Maximum Likelihood Estimation of New State Space Models for Operational Modal Analysis

Tesis Doctoral

Francisco Javier Cara Cañas
Ingeniero de Caminos, Canales y Puertos

2013
Maximum Likelihood Estimation of New State Space Models for Operational Modal Analysis

Tesis Doctoral

Autor: Francisco Javier Cara Cañas
Ingeniero de Caminos, Canales y Puertos

Directores: Enrique Alarcón Álvarez
Dr. Ingeniero de Caminos, Canales y Puertos

Jesús Juan Ruiz
Dr. Ingeniero Industrial

2013
Tribunal nombrado por el Magfco. y Excmo. Sr. Rector de la Universidad Politécnica de Madrid, el día ..... de ............ de 2013.

Presidente: D. ........................................................................................................

Vocal: D. ........................................................................................................

Vocal: D. ........................................................................................................

Vocal: D. ........................................................................................................

Secretario: D. ...................................................................................................

Suplente: D. ....................................................................................................

Suplente: D. ....................................................................................................

Realizado el acto de defensa y lectura de la tesis el día ..... de ............ de 2013 en la E.T.S. Ingenieros Industriales.

CALIFICACIÓN:

EL PRESIDENTE LOS VOCALES

EL SECRETARIO
To my parents, Salvador and Josefina
Acknowledgements

It is very difficult to express in words the gratitude I feel for the people who helped me making this work possible. First of all, I am specially grateful to my supervisors, prof. Enrique Alarcón and prof. Jesús Juan, for their support, encouragement and guidance throughout the course of this work. I thank them for giving me their time and advices, for sharing their scientific knowledge and for giving me all resources needed to accomplish this work.

I thank all my colleagues of the Centro de Modelado en Ingeniería Mecánica, Alberto Fraile, Ignacio del Rey, Lutz Hermanns, Irene Espinosa, Olalla Ortiz, Alfredo Grande, Sonia Fernández, Jaime Vega, Josu Barredo, Javier Fernández, and my colleagues of the Laboratory of Statistics, Camino González, María Jesús Sánchez, Carolina García, José Manuel Mira, Jaime Carpio and Eduardo Caro, for the stimulating and pleasant time I have spent working with them.

I would like to express my gratitude to prof. Guido de Roeck and his research team, from K.U.Leuven, for receiving me in a research stay in 2011, for their hospitality and for their generosity sharing their scientific knowledge and data. The Consejo Social de la Universidad Politécnica de Madrid is gratefully acknowledged for the financial support for this stay.

I would like to thank Dionisio Bernal, Edwin Reynders and Filipe Magalhães, for the time they devoted to evaluate my PhD dissertation, a process necessary to obtain the “International Mention”.

I would also like to thank Pilar Crespo from the Ministry of Fomento in Spain, who gave us the opportunity of working with data recorded at Tablate Bridge, what undoubt-edly constituted the seed of this work.

Last but not least, I am greatly indebted to my family, especially my parents, who have always supported and motivated me in an outstanding manner throughout my studies and life. And finally I would like to thank Cecilia, who has more contributed than she realizes, for her love and support.

Madrid, August 2013.
Abstract

The modal analysis of a structural system consists on computing its vibrational modes. The experimental way to estimate these modes requires to excite the system with a measured or known input and then to measure the system output at different points using sensors. Finally, system inputs and outputs are used to compute the modes of vibration.

When the system refers to large structures like buildings or bridges, the tests have to be performed in situ, so it is not possible to measure system inputs such as wind, traffic, ... Even if a known input is applied, the procedure is usually difficult and expensive, and there are still uncontrolled disturbances acting at the time of the test. These facts led to the idea of computing the modes of vibration using only the measured vibrations and regardless of the inputs that originated them, whether they are ambient vibrations (wind, earthquakes, ...) or operational loads (traffic, human loading, ...). This procedure is usually called Operational Modal Analysis (OMA), and in general consists on to fit a mathematical model to the measured data assuming the unobserved excitations are realizations of a stationary stochastic process (usually white noise processes). Then, the modes of vibration are computed from the estimated model.

The first issue investigated in this thesis is the performance of the Expectation-Maximization (EM) algorithm for the maximum likelihood estimation of the state space model in the field of OMA. The algorithm is described in detail and it is analysed how to apply it to vibration data. After that, it is compared to another well known method, the Stochastic Subspace Identification algorithm.

The maximum likelihood estimate enjoys some optimal properties from a statistical point of view what makes it very attractive in practice, but the most remarkable property of the EM algorithm is that it can be used to address a wide range of situations in OMA. In this work, three additional state space models are proposed and estimated using the EM algorithm:

- The first model is proposed to estimate the modes of vibration when several tests are performed in the same structural system. Instead of analyse record by record and then compute averages, the EM algorithm is extended for the joint estimation of the proposed state space model using all the available data.

- The second state space model is used to estimate the modes of vibration when the number of available sensors is lower than the number of points to be tested. In these cases it is usual to perform several tests changing the position of the sensors from one test to the following (multiple setups of sensors). Here, the proposed state space model and the EM algorithm are used to estimate the modal parameters taking into account the data of all setups.
And last, a state space model is proposed to estimate the modes of vibration in the presence of unmeasured inputs that cannot be modelled as white noise processes. In these cases, the frequency components of the inputs cannot be separated from the eigenfrequencies of the system, and spurious modes are obtained in the identification process. The idea is to measure the response of the structure corresponding to different inputs; then, it is assumed that the parameters common to all the data correspond to the structure (modes of vibration), and the parameters found in a specific test correspond to the input in that test. The problem is solved using the proposed state space model and the EM algorithm.
Resumen

El análisis modal de un sistema estructural consiste en calcular sus modos de vibración. Para estimar estos modos experimentalmente es preciso excitar el sistema con entradas conocidas y registrar las salidas del sistema en diferentes puntos por medio de sensores. Finalmente, los modos de vibración se calculan utilizando las entradas y salidas registradas.

