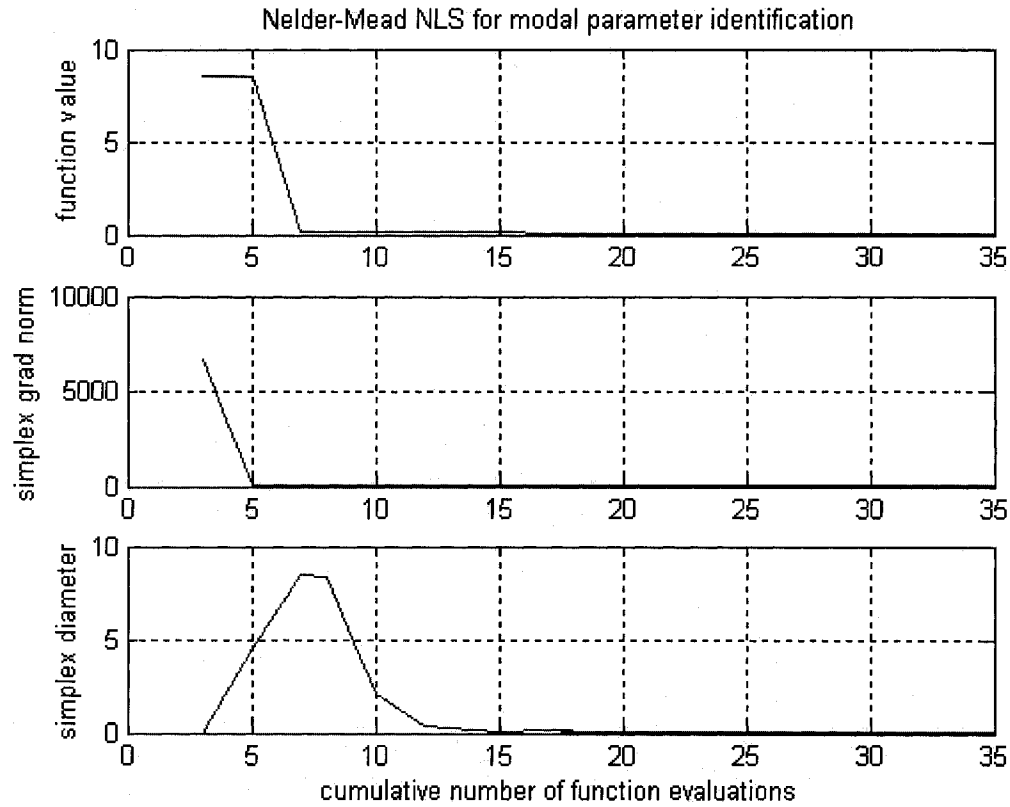


Figure 4.3: The convergence history of simplices

4.5.2 MDOF example

The second example refers to the six-story building shown in Figure 3.5. This structure is excited at the top floor with Gaussian white noise, and the acceleration responses are measured at the same floor. The damping ratios are assumed to be 5% for all the modes. The measurement error level is assumed to be 20% of the root-mean-square (rms) of the noise-free acceleration response at corresponding channels. Sampling rate of the data is 50Hz while the system is simulated at a time step of 0.01sec. A measurement of 5 sec duration is used for the identification.

As a first step, we try to identify only the lowest mode, with a poor initial guess {50, 51, 52} for a frequency value 6.7767, and {0.04, 0.05, 0.06} for damping ratio of 0.05. The search converges to a local minimum (Figure 4.4), and results in bad estimates.

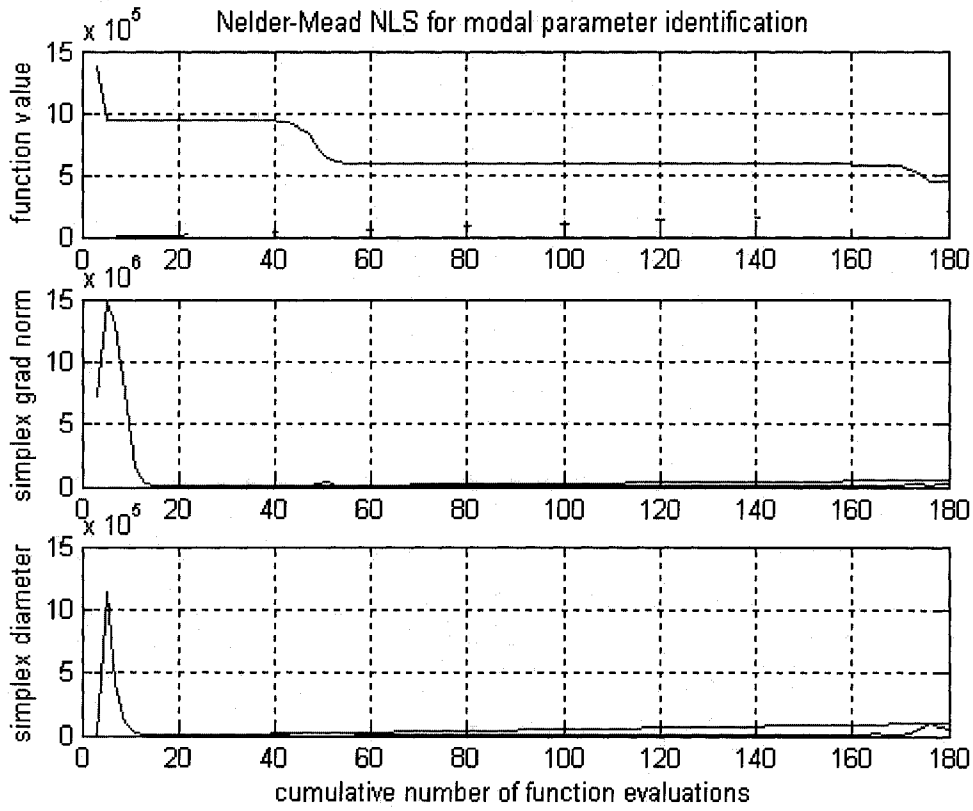


Figure 4.4: the convergence of simplices to a local minimum

Starting from a good initial guess: {5, 6, 7} for the frequency and {0.03 0.05 0.06} for damping, the search converges to the global minimum (Figure 4.5), and the results are good estimates.

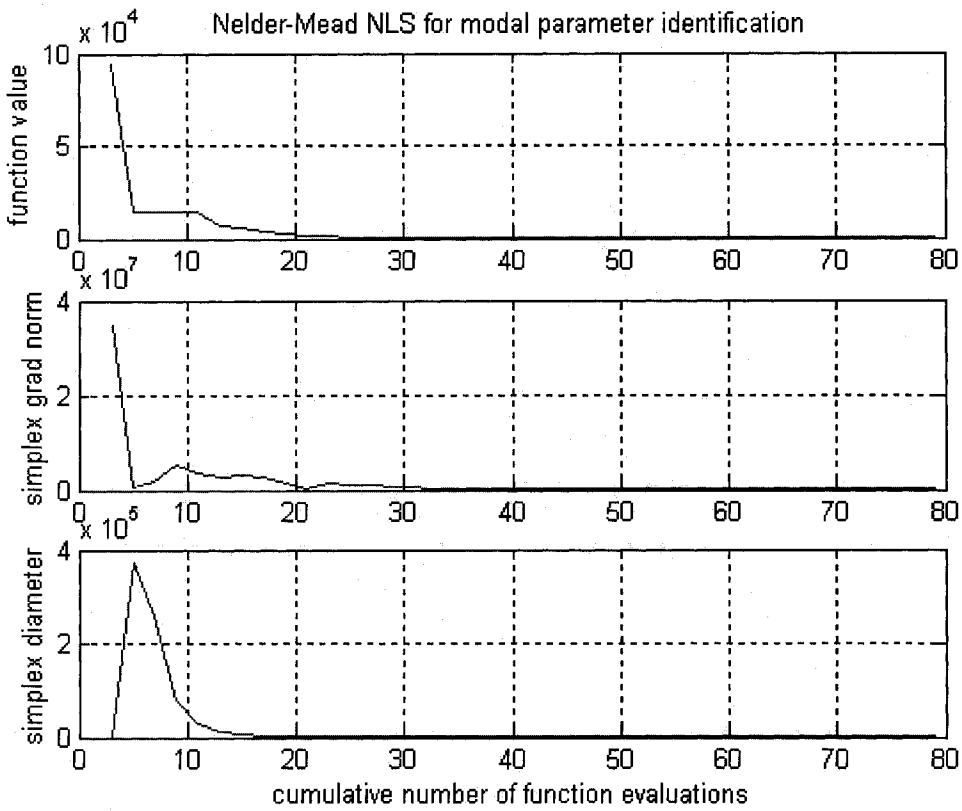


Figure 4.5: The convergence of simplices to the global minimum

Table 4.3: Identified parameters of the first mode with different initial guess (The exact parameters are (6.78, 0.05))

Initial guess		Identified parameters		Noise level
ω	ζ	ω	ζ	
50, 51, 52	0.04, 0.05, 0.06	237.35	0.807	20%
5, 6, 7	0.04, 0.05, 0.06	6.78	0.050	20%

We can also identify several modes at the same time by using a larger simplex. While providing good initial guesses:

Table 4.4: Initial guess for the identification of first three modal frequencies and damping

mode	Reference value		Initial guess (simplex centroid)	
	ω	ζ	ω	ζ
1	6.78	0.05	6	0.06
2	17.74	0.05	15	0.06
3	26.96	0.05	25	0.04

Here there are six parameters to be identified, each vertices of the simplex has a coordinate of six variables, and the simplex has seven vertices in total. The number of vertices of a simplex is the number of parameters plus one. On a two dimensional plane coordinating two parameters, the simplex is a triangle with three vertices; in the three dimensional space with three parameters as coordinates, the simplex is a tetrahedron with four vertices, and so on for higher dimensions.

We can generate the initial simplex coordinates manually if the number of parameters is small, or the dimension is low; however this manual generation would become too tedious if the numbers of parameters become large, as in this case we need to generate seven vertices for a simplex, and each vertices contains six number of coordinate values. Therefore, we adopt an automatic generation of initial simplex for this example (Kelley, 1999). This method generates a unit simplex of dimension n around the origin and then shifts its controid to the initial guess. This way guarantees that the simplex has a good geometry.

The converged simplex vertices are given in the following table:

Table 4.5: Identified parameters of first two modes with good initial guess

Name of parameter	Reference value	Identified value
1st natural frequency	6.78	6.77
2nd natural frequency	17.74	18.12
3rd natural frequency	26.96	27.14
1st mode damping ratio	0.05	0.050
2nd mode damping ratio	0.05	0.058
3rd mode damping ratio	0.05	0.058

A good agreement with the exact solution is obtained. Again, this proves that as long as the starting guess is good enough to the final results, the Nelder-Mead algorithm works well for our modal identification problem.

4.6 The Genetic Algorithm and its implementation

4.6.1 Basic procedures of the Genetic Algorithm (GA)

Originally proposed by Holland (Holland, 1975), Genetic Algorithm (GA) is a general purpose stochastic optimization method that mimics the heuristic concepts of natural selection and genetic operation. It belongs to the class of probabilistic algorithms, yet it is very different from random algorithms as it combines elements of directed and stochastic search. This approach has been shown to provide robust search in complex spaces (Michalewicz, 1999).