Cuando el sistema es una gran estructura como un puente o un edificio, los experimentos tienen que realizarse in situ, por lo que no es posible registrar entradas al sistema tales como viento, tráfico, … Incluso si se aplica una entrada conocida, el procedimiento suele ser complicado y caro, y todavía están presentes perturbaciones no controladas que excitan el sistema durante el test. Estos hechos han llevado a la idea de calcular los modos de vibración utilizando sólo las vibraciones registradas en la estructura y sin tener en cuenta las cargas que las originan, ya sean cargas ambientales (viento, terremotos, …) o cargas de explotación (tráfico, cargas humanas, …). Este procedimiento se conoce en la literatura especializada como Análisis Modal Operacional, y en general consiste en ajustar un modelo matemático a los datos registrados adoptando la hipótesis de que las excitaciones no conocidas son realizaciones de un proceso estocástico estacionario (generalmente ruido blanco). Posteriormente, los modos de vibración se calculan a partir del modelo estimado.

El primer problema que se ha investigado en esta tesis es la utilización de máxima verosimilitud y el algoritmo EM (Expectation-Maximization) para la estimación del modelo espacio de los estados en el ámbito del Análisis Modal Operacional. El algoritmo se describe en detalle y también se analiza como aplicarlo cuando se dispone de datos de vibraciones de una estructura. A continuación se compara con otro método muy conocido, el método de los Subespacios.

Los estimadores máximo verosímiles presentan una serie de propiedades que los hacen óptimos desde un punto de vista estadístico, pero la propiedad más destacable del algoritmo EM es que puede utilizarse para resolver un amplio abanico de situaciones que se presentan en el Análisis Modal Operacional. En este trabajo se proponen y estiman tres modelos en el espacio de los estados:

- El primer modelo se utiliza para estimar los modos de vibración cuando se dispone de datos correspondientes a varios experimentos realizados en la misma estructura. En lugar de analizar registro a registro y calcular promedios, se utiliza algoritmo EM para la estimación conjunta del modelo propuesto utilizando todos los datos disponibles.

- El segundo modelo en el espacio de los estados propuesto se utiliza para estimar los modos de vibración cuando el número de sensores disponibles es menor que
el número de puntos que se quieren analizar en la estructura. En estos casos es usual realizar varios ensayos cambiando la posición de los sensores de un ensayo a otro (múltiples configuraciones de sensores). En este trabajo se utiliza el algoritmo EM para estimar los parámetros modales teniendo en cuenta los datos de todas las configuraciones.

• Por último, se propone otro modelo en el espacio de los estados para estimar los modos de vibración en la presencia de entradas al sistema que no pueden modelarse como procesos estocásticos de ruido blanco. En estos casos, las frecuencias de las entradas no se pueden separar de las frecuencias del sistema y se obtienen modos espurios en la fase de identificación. La idea es registrar la respuesta de la estructura correspondiente a diferentes entradas; entonces se adopta la hipótesis de que los parámetros comunes a todos los registros corresponden a la estructura (modos de vibración), y los parámetros encontrados en un registro específico corresponden a la entrada en dicho ensayo. El problema se resuelve utilizando el modelo propuesto y el algoritmo EM.
3 Joint analysis of multiple records

3.1 Introduction ................................................................. 37
3.2 State space model for the analysis of multiple records ............. 38
  3.2.1 Classic state space model .......................................... 38
  3.2.2 Proposed joint state space model ............................... 39
3.3 Maximum likelihood estimation of the proposed model ............. 39
  3.3.1 E-step: expectation step ........................................... 41
  3.3.2 M-step: maximization step ....................................... 42
  3.3.3 Overall procedure ................................................ 43
3.4 Starting values for the EM algorithm ............................... 43
3.5 Example 1: simulated data ............................................ 44
3.6 Example 2: Table II Bridge ........................................... 47
3.7 Conclusions ............................................................... 48
3.8 Appendix ................................................................. 49

4 Analysis of multiple setups of sensors ............................... 55

4.1 Introduction ............................................................... 55
4.2 State space model for the analysis of multiple setups of sensors 56
4.3 Maximum likelihood estimation: EM algorithm ..................... 58
  4.3.1 E-step: expectation step ........................................... 59
  4.3.2 M-step: maximization step ....................................... 60
  4.3.3 Overall procedure ................................................ 61
4.4 Starting values for the EM algorithm ............................... 61
4.5 Example 1: simulated data ............................................ 62
  4.5.1 Description of the data ............................................ 62
  4.5.2 Proposed method with permanent-moving sensors ............. 62
  4.5.3 Multi-step approach .............................................. 63
  4.5.4 Proposed method with overlapped sensors ..................... 64
4.6 Example 2: Ninove footbridge ....................................... 66
  4.6.1 Bridge description and ambient vibrations tests ............. 66
  4.6.2 Proposed method .................................................. 67
  4.6.3 Multi-step approach using real measured data ............... 70
4.7 Example 3: gym building ............................................. 72
4.8 Conclusions ............................................................... 75
4.9 Appendix ................................................................. 77

5 Operational modal analysis with unobserved autocorrelated inputs .... 81

5.1 Introduction and problem description ............................... 81
5.2 State space model for OMA with unobserved autocorrelated inputs 83
5.3 Maximum likelihood estimation: EM algorithm ..................... 85
  5.3.1 E-step: expectation step ........................................... 86
  5.3.2 M-step: maximization step ....................................... 87
  5.3.3 Overall procedure ................................................ 88
5.4 Starting values for the EM algorithm ............................... 88
5.5 Example 1: simulated data ............................................ 89
  5.5.1 Description of the data ............................................ 89
C.2.4 Matrix Derivatives . . . . . . . . . . . . . . . . . . . . . . . . . . . 143
C.3 Jensen’s inequality . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 143

Bibliography 151
# List of Figures

2.1 Model of the ASCE benchmark structure ................................. 19
2.2 Modal parameters of the simulated benchmark problem ................ 20
2.3 Power spectral density of simulated accelerations ..................... 20
2.4 Evolution of the eigenvalues of A matrix with the EM algorithm ..... 23
2.5 Boxplots of SSI, EM1 and EM2 results for the 100 simulated cases ... 24
2.6 Test structure built in the University of British Columbia .......... 28
2.7 Stabilization diagram for the experimental data ..................... 29
3.1 Simulated structure ..................................................... 45
3.2 Tablate II Bridge ....................................................... 47
3.3 Permanent sensors and recording unit in Tablate II Bridge .......... 47
3.4 Stabilization diagram for Tablate II Bridge ......................... 49
4.1 Merging $M$ individual partial mode shapes into a global mode shape ... 56
4.2 Estimating the modal parameters processing $M$ setups at the same time ... 57
4.3 Modal parameters using simulated data with permanent-moving sensors ... 64
4.4 Ninove footbridge picture, geometry and setups ..................... 67
4.5 Stabilization diagram for the Ninove footbridge .................... 68
4.6 Identified modes at Ninove footbridge using the proposed method ... 69
4.7 Gym building pictures, geometry and sensor location ................ 75
4.8 Stabilization diagram corresponding to the gym .................... 77
4.9 Identified modes at the gym building ................................ 78
5.1 Inputs for simulated records 1 and 2 ................................ 91
5.2 Stabilization diagram for simulated record 1 ....................... 93
5.3 Stabilization diagram for simulated record 2 ....................... 94
5.4 Stabilization diagram for simulated record 3 ....................... 95
5.5 Stabilization diagram built using the joint state space model ...... 95
5.6 Stabilization diagram for the simulated data using the proposed model ... 96
5.7 Estimation of the inputs in records 1 and 2 .......................... 98
5.8 Sensor position and jumping place in the gym building ............. 99
5.9 Jumping force recorded by the force plate in the gym building ... 100
5.10 Vibration recorded by sensor 8 due to ambient loading and jumping 100
5.11 Stabilization diagrams for the gym using the proposed model ....... 101
5.12 Jumping load estimated using the vibration recorded by sensor number 8 ... 102
A.1 Zero-Order Hold approximation of $\sin x$ ................................ 117
# List of Tables

2.1 Modal parameters identified from one simulated time history response . . . 22  
2.2 Modal parameters identified from 100 simulated time history responses . . 27  
2.3 Resulting modal parameters for the experimental time history response . . 30  

3.1 Modal parameters estimated from multiple records (simulated data) . . . 46  
3.2 Multiple records from Tablate II Bridge . . . . . . . . . . . . . . . . . . . 48  
3.3 Modal parameters estimated from multiple records (Tablate II bridge) . . 50  

4.1 Modal parameters estimated from the simulated data . . . . . . . . . . . 65  
4.2 Modal parameters estimated by the proposed method . . . . . . . . . . . 66  
4.3 Modes estimated by the proposed method at Ninove footbridge . . . . . 70  
4.4 Natural frequencies estimated using a one-step approach in all the setups . 73  
4.5 Mean natural frequencies and mean damping ratios of the merged modes . 74  
4.6 MAC values computed between modes estimated at Ninove footbridge . . 74  
4.7 Points measured in each configuration of the gym building . . . . . . . . . 76  

5.1 Characteristics of the inputs for records 1 and 2 . . . . . . . . . . . . . . 90  
5.2 Modal parameters estimated from simulated data with non-white inputs . 93  
5.3 Modal parameters estimated from simulated data using the joint model . . 94  
5.4 Modal parameters estimated from simulated data using the proposed model 96
Chapter 1

Introduction

1.1 Research context

1.1.1 Modal analysis using measured data

The essential idea of modal analysis is to describe the vibrations of a structural system with simple components, the so-called vibration modes. Such is the importance of vibration modes that we could even say that understanding vibration problems requires understanding the modal properties.

In practice, the modal properties can be estimated from experimental data: measured displacement, velocities or accelerations are generally used [40]. Technical literature differentiates three different approaches [87]:

- **Experimental Modal Analysis (EMA).** The structure is excited by one or several measured dynamic forces, the response of the structure to these forces is recorded, and the modal parameters are extracted from the input and output measurements. Therefore, EMA methods are suitable for laboratory conditions, where the inputs, the outputs and the boundary conditions of the test are totally controlled and measured. In fact, EMA is a well-established and often used approach in mechanical engineering, as documented in [40], [51], [67].

- **Operational Modal Analysis (OMA).** In large structures, like bridges or buildings, it is difficult to apply known inputs. This led to the idea of using the vibrations due to environmental loads such as wind and traffic. Here the assumption that the unmeasured loads are realizations of a stationary stochastic process substitutes for the deterministic knowledge of driving forces. In technical literature, we can find different names for this technique:
  - Ambient modal analysis: the inputs acting in the structure are ambient inputs (wind, traffic, earthquakes, ...), not controlled inputs (hammers, shakers, ...).
  - Output-only modal analysis: the modal parameters are estimated using only the response of the structure.
  - Operational modal analysis or in-operation modal analysis: the vibrations are recorded when the structure is in its normal operation. This is the most used name in the literature and this is the name we are going to use in this work.
In civil engineering, OMA has become the primary modal testing method, and the number of reported case studies is abundant: roadway and railway bridges \cite{14, 50, 65, 66, 68}, footbridges \cite{52, 88}, silos \cite{37}, buildings and special structures \cite{29}, historical buildings \cite{86}, offshore platforms \cite{12}, high-rise buildings \cite{13}, dams \cite{50}, among others.

- Operational Modal Analysis with eXogenous inputs (OMAX). Last years, some researchers have used practical actuators in the modal test of large structures \cite{44, 88}. The amplitude of the artificial (measured) forces can be equal or even lower than the amplitude of the operational forces, and both, operational and artificial forces are included in the identified system model.

Modal parameters are used by engineers in many applications, for example:

- Model updating. The use of experimental data recorded at a structural system is used to calibrate a mathematical model in order to obtain more reliably predictions of its dynamical behaviour \cite{76}.

- Structural health monitoring. Changes in the vibration pattern are indicative of changes in the structure, for example, due to damage \cite{52, 88}.

- Vibration serviceability and control of vibrations. The vibration level of structural systems are measured and predicted for other scenarios. When comfort values are exceeded, specific devices are designed and incorporated to the system in order to reduce the vibrations \cite{38, 57}.

- Load estimation. Vibration measurements are used to estimate the forces acting in the structures \cite{64}.

This explains the increasing interest in developing accurate algorithms to estimate the modal parameters using measured vibration data.

1.1.2 The discrete state space model

Nowadays, performing modal analysis with experimental data typically consists of three basic steps \cite{24}:

1. data collection;
2. system identification;
3. extracting and validating a set of modal parameters.