One important property of GA is that they maintain a population of potential solutions; thus it is efficient to explore a wide search space and detect a promising "valley". Besides, it is a self-start method, no need for an initial guess. Despite these advantages, GA's slow convergence in fine-tuning is one of the major known drawbacks. The slow convergence of GA before providing an accurate solution is closely related to its failure to exploit local information.

In applying the Genetic Algorithm, the modal parameters to be optimized must be coded as chromosomes (genes) like a design vector. The initialization of all chromosomes in the population is done randomly. The bit-streams are then translated into floats and the corresponding fitness values of the chromosomes are calculated. The rank of every chromosome is determined according to its fitness. The stopping criterion is checked and the algorithm turns to the selection step. Thereby, a fixed percentage of the chromosomes from the population according to their rank are selected to be a part of the set of parents of the next generation. Another part of the chromosomes (those with medium-grade fitness) has also the opportunity to become parents, while a fixed percentage of the chromosomes are killed.

After the selection the evolution procedure is performed, which consists in a crossover and a mutation step. In order to generate offspring, in this work the single-point crossover method is used: two parents generate two offspring by passing on parts of their chromosomes. The method exchanges the bits of the chromosome's bitstream beginning at a single random point and stores them in the chromosomes of the offspring. Besides the single-point crossover, there are some different crossover methods, such as shuffle crossover or n-points crossover methods.

4.6.2 Mapping objective function values to fitness

Fitness values must be a nonnegative figure of merit to be maximized; it is therefore often necessary to map the underlying problem-dependent objective function to a fitness function from one or more mappings. For this defined optimization problem for modal identification, the objective is to minimize a squared error function of the residues, the following error-to-fitness transformation is used:

$$fitness(x) = \max(C_{\max} - g(x), 0) \tag{4.7}$$

where C_{max} may be taken as the largest value of squared error observed thus far; in the simulations, it is taken as a fixed value based on a cursory estimate; g(x) denotes the objective function value computed of each individual x.

4.6.3 The selection scheme and the elitist strategy

The selection scheme used in this work is the Standard proportional selection for maximization problems. The probability of selecting an individual among a population of N individuals is given by

$$P(i) = \frac{Fit(i)}{\sum_{j=1}^{N} Fit(j)} \tag{4.8}$$

This selection scheme is simple, however it has the potential problem that the best individual in the population may fail to reproduce; thus a modified elitist strategy is used along the selection to make sure that the best member survives; this strategy uses the best individual in the previous generation to replace the worst individual in the current generation if the best member of the current generation is worse then the best member of the previous generation.

4.6.4 Implementation of the Genetic Algorithm

At beginning, we need to fix the following parameters:

- search domain of each modal parameter;
- chromosome representation of modal parameters;
- the population size and the randomly generated starting points;
- the mutation probability and the crossover probability;
- the maximum number of successive generations.

The search domain of frequency and damping parameters can be defined using a lower bound and an upper bound based on *a priori* knowledge; usually the lower bound is defined as zero, and the upper bound of frequency is taken as highest significant frequency value, while the upper bound of damping can be taken as a high value that we believe to be an upper limit for this kind of structures, say 0.5 for most civil structures. One should reduce this range as small as possible, so that the program can converge faster.

This paper adopts a binary encoding of the modal parameters (number of chromosome bits controls the resolution of identified parameters).

The initial population is randomly generated in the search domain based on an uniform distribution. One benefit of the binary encoding is that each generation of

chromosomes will always have parameter values in the pre-defined search domain, even undergo mutation and crossover.

To avoid missing global minimum, it is preferable to start the genetic process with a large population, and to have a bigger variation step for each variable. However, the computation load will be increased accordingly, since for every individual in the population, we need to run a whole simulation to compute the fitness value. After a fixed number of generations, ideally, some of the best individuals in the population will accumulate in a small region containing the true solution during evolution, and they can be used as the initial simplex for Nelder-Mead search method.

4.7 Examples

4.7.1 SDOF example

Consider the single degree-of-freedom system shown in Figure 3.2. The Genetic Algorithm illustrated above is used to solve the identification of this simple model. The range for the initial population of k is chosen as 1~100 kN/cm, and the range for the initial population of damping ratio is chosen as 0.01~0.20. These are very wide search domain; we can safely assume the true values must be within this range. A moderate measurement noise level of 20% is assumed, i.e., the prediction error has a covariance that is 20% of the root-mean-square of the response.

The parameters used in a Genetic Algorithm to solve this problem are chosen as:

population size = 50

crossover rate = 0.8

mutation rate = 0.15

Using the simulated acceleration responses as the measurement and adding noise, after 100 generations, the best member in the last generation: k = 10.306 kN/cm, ζ = 0.049, corresponding to fitness value 299.096. After sorting, the best three individuals in the last generation are:

1) var(1) = 10.306 var(2) = 0.067 fitness = 298.011
2) var(1) = 10.306 var(2) = 0.051 fitness = 299.078
3) var(1) = 10.306 var(2) = 0.049 fitness = 299.096

var(1) is the stiffness, and var(2) is damping ratio. The best member having the largest fitness value is the third one, and it is chosen as the solution given by GA. If we want to fine-tune it, this is also a very good initial guess for Nelder-Mead to start with.

In the second test, a high level of noise (100%) is added into the measurement acceleration response. After 100 generations, the best member in the last generation: k = 10.405 kN/cm, ζ = 0.045, corresponding to a fitness value 135.210. Although the fitness value calculated is much different from the noise-free case, the identified parameters do not change much. This example demonstrates the robustness of GA against noise-pollution.

For this case, the best three individuals in the last generation are:

1) var(1) = 10.405 var(2) = 0.045 fitness = 135.210
2) var(1) = 10.405 var(2) = 0.040 fitness = 135.124
3) var(1) = 10.405 var(2) = 0.028 fitness = 133.820

We observe that the best member having the largest fitness value is the first one, and it is chosen as the solution given by GA. Again, we see if we want to fine-tune it, this is also a very good initial guess for Nelder-Mead to start with.

4.7.2 MDOF example

The second example identifies the six degree-of-freedom shear building model. Three modal frequencies and damping pairs are to be identified. The range for the initial population of all frequencies is chosen as 0.5~100, and the range for the initial population of damping ratios are chosen as 0.01~0.20. They are very wide search domain; we can safely assume the true values must be within this range. A moderate level of measurement noise of 20% is assumed.

The parameters used in a Genetic program to solve this problem are chosen as:

population size = 500
crossover rate = 0.8
mutation rate = 0.15

As a first attempt and using the simulated acceleration responses as the measurement and add no noise, after 100 generations, the best member in the last generation

corresponding to the smallest fitness value are given in Table 4.6.

Table 4.6: Best member in the last generation corresponding to the smallest fitness value

Parameter	Reference value	Identified using GA	Refined by Nelder-Mead
ω_1	6.78	6.33	6.71
ω_2	17.74	19.81	18.23
ω_3	26.96	27.92	27.35
ζ_1	0.05	0.058	0.051
ζ_2	0.05	0.061	0.051
ζ_3	0.05	0.043	0.048

As we can see, the solution of GA, though not very accurate, is a very good initial guess for Nelder-Mead to start with. The fourth column in Table 4.6 is the refined parameter values by Nelder-Mead.

To best hybridize the Genetic and Nelder-Mead algorithm, we can terminate the Genetic Algorithm earlier with fewer number of generations.

4.8 Comparison of the proposed approach with PEM and LSFD method

The basic idea of system identification is to produce a model from available response data such that when the model and the real system are subjected to the same inputs, the differences between the model outputs are the true system outputs should be as small as possible.

Literally, the proposed optimization approach can also be called a kind of Prediction Error Method, as both this approach and the PEM solve unknown parameters by a nonlinear optimization procedure to minimize the output error between the model output and the measurement. The difference between the proposed approach and the PEM is that this method minimized the prediction error using a direct modal model, rather than indirectly through an ARMAX type model.

Besides the PEM, another method which inspired the proposal of the deterministic optimization approach is the least square frequency domain method (LSFD) proposed by

Iwaniec and Uhl (Iwaniec, 2003). The LSFD identifies the modal model properties by fitting the FRF (minimize the difference between the frequency response between the measurement and the model).

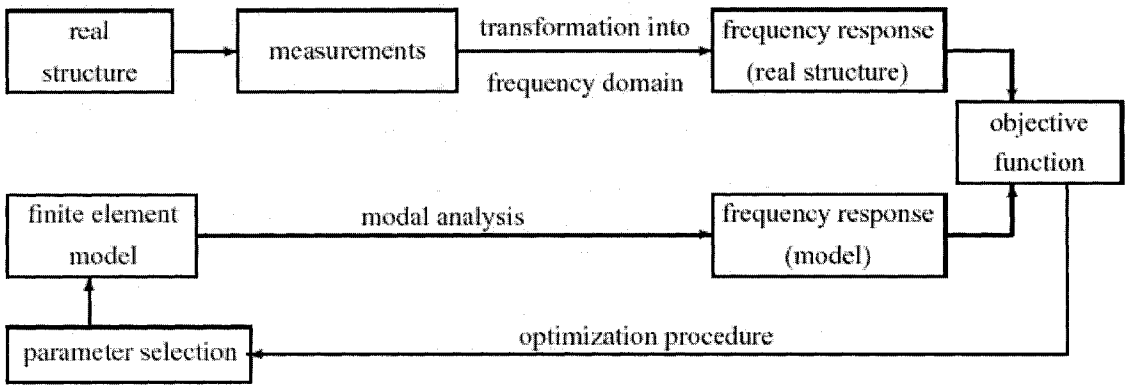


Figure 4.6: Flowchart of LSFD

The flowchart of LSFD is shown in Figure 4.7. Similarly, we can write the flow chart for the proposed deterministic optimization approach.

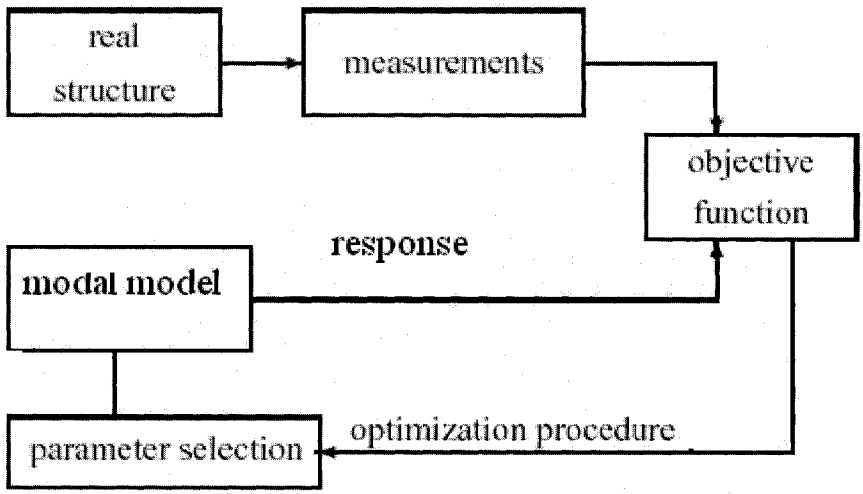


Figure 4.7: Flowchart of the proposed optimization approach

The formulation in this approach is more direct and simpler, however there is no available information of the derivatives of the objective function; for both PEM and LSFD, the derivatives can be computed without difficulty using the parameterized FRF. All of them lead to optimization problem to be solved for parameters. Therefore the PEM

and the LSFD, along with their approximations can also be called optimization approaches. Why the PEM family is so popular? The reason is that the ARMAX model in PEM and the parameterized FRF in LSFD offer explicit calculation of gradients and Hessians of the cost function, and therefore enables the application of general and very fast optimization algorithm such as Gauss-Newton and Levenberg-Marquardt methods. The PEM function implemented in MATLAB system identification toolbox (Ljung, 2004) uses Levenberg-Marquardt algorithms and the LSFD implemented in VIOMA (Virtual in Operational Modal Analysis toolbox) toolbox (Iwaniec, 2003) uses Gauss-Newton algorithm to solve the involved optimization problems.