The first step, corresponding to recording and preprocessing the data, is not going to be treated here (see \cite{9, 77} for details). The next step is to fit a mathematical model to the recorded data (system identification). In the time domain, the simplest model we can use is the ARX (see \cite{11})

\[
y_t + a_1 y_{t-1} + \cdots + a_n y_{t-n} = b_1 u_{t-1} + \cdots + b_m u_{t-m}.
\]

This difference equation is known as the Auto-Regressive with eXogenous inputs (ARX) model, and its application to the identification of structural systems can be found, for example, in \cite{4}. This model deals with a Single Input and a Single Output (SISO) in the
1.1 Research context

system. However, in structural mechanics, system inputs and outputs are recorded simultaneously in different points, what is usually called as Multiple Input-Multiple Output problems (MIMO). Although ARX models can be extended to MIMO problems (VARX models), another interesting approach to study MIMO problems is the state space model. Introduced by Kalman [58] in 1960, it was originally developed in the aerospace industry and quickly spread to many other fields of engineering, and later to economics, medicine, seismology, etc. The increasing popularity of state space modelling can be attributed to the fact that it has a very general form and is able to represent linear, non-linear, time-invariant and time-varying dynamics in a relatively compact form. In 1977 Hart and Yao [47] applied this methodology to system identification in structural dynamics. Since then, the space state model is one of the preferred frameworks for obtaining the vibrational modes of a mechanical structure.

The key property of the state space model in discrete-time is that the system at a time $k$ is fully characterized by the state vector $x_k$. Future states depend only on the current state $x_k$ and the input at time $k$, $u_k$. Thus, $x_k$ includes a complete “memory” of the system, and its time-evolution is governed by the following linear equation (the state equation)

$$x_{k+1} = Ax_k + Bu_k + w_k,$$

where

1. $Ax_k$ is an autoregressive component that links one state to the next;
2. $Bu_k$ is a deterministic component and captures the effect of the input variables;
3. $w_k$ is the input noise process due to unknown disturbances. It is usually modelled as independent random variables with zero mean and variance matrix $Q$.

An important property of the state variables $x_k$ is that they are not observed directly but only certain linear combinations of them are measured. In the case of structural systems, the system behaviour is observed by means of sensors distributed throughout the structure. Calling $y_k$ the vector containing the output measurements provided by the sensors at time $k$, the state space model is completed by the observation equation that relates the observations with states $x_k$ and inputs $u_k$

$$y_k = Cx_k + Du_k + v_k.$$  \(1.2\)

This linear equation also includes a random component $v_k$, the output noise process. It is also modelled as independent random variables with zero mean and variance matrix $R$. For the simplest case, the random processes $w_k$ and $v_k$ can be considered independent.

Equations (1.1) and (1.2) represent the state space model in discrete-time. They depend on a set of parameters: the transformation matrices $A$, $B$, $C$ and $D$, and the variance matrices $Q$ and $R$. In this case, system identification consists on estimating these parameters from the data collected in the measurement stage: the sensors measurements $Y = \{y_1, y_2, \ldots, y_N\}$, and the known inputs $U = \{u_1, u_2, \ldots, u_N\}$. This way of posing the problem is known as parametric system identification.

The model given by (1.1) and (1.2) is the basic state space representation of a system, also called the dynamic linear model. When the input data is not available (like in Operational Modal Analysis), the option is to use the model

$$x_{k+1} = Ax_k + w_k
y_k = Cx_k + v_k,$$
Introduction

where it is assumed that the system is driven by white noise stochastic inputs.

1.1.3 Model estimation: the EM algorithm

Modal parameters are computed from the state space model (essentially from the eigenvalues of matrix $A$). Therefore, the problem of estimating the vibration modes is reduced to estimate the state space model from measured data. There are many solid and proven methods for estimating the state space model from time series data. The available methods can be classified into two well defined categories: optimization methods and subspace methods.

- In the optimization approach, the parameters of the model are chosen so that the difference between the output predicted by the model and the measured output is minimized. The optimization is performed using gradient methods starting from an initial guess and the estimation involves an iterative, numerical search for the best fit. In particular for structural system identification, the procedure is numerically demanding, and local optima are a concern.

  A widely used method in control engineering is the Prediction Error Method (PEM). Ref. [63] is an excellent reference for this procedure and for studying the different models and methods related to the identification of linear dynamic systems. A more general and statistically appealing procedure is the maximum-likelihood estimation. Maximizing the log-likelihood with respect to the parameters falls into two main classes based on either gradients (Newton-Raphson or scoring [15]) or Expectation-Maximization (EM), [92].

- The alternative Subspace methods are based on regression and lead to easily implementable estimators carrying a relatively low computational cost. The most popular are the Canonical Variate Analysis (CCA) [61], the Numerical algorithms For Subspace state space system IDentification (N4SID) [78], and the Multivariate Output Error State sPace (MOESP) algorithm [96]. These algorithms are probably the most used ones in Operational Modal Analysis (see [87]) and they are usually known as the Stochastic Subspace Identification algorithms (SSI).

Although these two approaches to state space model estimation have led to a mature set of algorithms and methods, both methodologies have their limitations. The optimization approach suffers from the complex choice of initial settings and a slow convergence [70], and the subspace methods suffer from the lack of optimality [8, 27]. In this work, we are going to explore the optimization approach based on maximizing the likelihood function using the EM algorithm. This algorithm is a well-known tool for the computation of maximum likelihood estimates. It consists on an E-step (evaluating a conditional Expectation) and later on a M-step (which Maximizes the previous expectation) in an iterative loop until the convergence is reached.

The name EM was given by Dempster, Laird, and Rubin [31], who presented the general formulation of the EM algorithm, established its basic properties and provided many examples and applications. The EM algorithm applied to the mathematical model written in the state space form was mainly developed by Shumway and Stoffer [92, 93]. A review of the applications of the EM algorithm in many different contexts can be found in the monographs of Little and Rubin [62] and McLachlan and Krishnam [70].
The method is mainly used in a variety of incomplete-data problems arising in standard statistical situations such as linear models, contingency tables and loglinear models, random effects models and general variance-components models, time series and stochastic processes, among others. But these algorithms have also been profitably used in engineering, psychometry, econometrics, epidemiology, genetics, astronomy, etc. However, as far as we know, the application of the EM algorithm to maximum likelihood estimation of modal parameters in structures was not used before the present study.

1.2 Objectives

It might be clear from previous sections the importance of accurate algorithms to estimate the estate space model for modal analysis. This Thesis is devoted to explore the maximum likelihood estimation of this model using the EM algorithm. Two main issues are investigated:

- The first issue is the analysis of the EM algorithm in Operational Modal Analysis, and the comparison with other well-known methods.

- The second issue that is treated in this thesis is the development and estimation of new state space models to address some situations that arise in OMA (like multiple setups of sensors, non-white noise inputs,...) and which are difficult to solve using the traditional state space approach. The estimation of these new models using the EM algorithm is investigated.

1.3 Outline and organization of the text

The document is organized as follows:

Chapter 2 presents the EM algorithm and its application to estimate the modes of vibration from simulated and measured data in an OMA context. Although the EM algorithm for the state space model was derived in the 1980’s (see [92]), the application to OMA is challenging: first, because the models estimated in usual time series problems involve low model orders, but in OMA, high model orders are frequent; second, the EM is an iterative algorithm and therefore it needs a starting point; the construction of this initial point is not straightforward, specially when we work with high orders. In this Chapter we propose two alternative to overcome this problem. Finally, we compare the parameters estimated by the EM algorithm with the parameters estimated with the most used method in OMA, the SSI.

When dynamic tests are repeated several times at the same structural system, we obtain data recorded at different moments. We have used the term record for the data recorded in one test (one record may include the measurements of one or several sensors). The data corresponding to all tests have been called multiple records. If the modal parameters are estimated from each record, we have as many estimates for each mode as records, and sometimes certain modes are not estimated in all records. In Chapter 3 we propose to estimate the modal parameters taking into account all the records at the same time (joint estimation). For this, we introduce the joint state space model (equations, properties and estimation using the EM algorithm). Using this model we obtain only one estimate for each mode, and what is more important, this estimate is the most likely value taking into account all measured data.
Chapter 4 presents the estimation of the modal parameters from multiple measurement setups. This situation arises when the number of tested points is bigger than the number of available sensors (for example, because the structure is large, or because estimates are needed in a lot of points). Then, it is usual to repeat measurements but changing the positions of the sensors from one test to the next (sensor setups). Instead of processing each setup separately and then merge the results, we propose a new state space model to estimate the modal parameters of the structure processing all the setups at once. This model is estimated using the EM algorithm.

Another important problem in OMA is the presence of unmeasured autocorrelated inputs. In OMA the inputs are not known by definition, and the option is to assume they can be modelled as white noise processes. This hypothesis is valid in many situations, but when the input is clearly a non-white noise process (for example, with harmonic loads), we obtain spurious modes that correspond to the input and not to the system. Chapter 5 proposes to measure the vibrations of the system when different inputs are applied. Using the appropriate state space model, we can estimate the modal parameters (common to all the measurements) and the characteristics of the inputs (specific to one set of measurements or record). Chapter 5 shows that this model can be estimated using the EM algorithm.

In Chapter 6 the main conclusions are summarized and further lines of research are presented.

Appendix A presents some basic concepts of the state space model. Only the properties of the model used in this text are included.

Appendix B defines the modes of vibration and the relation between them and the state space model. The parametric approach to OMA using the state space model is relatively new, so these equations are mainly included in research works, and sometimes it is hard to find a justification for them ([81] and [87] are valuable references in this sense). In this Appendix the equations for computing the modal parameters from the state space model are derived in detail.

Finally, Appendix C includes some mathematical relationships that are used in the text.
Chapter 2

Operational Modal Analysis using the EM algorithm

2.1 Introduction

In this Chapter we investigate the performance of the Expectation Maximization algorithm for the estimation of the state space model in the context of Operational Modal Analysis. The characteristics of the EM algorithm are well documented (see for instance McLachlan and Krishnam, [70]). It leads in general to simple equations, has the property of increasing the loglikelihood at each iteration until convergence and it derives sensible parameter estimates. Consequently, it is a popular tool to derive maximum likelihood estimation.

In the case of the state space model, the EM algorithm was developed by Shumway and Stoffer in 1982 [92]. Since then, the method has been applied to a variety of problems: estimation of dynamic unobserved component models [97], estimation of covariance parameters in a state space model [60], non-linear system identification [90], speech recognition [33], ... However, this method has not received much attention in operational modal analysis. In this chapter we describe the method in detail and show how to use it in OMA. Finally, we compare the results with a well known method, the SSI.

2.2 Modal analysis with the state space model

2.2.1 State Space Equations

A vibrating structural system can be represented by the discrete-time stochastic state space model (see Appendices A and B):

\[ \begin{align*}
  x_t &= Ax_{t-1} + Bu_{t-1} + w_t \\
  y_t &= Cx_t + Du_t + v_t,
\end{align*} \tag{2.1} \]

where \( t \) denotes the time instant, of a total number \( N \), measured with constant sampling time \( \Delta t \); \( y_t \in \mathbb{R}^{n_o} \) is the measured output vector; \( u_t \in \mathbb{R}^{n_i} \) is the measured input vector; \( x_t \in \mathbb{R}^{n_s} \) is the state vector; \( n_o, n_i \) and \( n_s \) are the number of outputs, inputs and the order of the state vector, respectively; \( A \in \mathbb{R}^{n_s \times n_s} \) is the transition state matrix describing the dynamics of the system; \( B \in \mathbb{R}^{n_s \times n_i} \) is the input matrix; \( C \in \mathbb{R}^{n_o \times n_s} \) is the output matrix, which is describing how the internal states are transferred to the output measurements.
Operational Modal Analysis using the EM algorithm

$y_t$: $D \in \mathbb{R}^{n_o \times n_i}$ is the direct transmission matrix; $w_t \in \mathbb{R}^{n_s}$ is the process noise due to disturbances and modelling discrepancies, while $v_t \in \mathbb{R}^{n_o}$ is the measurement noise due to sensor inaccuracy. In this work, the noise processes $w_t$ and $v_t$ are considered uncorrelated white noise processes with $w_t \sim N(0, Q)$ and $v_t \sim N(0, R)$ (see Section $\text{A.4}$ for details).