4.9 Conclusions and remarks

The proposed deterministic optimization approach that solves modal identification problems as a nonlinear Least-Square optimization problem is very robust against noise corruption in the measurement; it gives good estimations of modal parameters if the global optimum of the objective function is correctly found by the optimization algorithm, and even the damping ratios can be accurately estimated, which is a remarkable feature for civil engineering applications. These advantages ensure it is worth trying in the input-output identification and is a promising direction for future research.

The solution of the formulated optimization is a very challenging task from computational point of view. However, it has been shown that global convergent methods, such as the Genetic Algorithm, can be used to search a good initial guess for local convergent methods to perform fine-tuning. With good initial iterations, the Nelder-Mead direct search method works well in this optimization problem for modal identification.

For a better direct optimization approach, specific optimization codes that combines the model-based methods and the continuation techniques, should be developed specifically for modal identification problem, as the research work done in many diverse research fields that involves a complicated, external non-smooth evaluation of the cost functions.

CHAPTER FIVE

PROBABLISTIC OPTIMIZATION USING BAYESIAN ESTIMATION

5.1 General

This statistical optimization approach is an application of the probabilistic Bayesian minimum-mean-square-error (MMSE) estimator in the identification of modal parameters. Due to the difficulty in evaluating the Bayesian MMSE estimation, a suboptimal estimator is sought instead, which is the maximum a posteriori (MAP) estimator. The objective function is therefore defined as the a posteriori probability density or its reduced likelihood function. A nonlinear optimization problem is then formed to maximize the Bayesian a posteriori Probability Distribution Function (PDF), and solved by the Nelder-Mead simplex method due to the lack of explicit gradient formulation which is a similar difficulty to the deterministic optimization approach described in the last chapter.

The most significant particularity of the Bayesian approach is that the parameter vector $\theta=\{\omega_1,\zeta_1,\omega_2,\zeta_2,...\}$ of interest is assumed to be a random variable or a random vector whose particular realization must be estimated. This is in contrast with the classical approaches of system identification that the parameters are assumed to be deterministic but unknown values.

The first significant advantage of the Bayesian approach is that it offers direct incorporation of our prior knowledge about the modal properties, if any, into the identification process. It is a fundamental rule of estimation theory that the use of prior knowledge will lead to an increase of the accuracy in the estimation. However it is difficult to make use of any prior knowledge directly in the identification process using deterministic approaches. In the Bayesian approach, the prior knowledge is carried out in the Bayes' theorem through a prior PDF, the resultant MMSE estimator is optimal with respect to the assumed prior PDF of θ.

The second advantage of the statistical approach is that it allows identifying not only the modal parameter values in the form of probabilistic means but also the associated

uncertainties in the form of probabilistic variances. The uncertainties of the identified modal parameters can be used in finite element model update based on identified modal parameters, where less weight can be associated to the parameters having more uncertainties. A widely employed technique in model updating is to use the inverse of the covariance of the identified parameters as the weighting matrix. Using non-statistical system identification methods, this covariance can only be obtained by Monte Carlo simulations, which is a heavy computational procedure.

The Bayesian estimation is based on the conditional PDF of observed parameters. Calculating the conditional PDF of the measurement given modal parameters requires the inversion of a response covariance matrix, which is computationally prohibitive since the number of measured data is usually very large. Two methods are combined to avoid this difficulty: one is by using the truncated expansion of the conditional PDF; the other one is by using the steepest descent algorithm to solve the inversion a much smaller matrix.

5.2 Parameter estimation using Bayesian approach

The Bayesian approach to parameter estimation assumes that the parameters to be estimated are a realization of the random variables. As a consequence, we assign a prior PDF, $P(\theta)$ to it. After some responses Y = $\{y_k, k=0,...N-1\}$ are measured, our state of knowledge about the parameter is included into conditional PDF $P(\theta \mid Y)$ (Probability density function of θ given observation Y). Using the minimum-mean-square-error (MMSE) estimator, optimal estimation is attempted by minimizing the Bayesian mean-square error (BMSE) defined as

(5.1) $$BMSE(\hat{\theta}) = E[(\theta - \hat{\theta})^2].$$

It should be emphasized that since θ is a random variable, this is a fundamentally different least-square than the deterministic least-square (LS) in some non-probabilistic approaches. The optimal solution to this estimator is the conditional mean of the parameter given observations Y, or the mean of the *a posteriori* PDF $P(\theta \mid Y)$.

(5.2) $$\hat{\theta} = E(\theta \mid Y)$$

The expectation operator is with the joint PDF of parameter and observations. The *a posteriori* PDF $P(\theta \mid Y)$ refers to the PDF of θ after the observation or measurements

have been made. The Bayesian idea is that the reduction of out uncertainty about θ and the increase in our knowledge about the parameters to be identified are represented in the *a posteriori* PDF $P(\theta \mid Y)$.

Using Bayes' theorem, the posterior PDF of the parameters given the measurement series Y = {y(0),y(1),...,y(N-1)}^{T} is given by (Kay, 1993)

$$P(\theta \mid Y) = \frac{P(\theta)P(Y \mid \theta)}{P(Y)} = \frac{P(\theta)P(Y \mid \theta)}{\int P(\theta)P(Y \mid \theta)d\theta}. \tag{5.3}$$

The numerator is the *a posteriori* PDF of θ, which is a production of the prior PDF $P(\theta)$ and the conditional PDF of Y given a particular set of parameter values. The denominator is a normalizing constant to ensure the posterior PDF integrate to unity which is necessary for a valid PDF function; it has no effect over the selection of θ. To bypass the difficulty in evaluating the integral of $P(\theta \mid Y)$ in calculating the mean, we seek a suboptimal estimate of the parameters by solving a nonlinear optimization problem:

$$\hat{\theta} = \min_{\theta} \; J(\theta) = -\ln(P(\theta)P(Y \mid \theta)) \tag{5.4}$$

where J denotes the cost function to be minimized, which is equivalent to maximize the a posteriori PDF, and therefore it is called MAP estimation in the statistics literature. Once the most probable parameter has been obtained by solving this optimization problem, the covariance of the parameters can also be computed by the inverse of the Hessian of the cost function at $\theta = \hat{\theta}$, that is:

$$R_{\hat{\theta}} = H^{-1}(J(\hat{\theta})) \tag{5.5}$$

The Hessian is computed using a numerical derivative. In other words, the identified modal parameters are also Gaussian random variables and have a distribution

$$\theta \sim N(\hat{\theta}, R_{\hat{\theta}}) \tag{5.6}$$

Prior PDF $P(\theta)$ incorporates any prior knowledge and engineering judgment about the parameters to be identified. If no prior knowledge is available, a constant value can be assigned indicating that we have no preference over any choice of θ. In this case, the remaining term $P(Y \mid \theta)$ is the likelihood function $L(\theta, Y) = P(Y \mid \theta)$, and the solution is reduced to the maximum likelihood estimation (MLS).

Based on our Gaussian assumptions of the excitations and response, the term

$P(Y \mid \theta)$ follows an N-variate Gaussian distribution with zero mean and covariance matrix R_y. Since a Gaussian distribution is defined completely by its mean and covariance, one needs only to find the solution of the covariance matrix for $P(Y \mid \theta)$ from the given parameters.

Given a set of parameter values, the response covariance can be computed from stochastic vibration theory. In the next two sections, we present the computation for the simple SDOF model, and then extend the formulation to MDOF case.

5.3 Formulations for SDOF system

In the case of SDOF, the parameter vector θ to be identified is comprised of $\theta = [\omega_0, \zeta, S_0, \sigma_v]$, where we have not only the natural frequency ω_0, damping ratio ζ, but also the input spectral density S_0 and spectral density of measurement noise σ_v. They are all needed to calculate the stochastic response. If the input power spectrum S_0 is known, we have a parameter vector of three entries.

The differential equation governing the motion of SDOF oscillator is given by:

$$\ddot{x}(t) + 2\zeta\omega_0\dot{x}(t) + \omega_0^2 x(t) = f(t) \tag{5.7}$$

where ω_0 and ζ are the natural frequency and damping ratio of the oscillator, respectively, and f(t) the ambient excitation. Assume that *f(t)* is a Gaussian white noise with constant spectral density:

$$S_f(\omega) = S_0 \tag{5.8}$$

From stochastic analysis of structural vibration, we know the displacement, velocity, and acceleration responses are also Gaussian random processes with zero-means, and auto-correlation functions given by (Lutes, 1997):

$$R_x(\tau) = \frac{\pi S_0}{2\zeta\omega_0^3} e^{-\zeta\omega_0|\tau|}\left[\cos(\omega_d\tau) + \frac{\zeta\omega_0}{\omega_d}\sin(\omega_d|\tau|)\right] \tag{5.9.1}$$

$$R_{\dot{x}}(\tau) = \frac{\pi S_0}{2\zeta\omega_0} e^{-\zeta\omega_0|\tau|}\left[\cos(\omega_d\tau) + \frac{\zeta\omega_0}{\omega_d}\sin(\omega_d|\tau|)\right] \tag{5.9.2}$$

$$R_{\ddot{x}}(\tau) = \frac{\pi S_0\omega_0}{2\zeta} e^{-\zeta\omega_0|\tau|}\left[\cos(\omega_d\tau) + \frac{\zeta\omega_0}{\omega_d}\sin(\omega_d|\tau|)\right] \tag{5.9.3}$$

where $\omega_d = \omega_0\sqrt{1-\zeta^2}$ is the damped natural frequency of the SDOF oscillator.

Assume that measured responses (usually acceleration series) are available at discrete times $\{t_k = k\Delta t, k = 0,...,N-1\}$ and that there is measurement noise and modeling error, meaning there is a difference between the measured response y(k) and the model response x(k). That we have a Bayesian linear model

(5.10) $y(k) = x(k) + v(k)$ for k=0,...,N-1

Furthermore, the noise process v is assumed to be white Gaussian noise (WGN) with zero mean and covariance R_v.