We also need an initial condition for the difference equation (2.1a): we shall assume that $x_0$ is a Gaussian random variable of known mean $\bar{x}_0$ and known covariance $P_0$. Further, we shall assume that $x_0$ is independent of $w_t$ and $v_t$ for any $t$.

In the case of OMA, only the outputs of the structural system are measured, while the input sequence $u_t$ remains unmeasured. Therefore, the option is to use the state space model:


given by the state space model (2.2)

\begin{align}
\mathbf{x}_t &= A \mathbf{x}_{t-1} + \mathbf{w}_t \quad (2.2a) \\
\mathbf{y}_t &= C \mathbf{x}_t + \mathbf{v}_t. \quad (2.2b)
\end{align}

In this model, the white noise assumptions for $w_t$ and $v_t$ and the conditions for the initial state $x_0$ are also valid:

\begin{align}
w_t &\sim N(0, Q), \quad v_t \sim N(0, R), \quad x_0 \sim N(\bar{x}_0, P_0). \quad (2.3)
\end{align}

2.2.2 System identification and modal analysis

The system identification problem investigated here can be defined as the determination of the parameters of the state space model (2.2), that is,

\begin{align}
\theta = \{A, C, Q, R, \bar{x}_0, P_0\}, \quad (2.4)
\end{align}

using the output measurements $\{y_1, y_2, \ldots, y_N\}$ available for $N$ time steps.

According to Appendix $\text{B}$, the natural frequencies and modal damping ratios can be retrieved from the eigenvalues of $A$, and the mode shapes can be evaluated using the corresponding eigenvectors and the output matrix $C$. It is usual to express the $j$th eigenvalue of $A$ as (the eigenvalues of $A$ come in complex conjugate pairs and each pair represents one physical vibration mode):

\begin{align}
\lambda_j &= \exp \left( \left( -\zeta_j \omega_j \pm i \omega_j \sqrt{1 - \zeta_j^2} \right) \Delta t \right), \quad (2.5)
\end{align}

where $\Delta t$ is the time step. Therefore

\begin{align}
\omega_j &= \frac{|\ln (\lambda_j)|}{\Delta t}, \quad (2.6) \\
\zeta_j &= -\frac{\text{Real} \left[ \ln (\lambda_j) \right]}{\omega_j \Delta t}. \quad (2.7)
\end{align}

If proportional damping is admitted, $\omega_j$ is the undamped natural frequency and $\zeta_j$ is damping ratio of mode $j$. The $j$th mode shape $\psi_j \in \mathbb{R}^{n_o}$ evaluated at sensor locations can be obtained using the following expression:

\begin{align}
\psi_j &= C v_j, \quad (2.8)
\end{align}

where $v_j$ is the complex eigenvector of $A$ corresponding to the eigenvalue $\lambda_j$. 

\[ y_t; D \in \mathbb{R}^{n_o \times n_i} \text{ is the direct transmission matrix; } w_t \in \mathbb{R}^{n_s} \text{ is the process noise due to disturbances and modelling discrepancies, while } v_t \in \mathbb{R}^{n_o} \text{ is the measurement noise due to sensor inaccuracy. In this work, the noise processes } w_t \text{ and } v_t \text{ are considered uncorrelated white noise processes with } w_t \sim N(0, Q) \text{ and } v_t \sim N(0, R) \text{ (see Section A.4 for details).} \]

\[ \text{We also need an initial condition for the difference equation (2.1a): we shall assume that } x_0 \text{ is a Gaussian random variable of known mean } \bar{x}_0 \text{ and known covariance } P_0. \text{ Further, we shall assume that } x_0 \text{ is independent of } w_t \text{ and } v_t \text{ for any } t. \]

\[ \text{In the case of OMA, only the outputs of the structural system are measured, while the input sequence } u_t \text{ remains unmeasured. Therefore, the option is to use the state space model:} \]

\[ x_t = A x_{t-1} + w_t \quad (2.2a) \]
\[ y_t = C x_t + v_t. \quad (2.2b) \]

\[ \text{In this model, the white noise assumptions for } w_t \text{ and } v_t \text{ and the conditions for the initial state } x_0 \text{ are also valid:} \]
\[ w_t \sim N(0, Q), \quad v_t \sim N(0, R), \quad x_0 \sim N(\bar{x}_0, P_0). \quad (2.3) \]

\[ 2.2.2 \text{ System identification and modal analysis} \]

\[ \text{The system identification problem investigated here can be defined as the determination of the parameters of the state space model (2.2), that is,} \]
\[ \theta = \{A, C, Q, R, \bar{x}_0, P_0\}, \quad (2.4) \]

\[ \text{using the output measurements } \{y_1, y_2, \ldots, y_N\} \text{ available for } N \text{ time steps.} \]

\[ \text{According to Appendix B, the natural frequencies and modal damping ratios can be retrieved from the eigenvalues of } A, \text{ and the mode shapes can be evaluated using the corresponding eigenvectors and the output matrix } C. \text{ It is usual to express the } j \text{th eigenvalue of } A \text{ as (the eigenvalues of } A \text{ come in complex conjugate pairs and each pair represents one physical vibration mode):} \]
\[ \lambda_j = \exp \left( \left( -\zeta_j \omega_j \pm i \omega_j \sqrt{1 - \zeta_j^2} \right) \Delta t \right), \quad (2.5) \]

\[ \text{where } \Delta t \text{ is the time step. Therefore} \]
\[ \omega_j = \frac{|\ln (\lambda_j)|}{\Delta t}, \quad (2.6) \]
\[ \zeta_j = -\frac{\text{Real} \left[ \ln (\lambda_j) \right]}{\omega_j \Delta t}. \quad (2.7) \]

\[ \text{If proportional damping is admitted, } \omega_j \text{ is the undamped natural frequency and } \zeta_j \text{ is damping ratio of mode } j. \text{ The } j \text{th mode shape } \psi_j \in \mathbb{R}^{n_o} \text{ evaluated at sensor locations can be obtained using the following expression:} \]
\[ \psi_j = C v_j, \quad (2.8) \]

\[ \text{where } v_j \text{ is the complex eigenvector of } A \text{ corresponding to the eigenvalue } \lambda_j. \]
2.3 Likelihood function for Gaussian state space models

Given $N$ measurements of the outputs $Y_N = \{y_1, y_2, \ldots, y_N\}$, the likelihood is given by

$$L_{Y_N}(\theta) = f_\theta(y_1, y_2, \ldots, y_N), \quad (2.9)$$

where $\theta$ are the parameters of the state space model (2.2), that is, Eq. (2.3), and $f_\theta(\cdot)$ denotes a generic joint density function with parameters represented by $\theta$. We can factor (2.9) as

$$f_\theta(y_1, \ldots, y_N) = f_\theta(y_N|Y_{N-1})f_\theta(Y_{N-1}) = \cdots = f_\theta(y_N|Y_{N-1}) \cdots f_\theta(y_2|Y_1) \cdot f_\theta(y_1|Y_0),$$

where $f_\theta(y_1|Y_0) = f_\theta(y_1)$. Then, Equation (2.10) becomes

$$L_{Y_N}(\theta) = \prod_{t=1}^{N} f_\theta(y_t|Y_{t-1}). \quad (2.10)$$

For model (2.2), $y_t|Y_{t-1}$ is normal distributed, with mean and variance

$$E(y_t|Y_{t-1}) = E(Cx_t + v_t|Y_{t-1}) = CE(x_t|Y_{t-1}) + E(v_t|Y_{t-1}) = Cx_t^{t-1},$$

$$\text{Var}(y_t|Y_{t-1}) = \text{Var}(Cx_t + v_t|Y_{t-1}) = C\text{Var}(x_t|Y_{t-1})C^T + \text{Var}(v_t|Y_{t-1}) = CP_t^{t-1}C^T + R,$$

where $E(x_t|Y_{t-1}) = x_t^{t-1}$ and $\text{Var}(x_t|Y_{t-1}) = P_t^{t-1}$ (see Section A.34 for details). Then

$$f_\theta(y_t|Y_{t-1}) = \frac{\exp \left( -\frac{1}{2} (y_t - E(y_t|Y_{t-1}))^T \text{Var}(y_t|Y_{t-1})^{-1} (y_t - E(y_t|Y_{t-1})) \right)}{(2\pi)^{\frac{y_t}{2}} |\text{Var}(y_t|Y_{t-1})|^{\frac{1}{2}}}, \quad (2.11)$$

An alternative way of writing this formula is using the innovations $\epsilon_t$ defined in Eq. (A.34) and (A.55)

$$\epsilon_t = y_t - Cx_t^{t-1}, \quad \epsilon_t \sim N(0, \Sigma_t), \quad \Sigma_t = CP_t^{t-1}C^T + R.$$

Substituting in (2.11) we have

$$f_\theta(y_t|Y_{t-1}) = \frac{1}{(2\pi)^{\frac{y_t}{2}} |\Sigma_t|^{\frac{1}{2}}} \exp \left( -\frac{1}{2} \epsilon_t^T \Sigma_t^{-1} \epsilon_t \right). \quad (2.12)$$

Thus, the likelihood (2.10) is computed as

$$L_{Y_N}(\theta) = \frac{1}{(2\pi)^{\frac{N+y}{2}} \prod_{t=1}^{N} |\Sigma_t(\theta)|^{\frac{1}{2}}} \exp \left( \sum_{t=1}^{N} -\frac{1}{2} \epsilon_t(\theta)^T \Sigma_t^{-1}(\theta) \epsilon_t(\theta) \right),$$

where it has been emphasized the dependence of the innovations on the vector $\theta$. In practice we generally work with the logarithm of the likelihood:

$$\log L_{Y_N}(\theta) = l_{Y_N}(\theta) = -\frac{n_\theta N}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{N} \left( \log |\Sigma_t(\theta)| + \epsilon_t(\theta)^T \Sigma_t(\theta)^{-1} \epsilon_t(\theta) \right). \quad (2.13)$$
The quantities $\epsilon_t$ and $\Sigma_t$ are calculated routinely by the Kalman filter (Property \ref{property:kalman}), so $\log L_{Y_N}$ is easily computed from the Kalman filter output. This representation of the likelihood was first given by \cite{91}. Harvey \cite{48} refers to it as the prediction error decomposition.

The method of maximum likelihood estimates the parameters $\hat{\theta}$ by maximizing the logarithm of the likelihood, $\log L_{Y_N}(\theta)$. The usual procedure is to use a Newton-Raphson algorithm (see \cite{45} for details). But the method we have chosen to maximize the likelihood, Eq. \ref{eq:likelihood}, is the Expectation-Maximization algorithm. The method is described in the following section.

### 2.4 Maximum likelihood estimation using the EM algorithm

The basic idea is that if the states, $X_N = \{x_0, x_1, x_2, ..., x_N\}$, could be observed in addition to $Y_N = \{y_1, y_2, ..., y_N\}$, the complete data could be considered, with joint density function $f_\theta(X_N, Y_N)$. Then, the likelihood would be computed as

$$L_{X_N, Y_N}(\theta) = f_\theta(X_N, Y_N) \Rightarrow l_{X_N, Y_N}(\theta) = \log (L_{X_N, Y_N}(\theta)) .$$

But the states $X_N$ are not known in the state space model. Let us assume that we know a trial value for the parameters, $\theta_j$. The method proposes to compute the conditional expectation $E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j]$ and to maximize $E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j]$ with respect to $\theta$, what leads to a revised estimate $\theta_{j+1}$ with

$$E[l_{X_N, Y_N}(\theta_{j+1})|Y_N, \theta_j] > E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j]. \quad \text{(2.14)}$$

Then, $\theta_{j+1}$ is taken as a new trial value of $\theta$ and the process is repeated until adequate convergence is achieved \cite{98}. This process is known as the Expectation-Maximization algorithm.