(5.11) $R_v = E[vv^T] = \sigma_v^2$

The PDF of the measurement {y(k), k=0,...,N-1} for the given parameters θ is also Gaussian, given by:

(5.12) $$P(Y \mid \theta) = \frac{1}{(2\pi)^{N/2}\left|R_y\right|^{1/2}} e^{\left(-\frac{1}{2}Y^T R_y^{-1} Y\right)}$$

where Y denotes the measurement series {y(k), k =0,...,N-1} in a column vector, and |R| denotes the determinant of a matrix R. This is a N-variate Gaussian distribution with zero-mean and covariance matrix R_y which is a function of the parameters θ. For the SDOF problem, we have (Sayed, 2003):

(5.13) $R_y(\theta) = R_x + R_v$

where the autocorrelation R_x is a Toeplitz matrix with the first row as:

$\{R_x((i-1)\Delta t)$, i=1,...,N$\}$

The MATLAB function "R_x=toeplotz(first row of R_x)" can be used to generate the whole Toeplitz matrix providing its first row.

5.4 Formulations for the MDOF system

Although the number of vibration modes equals the number of degrees of freedom, only the lower modes contribute significantly to the response, we only identify the modal parameters corresponding to these modes. The parameter vector θ to be identified are as follows:

1) Modal frequencies ω_i and modal damping ratios ζ_i of the lowest N_m number of

modes.

2) The modal shape vector components at observed degrees (one degree of freedom must be used as the normalized point which has a component of unity).
3) The parameters defining the forcing spectral $S_g(\omega)$, if not known, and the noise covariance R_v.

In modal analysis, the scaling of each mode shape is chosen such that one of its components corresponding to a measured DOF must be equal to unity. Since such scaling is arbitrary, the modal vectors can be identified only up to a constant scaling factor in system identification of modal properties. Therefore if there is only one single output, the mode shape is not presented in the estimation parameter vector.

We need an expression for the response covariance matrix of MDOF systems which is counterpart of the SDOF equations (5.9). The equation of motion for a damped dynamic system with n-degrees of freedom (DOF) using finite element method can be formulated as:

(5.14) $$M\{\ddot{x}\} + C\{\dot{x}\} + K\{x\} = \{f(t)\}$$

Here M, C, K$\in R_{n\times n}$ denote the mass, damping and stiffness matrix, whereas {f(t)} denotes the load vector. The load vector is modeled by:

(5.15) $$\{f(t)\} = T\{g(t)\}$$

where T is a force distribution matrix, and *g(t)* is a Gaussian stationary stochastic process with zero mean and spectral density $S_g(\omega)$.

The modal matrix $\Phi = \{\Phi_k\}$ (k=1...n) consists of the normalized eigenvectors that are solutions of the eigen-problem $K = \lambda M$ with respect to the eigenvalues. The modal transformation using Φ transforms the mass and the stiffness matrix into diagonal matrices. In most cases, the proportional damping is used, and therefore the damping matrix C is also transformed into a diagonal matrix by modal transformation. The uncoupled equations in the modal coordinates are

(5.16) $$\ddot{q}_k(t) + 2\zeta_k\omega_k\dot{q}_k(t) + \omega_k^2 q_k(t) = \frac{p_k(t)}{m_k}, \qquad k=1,\ldots,n$$

where $\{p_k(t)\} = \Phi^T T\{g(t)\}$ is the modal forcing vector, $\{q(t)\} = \Phi^T\{x(t)\}$ is the modal

coordinate vector, and m_k is the modal mass. Now, the modal forcing vector $\{p_k(t)\}$ is a Gaussian stationary stochastic process with zero mean and spectral density matrix

(5.17) $$S_p(\omega) = \Phi^T M^{-1} T S_g(\omega) T^T M^{-1} \Phi$$

The covariance matrix of the modal force function is the inverse Fourier transform of the spectral density matrix

(5.18) $$R_f = \int_{-\infty}^{\infty} S_f(\omega) e^{j\omega t} dt$$

Having the input spectrum in the modal coordinate, we can calculate the response spectrum using modal composition method. The spectral density of the response in modal coordinates is

(5.19) $$S_q^{r,s}(\omega) = \frac{S_p^{r,s}(\omega)}{\left[\left(\omega_r^2 - \omega^2\right)^2 + 2j\zeta_r\omega_r\omega\right]\left[\left(\omega_s^2 - \omega^2\right)^2 + 2j\zeta_s\omega_s\omega\right]}.$$

where $j = \sqrt{-1}$; the denominator is the harmonic transfer function for MDOF system.

Since the transformation from the modal coordinates to the physical coordinates is given by:

(5.20) $$\{x(t)\} = \Phi\{q(t)\}$$

The transformation from the modal response spectrum to the physical response spectrum is:

(5.21) $$S_x(\omega) = \Phi S_q(\omega) \Phi^T$$

The response spectral density can be expressed as follows (Lutes 1997):

(5.22) $$S_x^{l,k}(\omega) = \sum_{r=1}^{Nm}\sum_{s=1}^{Nm} \phi_l^{(r)}\phi_k^{(s)} \frac{S_p^{r,s}(\omega)}{\left[\left(\omega_r^2 - \omega^2\right)^2 + 2j\zeta_r\omega_r\omega\right]\left[\left(\omega_s^2 - \omega^2\right)^2 + 2j\zeta_s\omega_s\omega\right]}$$

and the state response correlation function is given by a inverse Fourier transform of the state response spectrum.

(5.23) $$R_x^{l,k} = \int_{-\infty}^{\infty} S_x^{l,k}(\omega) e^{j\omega t} dt$$

A single-output identification problem is considered, the measured responses are available at a single node at discrete times $t_k = k\Delta t$, where k varies between 0 to N-1. Usually measurement noise and modeling error are inevitable, meaning that there is a difference between the measured response y(k) and the model response corresponding to

the measurement; we can write the Bayesian linear model that is counterpart to the SDOF case equation (5.10):

(5.24) $$\{y(k)\} = H\{x(k)\} + \{v(k)\}$$

Where H is the observation matrix consisting of zeros and ones, y(k) is the measurement data vector, the noise process v is assumed to be white Gaussian noise (WGN) with zero mean and covariance R_v.

(5.25) $$R_v = E[vv^T] = \sigma_v^2 \mathrm{I}$$

where I denotes the unit matrix that is given by I=diag(1,…,1).

Since the measured responses or the observations are linear function of the model response, the measurements *y(k)* is zero-mean and the covariance of the measurement is

(5.26.1) $$R_y = E[(y-\bar{y})(y-\bar{y})^T] = E[yy^T] = HR_xH^T + R_v$$

and

(5.26.2) $$R_{xy} = R_xH^T$$

where R_x denotes the covariance matrix of the model response x(t) calculated using modal superposition method, and R_v is the covariance of measurement noise.

5.5 Calculating the cost function

The conditional PDF can be computed using the equation (5.12) along with equation (5.13) for SDOF systems and equation (5.26) for MDOF systems. Nonetheless, there is still a difficulty in the calculation of conditional PDF: in the formula of the multi-variate Gaussian distribution, we need to calculate the determinant and the inverse of an N-by-N covariance matrix R_y. However, the number of measurement points is usually a large number, and hence make the computation of PDF expensive, even practically infeasible. To overcome this difficulty, a combination of two techniques is adopted to bypass the matrix inversion and enable the numerical computation of the cost function. The first is to expand the conditional PDF of total response vector as a multiplication of conditional PDF of response at each individual time step, and approximate the response process as a finite-order Markov process; the second is to use an iterative optimization algorithm to solve the conditional mean and covariance of each individual conditional PDF.

5.5.1 Expansion and truncation of the conditional PDF

Beck and Yuen et al. (Beck, 1998) (Yuen, 2005) developed an approximate expanded formulation for the calculation of the conditional PDF; using Bayes' rule, the likelihood function $P(\theta \mid Y)$ can be written in terms of the transition probability density as:

$$P(Y \mid \theta) = P(y(0),..., y(N_p - 1) \mid \theta) \prod_{k=Np}^{N-1} P(y(k) \mid \theta; y(0),..., y(k-1)) \tag{5.27}$$

where N_P is chosen as a number much smaller than the total measurement point number N. The measurements at different time steps Y = {y(k),k=0,…,N-1} are not independent to each other, however, the autocorrelation of the model responses R_x is an oscillatory decreasing function, the envelope of this correlation is monotonically decreasing.

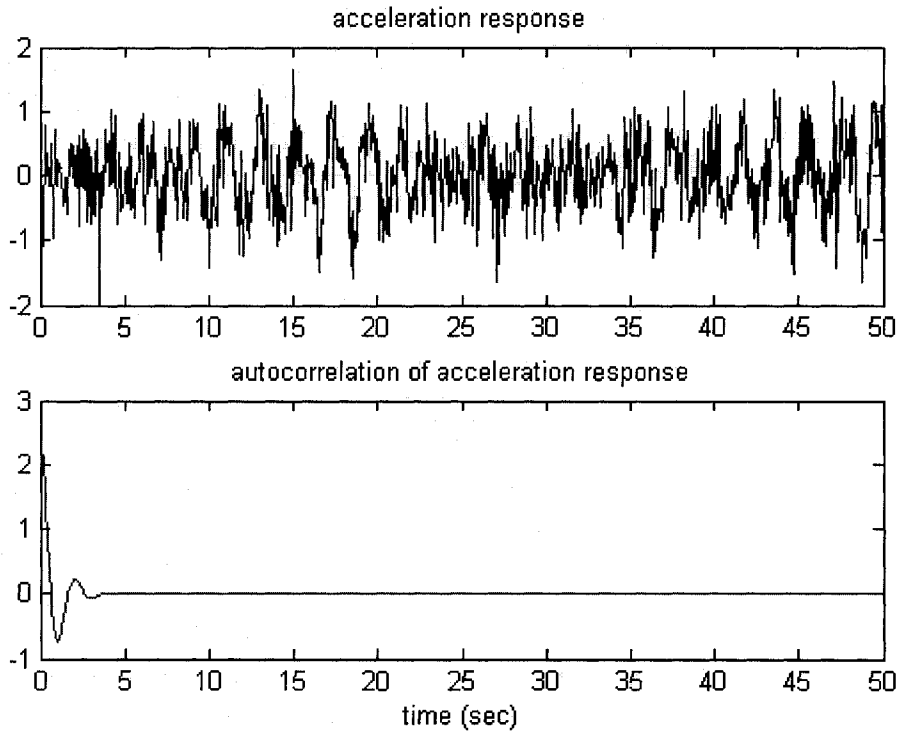


Figure 5.1: Typical acceleration response and its autocorrelation

Therefore, measurement points with a large time interval in between can be regarded as independent; in this respect the measurement series can be regarded as a Markov

process of order k if $R_x(\Delta t) \cong 0$. A process x(t) is Markov of order k if

$$P(x_m \mid x_{m-1}, x_{m-2}, \ldots, x_1) = P(x_m \mid x_{m-1}, x_{m-2}, \ldots, x_{m-k}) \quad \text{for k<m} \tag{5.28}$$

For the Markov-k process, we can approximate the transition probability density in the sense that the data points too far in the past have no effect on the statistical behaviour of a present point. Thus the equation of $P(Y \mid \theta)$ can be greatly simplified as:

$$P(Y \mid \theta) \cong P(y(0), \ldots, y(N_p - 1) \mid \theta) \prod_{k=Np}^{N-1} P(y(k) \mid \theta; y(k - N_p), \ldots, y(k-1))$$

That is, the conditional probability of current response depending on all the past observation data can be approximated by conditional probabilities depend only the last N_P observation data points. Since the Markov order N_P can be chosen as a number much smaller than N, the calculation of probability density function would thus require the inversion and determinant of a N_P–by-N_P matrix rather than the original N-by-N one. The choice of N_P is such that the correlation functions must have decayed to very small values over this time lag. Through numerical testing, it has been found that a value of N_P of the order of $(5 \sim 10 T_{max}/\Delta t)$ is sufficient, where T_{max} is the fundamental period of the structure, and Δt the sampling time step. Using a larger value of N_P gives no improvements in the calculation of conditional PDF.

To start the computation, one need to calculate a reduced PDF of the first N_P observation points $P(y(0), \ldots, y(N_p - 1) \mid \theta)$ in the same way with the complete PDF of N observation points; the inversion of the N_P-by-N_P Toeplitz matrix would not then be a difficult task..

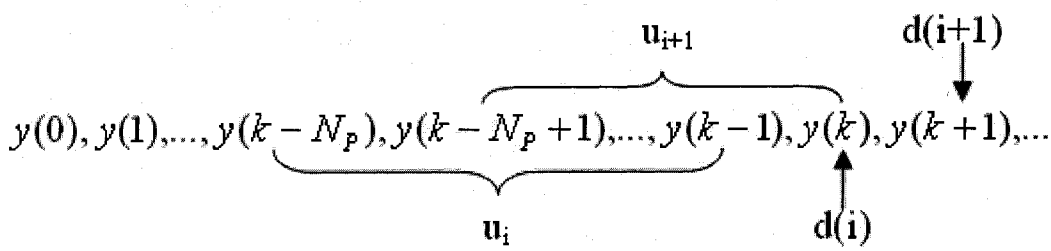
Next, we need an expression for the conditional probabilities $P(y(k) \mid \theta; y(k - N_p), \ldots, y(k-1))$, k=$N_P$,…,N-1. Since both the model response x(t) and the measurement noise v(t) are zero-mean Gaussian processes, y(t) is also a zero-mean Gaussian process.

Following the notation used by statistical signal processing literatures (Sayed, 2003), we write:

$$\begin{cases} \text{d(i)} = \text{y(k) is a scalar variable} \\ u_i = [y(k - N_P), y(k - N_P + 1), \ldots, y(k-1)] \text{ is a row vector} \end{cases}$$

We use i, instead of k, for the index of d and u to avoid confusion. Both of d(i) and u_i

change when we move forward to the end of measurements.



For Gaussian processes, the optimal estimator in the mean-square sense is the Bayesian minimum-mean-square-error (MMSE) estimator. The mean of y(k) given previous N_P number of observations {y(k-N_P),...,y(k-1)} is $\hat{\mu}(k)$ given by

(5.29.1) $$\hat{\mu}(k) = E[y(k) \mid y(k-N_P),..., y(k-1)] = R_u^{-1} R_{du} Y_{N_P}$$

The corresponding mean square error $\tilde{\mu}(k) = y(k) - \hat{\mu}(k)$ (also called the covariance matrix of the prediction error) is

(5.29.2) $$R_{\tilde{\mu}}(k) = E[\tilde{\mu}(k)\tilde{\mu}(k)^T] = R_d - R_{du}^T R_u^{-1} R_{du}$$

where Y_{Np} is the vector {y(k-N_P),...,y(k-1)}. R_d has a dimension of 1-by-1; and R_{du} has a dimension 1-by-N_P and R_u has a dimension N_P-by-N_P; they are submatrices of the covariance matrix of Z={y(k-N_P),...,y(k)}.

(5.30) $$R_Z = \begin{bmatrix} R_d & R_{du} \\ R_{du}^T & R_u \end{bmatrix}$$

Knowing the mean and covariance, the Gaussian distribution can be completely defined, that is:

$$P(y(k) \mid \theta; y(k-N_P),..., y(k-1)) = \frac{1}{(2\pi)^{1/2} |R_{\tilde{\mu}}|^{1/2}} \exp\left\{-\frac{1}{2}[y(k) - \hat{\mu}(k)]^T R_{\tilde{\mu}}^{-1}(k)[y(k) - \hat{\mu}(k)]\right\}$$

(5.31) $$= \frac{1}{(2\pi)^{1/2} |R_{\tilde{\mu}}|^{1/2}} \exp\left\{-\frac{1}{2}\tilde{\mu}(k)^T R_{\tilde{\mu}}^{-1}(k)\tilde{\mu}(k)\right\}$$

It is worth noting that the error covariance $R_{\tilde{\mu}}(k)$ is actually independent of the index k; it is evaluated only once during the analysis.

5.5.2 Steepest descent algorithm for calculating conditional PDF

The expanded form for conditional PDF requires the inversion of matrix R_u at each

time step. If the fundamental period of structure is low and/or the sampling rate is high, the size of the R_u can be large; the numerical error in calculating $R_u^{-1}R_{du}$ can jeopardize the final prediction. The inversion of matrix R_u can be avoided by reworking the recursion into a simple optimization form for the MMSE estimation (Sayed, 2003):

(5.32) $$\min_{w_i} E\left\|d - uw_i\right|^2\Big]$$

This optimization problem can be solved iteratively using the steepest descent algorithm, and the converged optimal solution is (Sayed, 2003):

(5.33) $$w^o = R_u^{-1}R_{du}$$

This steepest descent algorithm for computing data statistics $R_u^{-1}R_{du}$ is the basis for Least-Mean-Square (LMS) adaptive filtering (Sayed 2003), it is summarized as follows:

Step-1: let the weight $w_{-1}=0$ *as the initial guess.*

Step-2: choose a stepsize μ *that*

$$0 < \mu < \frac{2}{\lambda_{\max}}$$

where $\lambda_{\max}$ *is the maximum eigenvalue of the covariance matrix* R_u.

Step-3: iterate for $i \geq 0$.

$$w_i = w_{i-1} + \mu(R_{du} - R_u w_{i-1})$$

With the converged weight w, we can compute the MMSE estimation mean and covariance using equation (5.29).

The choice of a prior PDF is critical in Bayesian estimation. In this work, all variables are assumed to be Gaussian. Gaussian PDFs are quite useful due to their mathematical properties. Furthermore, Gaussian PDFs occur naturally in many practical applications according to Kolmogorov's central limit theory (Lutes, 1997).

5.6. Examples:

5.6.1 Example-1: Transient response of a SDOF oscillator

In this example, we consider the single degree-of-freedom (SDOF) spring-mass system presented in Figure 3.2 in Chapter 3, subjected to white Gaussian noise with spectral intensity S_0. A fourth-order Runge-Kutta method is used to integrate the system's response under generated noisy input. The parameters to be identified are $\theta = [\omega_0, \zeta, S_0, S_v]$, including natural frequency ω_0, damping ratio ζ, input spectral density S_0 and spectral density of measurement noise σ_v. The parameter values used to generate the simulated data are S_0 =1.0 cm^2/sec^3, and the noise variance σ_v^2 =0.251 for measured acceleration response. The chosen value of noise variance corresponds to a 10 per cent root-mean-square (rms) prediction-error level, i.e., the noise is 10% of the rms of the noise-free response. The time step used to generate the data is 0.01 sec. However, a much larger sampling time step was chosen (Δt = 0.1 sec) and the total time interval is T =1000 sec, so that the number of data points is N =10000. It is assumed in this identification that we have no prior information about the model, and the prior PDF of parameters is assumed a non-informative one. As explained in section 5.2, this reduces our identification to a Maximum Likelihood estimation.

The Nelder-Mead simplex method is used to solve the formulated optimization problem. Figure (5.2) shows a typical simulated acceleration measurement with noise added in it. Figure (5.3) shows a typical convergence history of the Nelder-Mead method.

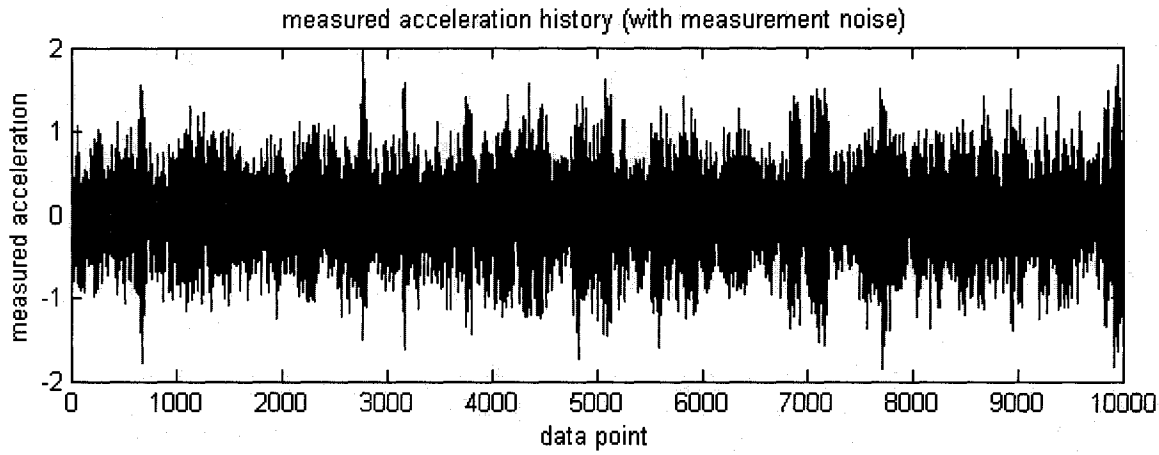


Figure 5.2: The acceleration measurement with noise

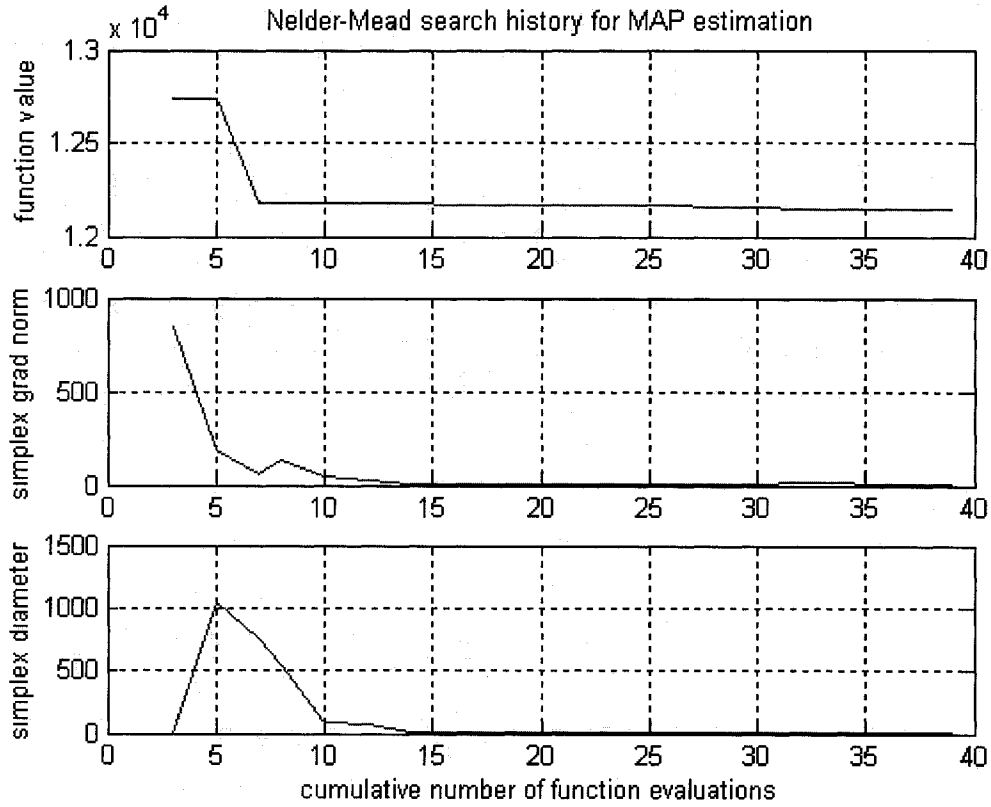


Figure 5.3: The convergence history of the Nelder-Mead method

The identified results using one simulated output acceleration series is presented in Table 5.1. The standard deviation is the square root of variances of the unknown parameters, and the coefficient of variation is the ratio of standard deviation and true value. The value N_P is chosen as 100 which corresponds to a time lag of twenty times the

fundamental period.

Table 5.1: Identification results

Parameter	Reference value	Identified mean	Standard deviation	Coefficient of variation
ω_o	3.16	3.17	0.0052	0.0016
ζ	0.05	0.050	0.0017	0.0340
S_0	1.0	1.012	0.1106	0.1106
S_v	0.251	0.242	0.0124	0.0255

Figure (5.4) shows the conditional PDF $P(\omega_0 \mid Y, \zeta, S_0, S_v)$, and the conditional PDF $P(\zeta \mid Y, \omega_0, S_0, S_v)$. They are Gaussian distributions with very small covariance. We can take the mean as the identified values for the parameters.

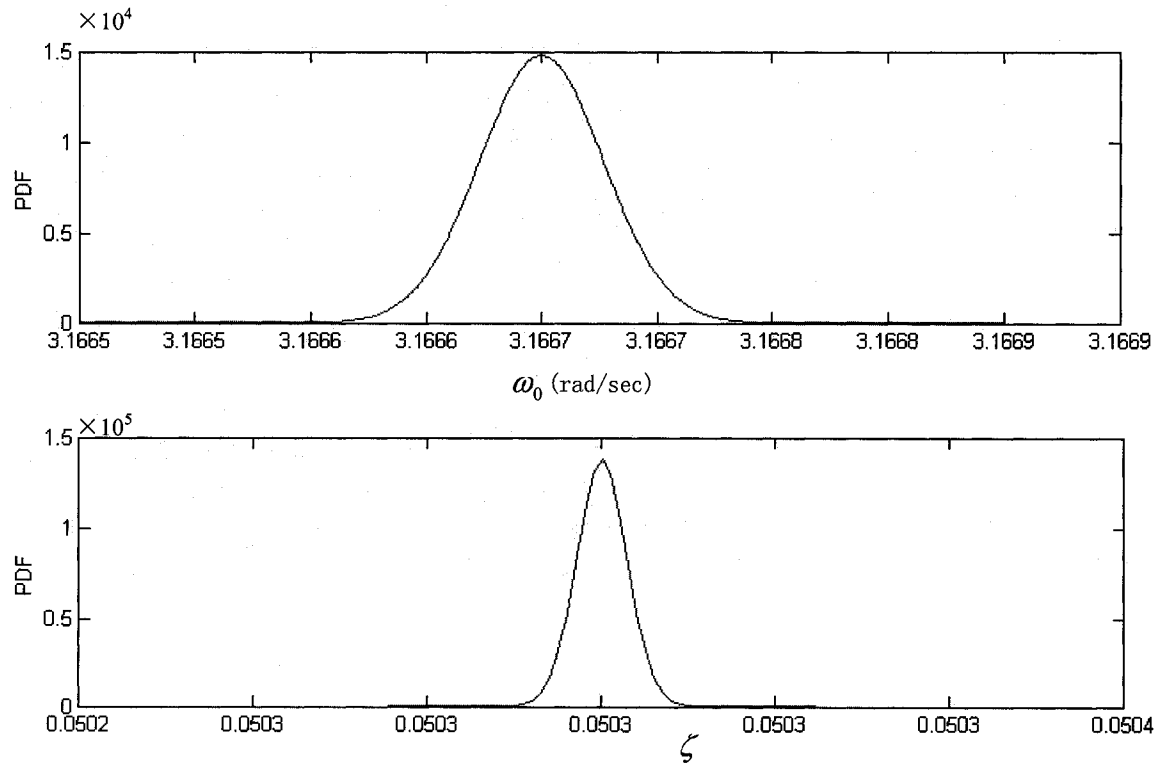


Figure 5.4: The probabilistic distribution of frequency and damping given observations

5.6.2 Example 2: six storey building

The second example refers to the six-story building shown in Figure 3.5 (chapter 3). This is the same model used in (Huang, 2001). This structure is excited at its base with Gaussian white noise. Acceleration responses are measured at the top floor and the third floor. The base excitation is equivalent to application of the same acceleration excitation to every floor. The damping ratios are chosen to be 5% for all modes. The measurement noise level is assumed to be 10% of the rms of the noise-free acceleration response at corresponding channels.

The time step used to generate the data is 0.01 sec. However, a larger sampling time step was chosen ($\Delta t = 0.02$ sec) and the total time interval is T =100 sec, so that the number of measured data points is N =5000. It is assumed in this identification that we have no prior information about the model, and the prior PDF of parameters is assumed a non-informative one. As explained in section 5.2, this reduces our identification to Maximum Likelihood estimation.

The identification using this statistical optimization approach is carried out using a single measurement series from the top storey and from the third storey respectively. Figure (5.5) and (5.6) shows the spectral density estimated using Bartlett method corresponding to the 3^{rd} and 6^{th} floor. Since the spectrum at different floor is different, one can expect naturally that the identification results obtained using measurements from different floors is also different.

In both cases, the value of the assumed Markov order N_P is chosen as 300, which covers a time lag six times the fundamental period.

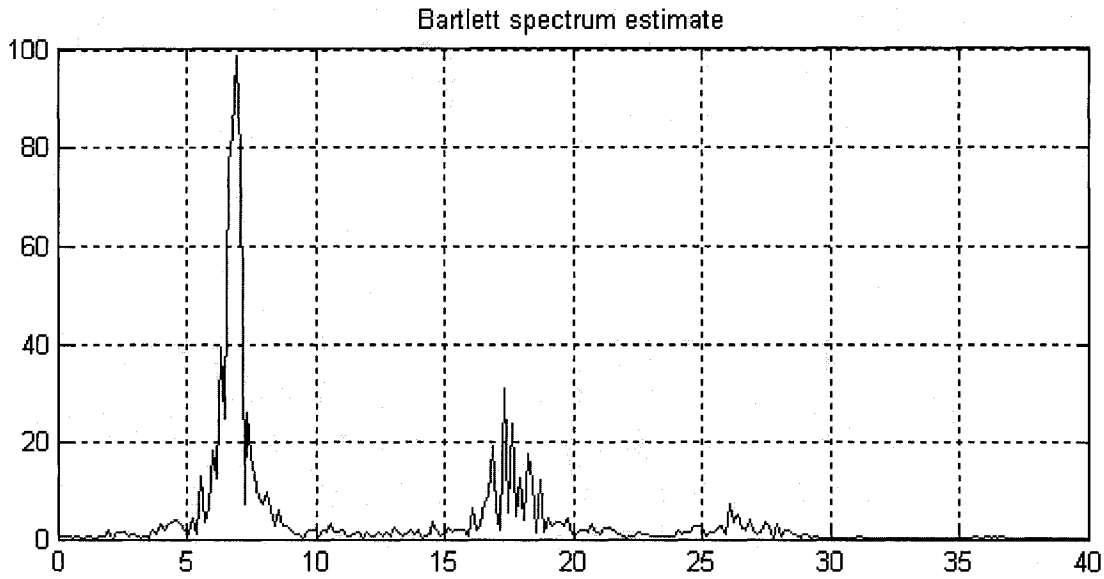


Figure 5.5: PSD plot at top floor

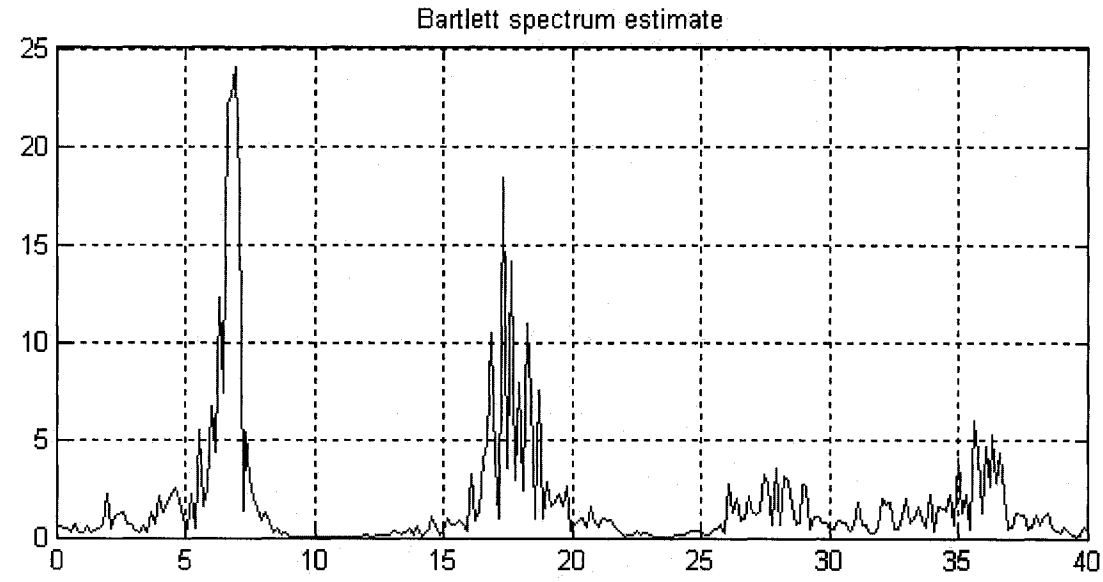


Figure 5.6: PSD plot at third floor

Table 5.2: Identification results from measurement at the top floor

Parameter	Reference value	Identified mean	Standard deviation	Coefficient of variation
ω_1	6.78	6.92	0.343	0.051
ω_2	17.74	16.95	0.994	0.056
ζ_1	0.05	0.0545	0.008	0.156
ζ_2	0.05	0.0549	0.007	0.144
S_P^1	18.51	17.215	3.253	0.176
S_P^2	2.44	2.879	0.789	0.324
$S_P^{1,2}$	-6.72	-6.218	1.893	0.282
σ_v^2	0.095	0.127	0.030	0.316

Table 5.3: Identification results from measurement at the third floor

Parameter	Reference value	Identified mean	Standard deviation	Coefficient of variation
ω_1	6.78	6.47	0.512	0.076
ω_2	17.74	17.25	0.734	0.0414
ζ_1	0.05	0.06	0.013	0.254
ζ_2	0.05	0.049	0.005	0.098
S_P^1	18.51	19.974	3.862	0.209
S_P^2	2.44	2.715	0.564	0.231
$S_P^{1,2}$	-6.72	-7.317	1.774	0.264
σ_v^2	0.095	0.105	0.029	0.301

In both cases, we see that the coefficient of variation of the identified frequencies are much smaller than the COV (Coefficient of Variation) of the identified damping ratios, this fact indicates that frequencies are identified much better than damping. Similar to most identification methods, accurate identification of damping is more difficult than frequencies.

In the first case, the COV of the first two identified modes are similar, but in the second case, the COV of the second mode is significantly smaller than COV of the first mode. This should not be a surprise; the first mode amplitude is larger at the top floor, while in the second mode, the third floor amplitude is larger in the second mode than its relative amplitude in the first mode. Therefore, the measurement at the third floor should contain more information about the second mode. Choosing the measurement point is of the most importance if higher modes identification is of interest.

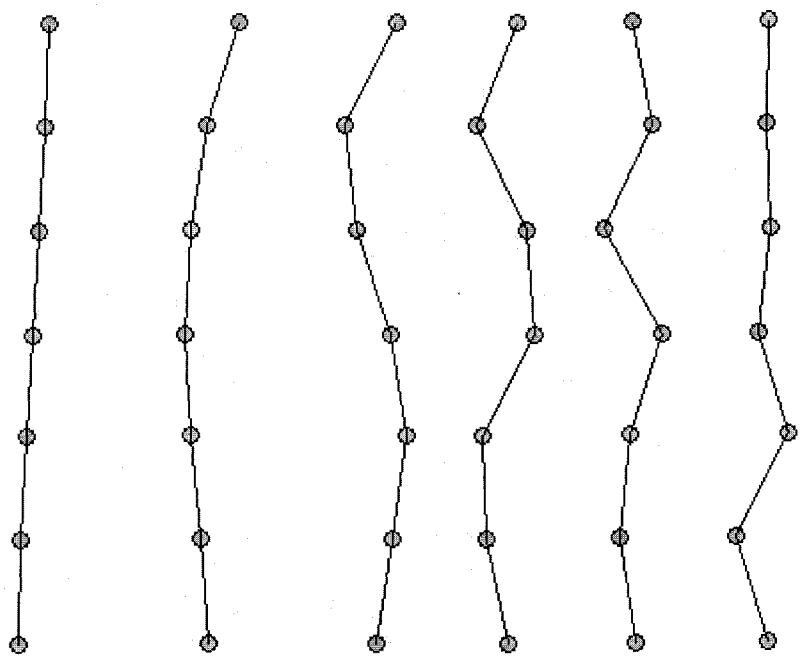


Figure 5.7: The six mode shapes (simulated in MATLAB)

It is a common sense that the identification of mode shapes requires the measurement at multiple nodes. For output-only identification, elements in modal shape vectors can only be identified at measured points. If there is only a single output, then it is not possible to identify the modal shape, since the depiction of modal shape requires at least

two points.

For input-output type identification, the situation is a little different, that the shape vector at input points can also be identified, since the calculation of the output response will require the knowledge of modal shape vectors at the input points also.

In this optimization approach, the identifiability of any parameter depends on whether there is a definable functional relationship between the parameter and the cost function. From equation (5.22), we see if the excitation is at the same floor with the measurement, then the modal shape is unidentifiable, however in the base-excitation case, the input is equivalent to the applied acceleration to all the floors at the same time; this problem, though only one single-output is measured, is equivalent to a multi-input-single-output problem. One expects that the modal shapes are identifiable using this optimization approach. Nonetheless, the identifiability also depends on the power of optimization algorithm. Multiple-variate identification problems are usually the most difficult task in optimization, and adding a modal vector to the parameter vector θ would increase the size tremendously, especially in the case when the excitation is applied to every node of the structure. Therefore, an alternative is used here such that the spectrum of modal excitations is sought in the parameter vector instead of a full list of modal vectors. We use equation (5.19) instead of (5.22) in computing the objective function of a MDOF system. From equation (5.19), we see that the modal excitation, if identified, hide all the effects of modal vectors and simplifies the optimization problem considerably.

5.7 Concluding remarks

The presented probabilistic optimization approach, which is an application of the Bayesian MMSE estimator, is able to identify the modal parameters with good accuracy. It is based on the output measurement only and thus is one output-only identification method.

The application of this approach is based on the assumption that the input excitations is a Gaussian random process and the response is also a Gaussian random process, the structure is linear, the measurement noise and modeling error can also be modeled as Gaussian processes. Therefore the unknown modal parameters are also random variables with a multi-variate Gaussian PDF.

The success of this identification procedure depends on the validity of those assumptions. Although applicable to most cases, we should always check these assumptions carefully before put trust on the identified results; although these assumptions are valid in many ambient vibration tests, there are some cases that these assumptions are not valid, such as the intermittency in the ocean swell and wind excitations on structures (Simiu, 1996). The sensitivity of the results to the deviation of Gaussian distribution is unclear at this moment (Sayed, 2003).

This probabilistic approach has two significant advantages over the non-probabilistic methods: the first is that the prior knowledge can be incorporated into the identification process through in inclusion of the a priori PDF of modal parameters in equation (5.4); in some cases, there are a large number of similar structures to be identified (Ghrib, 2004), the prior knowledge can be accumulated with increased number of identified structures. The second is that the uncertainties can also be obtained through the form of probabilistic covariance; this uncertainty can be pursued further in model updating or validations.

Similar to the deterministic optimization approach presented in chapter 4, this approach also require the solution of a multi-variable optimization problem. The ability of these approaches in identifying a large number of modes is obviously dependent on the performance of the optimization technique. The direct solver for optimization problems is one current active research subject in many fields of engineering.

CHAPTER SIX

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

6.1 Summary

The identification of modal parameters has attracted much research attention in the past twenty years. There is an ever increasing demand for its application in civil engineering structures for health monitoring, damage identification and model updating.

An investigation of existing methods for the identification of modal parameters is performed, with special attention on the peak-picking method, Prediction Error Method and subspace identification methods which are commonly considered as the simplest, the most accurate, and the most powerful tools in modal identification.

Two optimization based approaches for modal identification are then presented in this thesis. The first one is a deterministic optimization approach for input-output identification; it solves modal identification problems as a nonlinear Least-Square optimization problem minimizing the Residual Sum of Squares of the error between the measured output and the output from a modal model. The second one is a probabilistic optimization approach for output-only identification; this approach is an application of the Bayesian MMSE estimator to modal identification problem, it solves the modal parameters by maximizing the a posteriori PDF of the modal parameters given observation/measurements.

These approaches require the solution of nonlinear, nonconvex, multi-variable optimization problems, and there is no derivative information available to the formulated objective functions. Two direct optimization algorithms that require the evaluation of function value only, the Nelder-Mead simplex method and the Genetic Algorithm, are combined to solve the optimum. This combination follows the exploration/exploitation paradigm, that the global optimizer GA is used to explore the search domain and locate a good initial guess for the local optimizer Nelder-Mead method to start with, and find the global minimum.

6.2 Conclusions

The following conclusions are drawn from the present study:

1) The proposed deterministic optimization approach is robust against noise corruption in the measurement; it gives good estimations of modal parameters if the global optimum of the objective function is correctly found by the optimization algorithm; even the damping ratios can be accurately estimated, which is a remarkable feature in civil engineering applications.
2) The presented probabilistic optimization approach is able to identify the modal parameters with good accuracy.
3) The probabilistic optimization approach has two significant advantages over deterministic methods: the first is that the prior knowledge can be easily incorporated into the identification process; the second is that not only the unknown parameters, but also the uncertainties associated with the identification can be obtained.
4) The proposed combined Genetic and Nelder-Mead algorithms are found to be efficient in solving direct optimization problems.

6.3 Suggestions for future work

The follow topics are recommended for future research:

1) The development of ad hoc optimization methods for these two optimization problems for modal identification. If the proposed optimization approaches are to be successful, especially for multi-mode identification, specific optimization codes that combine the model-based methods and the continuation techniques should be developed.
2) The extension of the probabilistic approach to input-output identifications with a measured deterministic sequence of input excitation.
3) Combination of probabilistic identification approach with noise-cancellation filtering. Since this approach is based on the idea of Bayesian estimation, there are many adaptive filtering tools that can be incorporated; the application of those tools in the modal identification problem is yet a field to be explored.

REFERENCES

1. Balmes, E. (1997): Structural dynamics toolbox for use with MATLAB, user's guide, version 3. Scientific Software Group, Sevres, France
2. Basseville, M., Abdelghani, M., Benveniste, A. (2000): Subspace-based fault detection algorithms for vibration monitoring. Automatica, 36(1), 101-109
3. Beck, J. L., and Katafygiotis, L. S. (1998): Updating models and their uncertainties. I: Bayesian statistical framework. Journal of Engineering Mechanics, 124(4), 455–461
4. Berghen, F. V. (2003): CONDOR: a constrained, non-linear, derivative-free parallel optimizer for continuous, high computing load, noisy objective functions. Ph.D. thesis, Universite Libre de Bruxelles, Belgium
5. Berghen, F. V., and Bersini, H. (2004): Optimization algorithm for Non-Linear, Constrained, Derivative-free optimization of Continuous, High-computing-load Functions. Technical report, Universite Libre de Bruxelles, Belgium
6. Bernstein, S., and Rishter, M. (2003): The use of genetic algorithms in finite element model identification. Research report, Deutsche Forschungsgemeinschaft
7. Capecchi, D. (1989): Difference models for identification of mechanical linear systems in dynamics. Mechanical Systems and Signal Processing, 3(2), 157-172
8. Chakraverty, S. (2005): Identification of structural parameters of multi-storey shear buildings from modal data. Earthquake Engineering and Structural Dynamics. 34, 543-554
9. Chambers, L. (2001): The practical handbook of Genetic Algorithms applications. Chapman & Hall/CRC
10. Chase, J. G., Hwang, K. L., Barroso, L. R., Mander, J. B. (2005): A simple LMS-based approach to the structural health monitoring benchmark problem. Earthquake Engineering and Structural Dynamics, 34, 575-594
11. Chase, J. G., Spieth, H. A., Blome, C. F., Mander, J. B., (2005): LMS-based structural health monitoring of a non-linear rocking structure. Earthquake Engineering and Structural Dynamics, 34, 909-930

12. Chelouah, R., and Siarry, P. (2003): Genetic and Nelder–Mead algorithms hybridized for a more accurate global optimization of continuous multiminima functions. European Journal of Operational Research, 148, 335–348
13. Chiang, D. Y., Huang, S. T. (1997): Modal parameter identification using simulated evolution. AIAA Journal, 35(7), 1204-1208
14. Chong, E. K. P., and Zak, S.H. (1996): An Introduction to Optimization. John Wiley & Sons
15. Conn, A. R., Scheinberg, K., Toint, P. L. (1997): Recent progress in unconstrained nonlinear optimization without derivatives. Mathematical Programming, 79, 397-414
16. Cooley, J. W., and Tukey, J. W., (1965): An algorithm for the machine calculation of complex Fourier series. Mathematics of Computation, 19, 297-301
17. Davidon, W. C. (1991): Variable metric method for minimization. SIAM Journal of Optimization, 1, 1-17
18. Dennis, J. E., and Torczon, V. (1991): Direct search methods on parallel machines. SIAM Journal of Optimization, 1, 448-474
19. Dunn, S. A. (1997): Modified genetic algorithm for the identification of aircraft structures. Journal of Aircraft, 34(2), 251-253
20. Ewins, D. J. (2000): Modal testing: theory, practice and application. 2nd ed., Research Studies Press Ltd., Baldock, Hertfordshire, England
21. Gersch, W., and Luo, S. (1972): Discrete time series synthesis of randomly excited structural system response. Journal of the Acoustical Society of America, 51, 402-408
22. Ghrib, F., and Mamedov, H. (2004): Period formulas of shear wall buildings with flexible bases. Earthquake Engineering and Structural Dynamics, 33, 295-314
23. Hayes, M. H. (1996): Statistical digital signal processing and modeling. John Wiley & Sons, Inc.
24. Holland, J. (1975): Adaptation in natural and artificial systems. Ann Arbor, University of Michigan Press
25. Hooke, R., and Jeeves, T. A. (1961): "Direct search" solution of numerical and statistical problems. Journal of the ACM, 8, 212-229
26. Huang, C. S. (2001): Structural identification from ambient vibration measurement using the multivariate AR model. Journal of Sound and vibration, 241(3), 337-359

27. Ibrahim, S. R. (1977): Random decrement technique for modal identification of structures. Journal of Spacecraft and Rockets, 140, 696-700
28. Ibrahim, S. R., and Mikulcik, E. C. (1976): The experimental determination of vibration parameters from time response. Shock and Vibration Bulletin, 46(5), 183-198
29. Ibrahim, S. R., Asmussen, J. C., Brincker, R. (1999): Statistical theory of the vector random decrement technique. Journal of Sound and Vibration, Vol. 226, 329-244
30. Inman, D. J. (1994): Engineering Vibration. Englewood Cliffs, NJ: Prentice-Hall
31. Iwaniec, J., and Uhl, T. (2003): The application of the nonlinear least-squares frequency-domain method to estimation of the modal model parameters. Research report, AGH University of Science and Technology, Poland
32. Jaishi, B., and Ren, W. X. (2005): Structural Finite Element Model Updating Using Ambient Vibration Test Results. Journal of Structural Engineering, 131(4), 617-628
33. Juang, J. N. (1999): Optimized system identification. Research report, NASA/TM-1999-209711.
34. Juang, J. N., and Pappa, R. S. (1985): An eigensystem realization algorithm for modal parameter identification and modal reduction. Journal of Guidance and Control Dynamics, AIAA, 8(5), 620-627
35. Juang, J. N., and Pappa, R. S. (1986): Effect of noise on modal parameters identified by the eigensystem realization algorithm. Journal of Guidance and Control Dynamics, AIAA, 9(3), 294-303
36. Katafygiotis, L. S., Papadimitriou, C., Lam, H. F. (1998): A probabilistic approach to structural model updating. Journal of Soil Dynamics and Earthquake Engineering, 17(78), 495–507.
37. Katafygiotis, L. S., Yuen, K. V, Beck, J. L. (2001): Bayesian spectral density approach for modal updating using ambient data. Earthquake Engineering and Structural Dynamics, 30, 1103-1123
38. Katayama, T., Picci, G. (1999): Realization of stochastic systems with exogenous inputs and subspace identification methods. Automatica, 35, 1635-1652
39. Kay, S. M. (1993): Fundamentals of statistical signal processing, estimation theory. Prentice Hall

40. Kelley, C. T. (1999): Iterative methods for optimization. SIAM, Philadelphia
41. Kitada, Y. (1998): Identification of nonlinear structural dynamic systems using wavelets. Journal of Engineering Mechamics, 124(10), 1059–1066.
42. Koh, C.G., Chen, Y. F., Liaw, C. Y. (2003): A hybrid computational strategy for identification of structural parameters. Computers and Structures, 81, 107-117.
43. Lee, Z. K., Wu, T. H., Loh, C. H. (2003): System identification on the seismic behavior of an isolated bridge. Earthquake Engineering and Structural Dynamics, 32, 1797–1812
44. Ljung, L. (1999): System identification: theory for the user, 2nd Ed. Prentice Hall, 1999
45. Ljung, L. (2003): Aspects and Experiences of User Choices in Subspace Identification. Report no.: LiTH-ISY-R-2564, the 13th IFAC Symposium on System Identification, Rotterdam, The Netherlands
46. Ljung, L. (2004): User's manual, The MATLAB system identification toolbox, version 6.0
47. Lus, H., Betti, R., Longman, R. W. (1999): Identification of linear structural systems using earthquake-induced vibration data. Earthquake Engineering and Structural Dynamics, 28, 1449-1467
48. Lutes, L. D., and Sarkani, S. (1997): Stochastic analysis of structural and mechanical vibrations. Prentice Hall
49. Maeck, J. (2003): Damage Assessment of Civil Engineering Structures by Vibration Monitoring. PhD thesis, Department of Civil Engineering, K.U. Leuven, Belgium
50. Michalewicz, Z. (1996): Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag, Heidelberg
51. Overschee, P. V., Moor, B. D. (1997): Subspace identification for linear systems: Theory - Implementation - Applications. Katholieke Universiteit Leuven, Belgium, Kluwer Academic Publishers, Boston/London/Dordrecht
52. Peeters, B. (2000): System Identification and Damage Detection in Civil Engineering. PhD thesis, Department of Civil Engineering, K.U. Leuven, Belgium, December-2000
53. Petsounis, K. A., and Fassois, S. D. (2001): Parametric time domain methods for the

identification of vibrating structures-a critical comparison and assessment. Mechanical Systems and Signal Processing, 15(6), 1031-1060

54. Pridham, B. A. and Wilson, J. C. (2004): Identification of base-excited structures using output-only parameter estimation. Earthquake Engineering and Structural Dynamics, 33, 133–155
55. Quek, S. T., Wang, W., Koh, C. C. (1999): System identification of linear MDOF structures under ambient excitation. Engineering and Structural Dynamics, 28, 61-77
56. Rustagi, J. S. (1994): Optimization techniques in statistics. Academic Press
57. Salawu, O. S. (1997): Detection of structural damage through changes in frequency: A review. Engineering Structure, 19(9), 718-723
58. Sayed, A. H. (2003): Fundamentals of Adaptive Filtering. John Wiley & Sons
59. Simiu, E., and Scanlan, R. H. (1996): Wind effects on structures. 3rd edition, Wiley-Interscience, New York
60. Sjoberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, B., Hjalmarsson, H., Juditsky, A. (1995): Nonlinear Black-box modeling in system identification: a united overview. Automatica, 31(22), 1691-1724
61. Soize, C., and Lefur, O. (1997): Modal identification of weakly non-linear multidimensional dynamical systems using a stochastic linearization method with random coefficients. Mechanical Systems and Signal Processing, 11(1), 37-49
62. Structural dynamics @ 2000: current status and future directions / edited by D.J. Ewins, D.J. Inman. Research Studies Press LTD, Baldock, Hertfordshire, England
63. Vold, H., and Rocklin, G. F. (1982): The numerical implementation of a multi-input modal estimation method for mini-computers. International Modal Analysis conference Proceedings, Nov. 1982
64. Worden, K., and Tomlinson, G. R. (2001): Nonlinearity in structural dynamics, Detection, Identification and Modeling. IOP Publishing Ltd.
65. Wright, M. H. (1995): Direct search methods: once scorned, now respectable. In "Numerical Analysis" 1995, pages 191 208. Addison Wesley Longman, Harlow, United Kingdom
66. Xu, B., Wu, Z., Chen, G., Yokoyama, K. (2004): Direct identification of structural parameters from dynamic responses with neural networks. Engineering Applications

of Artificial Intelligence, 17, 931-943

67. Yang, J. N., Lei, Y., Pan, S., Huang, N. (2003): System identification of linear structures based on Hilbert–Huang spectral analysis, Part 1: normal modes. Earthquake Engineering and Structural Dynamics, 32, 1443–1467
68. Yang, Q. J., Zhang, P. Q., Li, C. Q., and Wu, X. P. (1994): A system theory approach to multi-input multi-output modal parameters identification methods. Mechanical Systems and Signal Processing, 8(2), 159–174.
69. Yuen, K. V, Beck, J. L., Katafygiotis, L. S. (2002): Probabilistic approach for modal identification using non-stationary noisy response measurements only. Earthquake Engineering and Structural Dynamics, 31, 1007-1023
70. Yuen, K. V., and Katafygiotis, L. S. (2001): Bayesian time-domain approach for modal updating using ambient data. Probabilistic Engineering Mechanics, 16, 219-231
71. Yuen, K. V., and Katafygiotis, L. S. (2002): Bayesian Modal Updating using Complete Input and Incomplete Response Noisy Measurements. Journal of Engineering Mechanics, 128(3), 340-350
72. Yuen, K. V., and Katafygiotis, L. S. (2005): Model updating using noisy response measurements without knowledge of the input spectrum. Earthquake Engineering and Structural Dynamics, 34, 167 187
73. Yun, C. B., Lee, H. J., Lee, C. G. (1997): Sequential prediction error method for structural identification. Journal of Engineering Mechanics, Vol. 123, No. 2, 115-122
74. Zeldin, B. A., Spanos, P. D. (1998): Spectral identification of nonlinear structural systems. Journal of Engineering Mechanics, 124(7), 728–733.
75. Zhang, J., Gao, F., Zhong, Q., Dai, Y. (2002): Spectrum estimation by using of neural network method. Proceedings of the 9th International Conference on Neural Information Processing (ICONIP'O2), Vol. 5

VITA AUCTORIS

Li Li was born in 1973 in Hefei, China. He completed a diploma in construction management at the Anhui University, Hefei, China in 1992 before he worked as a highway and bridge technician in Anhui provincial ministry of highway transportation from 1992 to 1995. He went on to the Xi'an Highway University, Xi'an, China where he obtained a Master of Engineering degree in Bridge Engineering in 1998. Afterwards he had worked as bridge engineer at a consulting company in Beijing, China, from 1998 to 2003.

Presently Li Li is a candidate for Master of Applied Science degree in Civil Engineering with specialization in Structural Engineering at the University of Windsor, Windsor, Ontario, Canada and expects to graduate in October 2005.

Conference presentation:

Li L.: Modal identification using a genetic and Nelder-Mead approach. 5th Annual MOPTA Conference (Modeling and Optimization: Theory and Applications), Windsor, Ontario, July, 2005