The EM algorithm maximizes the conditional expectation $E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j]$, but the objective is to maximize the likelihood, Eq. \ref{eq:likelihood}. We are going to prove now that to maximize $E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j]$ is equivalent to maximize Eq. \ref{eq:likelihood}, consider the states, $X_N = \{x_0, x_1, x_2, ..., x_N\}$, and the observations, $Y_N = \{y_1, y_2, ..., y_N\}$, with the joint density function

$$f_\theta(X_N, Y_N) = f_\theta(X_N|Y_N) \cdot f_\theta(Y_N).$$

The likelihood is then

$$L_{X_N, Y_N}(\theta) = L_{X_N|Y_N}(\theta) \cdot L_{Y_N}(\theta),$$

and the logarithm of the likelihood

$$\log L_{X_N, Y_N}(\theta) = \log L_{X_N|Y_N}(\theta) + \log L_{Y_N}(\theta),$$

$$l_{X_N, Y_N}(\theta) = l_{X_N|Y_N}(\theta) + l_{Y_N}(\theta).$$

But $l_{X_N, Y_N}(\theta)$ and $l_{X_N|Y_N}(\theta)$ are functions of the states $X_N$, which are not observed. Given a trial value for the parameters, $\theta_j$, we can compute the expected value of the above equation

$$E[l_{Y_N}(\theta)|Y_N, \theta_j] = E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j] - E[l_{X_N|Y_N}(\theta)|Y_N, \theta_j]. \quad \text{(2.15)}$$

Let $\theta_{j+1}$ be the parameters estimated by the EM algorithm. Then

$$E[l_{Y_N}(\theta_{j+1})|Y_N, \theta_j] = E[l_{X_N, Y_N}(\theta_{j+1})|Y_N, \theta_j] - E[l_{X_N|Y_N}(\theta_{j+1})|Y_N, \theta_j],$$

$$E[l_{X_N, Y_N}(\theta_{j+1})|Y_N, \theta_j] > E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j].$$
2.4 Maximum likelihood estimation using the EM algorithm

\[ l_{Y_N}(\theta_j) = E[l_{X_N,Y_N}(\theta_j)|Y_N, \theta_j] - E[l_{X_N|Y_N}(\theta_j)|Y_N], \]

where it has been considered \( E[l_{Y_N}(\theta_j)|Y_N, \theta_j] = l_{Y_N}(\theta_j). \) Subtracting these equations

\[ E[l_{Y_N}(\theta_{j+1})|Y_N, \theta_j] - l_{Y_N}(\theta_j) = \left( E[l_{X_N,Y_N}(\theta_{j+1})|Y_N, \theta_j] - E[l_{X_N|Y_N}(\theta_j)|Y_N, \theta_j] \right) + \right) + \left( E[l_{X_N|Y_N}(\theta_j)|Y_N, \theta_j] - E[l_{X_N|Y_N}(\theta_{j+1})|Y_N, \theta_j] \right). \]

According to Eq. (2.14)

\[ E[l_{X_N,Y_N}(\theta_{j+1})|Y_N, \theta_j] - E[l_{X_N,Y_N}(\theta_j)|Y_N, \theta_j] > 0. \]

On the other hand, taking into account the Kullback-Leibler divergence (see Eq. C.30)

\[ E[l_{X_N|Y_N}(\theta_j)|Y_N, \theta_j] - E[l_{X_N|Y_N}(\theta_{j+1})|Y_N, \theta_j] = E[l_{X_N|Y_N}(\theta_j) - l_{X_N|Y_N}(\theta_{j+1})|Y_N, \theta_j] = \int_{-\infty}^{+\infty} \log \left( \frac{f_\theta(X_N|Y_N)}{g_\theta(X_N|Y_N)} \right) f_\theta(X_N|Y_N) d(X_N|Y_N) \geq 0. \]

So it is verified

\[ E[l_{Y_N}(\theta_{j+1})|Y_N, \theta_j] - l_{Y_N}(\theta_j) \geq 0. \]

In conclusion, to maximize \( E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] \) with respect to \( \theta \) is equivalent to maximize \( \text{Eq. 2.13} \), and vice versa.

Summarizing, the EM algorithm is an iterative method to maximize the log-likelihood \( \text{Eq. 2.13} \) given by the following steps

1. The Expectation step is to compute \( E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] \).
2. The Maximization step consists on maximizing \( E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] \) obtaining \( \theta_{j+1} \).
3. Repeat steps 1 and 2 until \( \theta_{j+1} \) and \( \theta_j \) are closed to each other.

2.4.1 Expectation step

First we are going to derive \( l_{X_N,Y_N}(\theta) \) assuming both, the observations \( Y_N \) and the states \( X_N \), are known. Letting \( f_\theta(\cdot) \) denote a generic density function with parameters represented by \( \theta \), we can write

\[ f_\theta(X_N, Y_N) = f_\theta(x_0, x_1, \ldots, x_N, y_1, y_2, \ldots, y_N) = \]

\[ = f_\theta(x_0) f_\theta(x_1|x_0) f_\theta(x_2|x_0, x_1) \cdots f_\theta(y_1|x_0, \ldots, x_N) \cdots f_\theta(y_N|x_0, \ldots, x_N, y_1, \ldots, y_{N-1}). \]

On the other hand

\[ f_\theta(x_t|x_{t-1}, x_{t-2}, \ldots, x_0) = f_\theta(x_t|x_{t-1}) = f_w(x_t - Ax_{t-1}), \quad (2.16) \]

where \( f_w(\cdot) \) denotes the \( n_x \)-variate normal density with mean zero and covariance matrix \( Q \). This equation means that the process is Markovian, linear and Gaussian. On the other hand

\[ f_\theta(y_t|x_t, Y_{t-1}) = f_\theta(y_t|x_t) = f_v(y_t - Cx_t), \quad (2.17) \]
where \( f_\sigma(\cdot) \) denotes the \( n_\sigma \)-variate normal density with mean zero and covariance matrix \( R \). This equation means that the observations are conditionally independent given the states, and the observations are linear and Gaussian. Then

\[
f_\theta(X_N, Y_N) = f_{\bar{x}_0, P_0}(x_0) \prod_{t=1}^{N} f_{A,Q}(x_t - A x_{t-1}) \prod_{t=1}^{N} f_{C,R}(y_t - C x_t). \tag{2.18}
\]

where, under Gaussian assumption

\[
f_{\bar{x}_0, P_0}(x_0) = \frac{1}{(2\pi)^{n_0/2}|P_0|^{1/2}} \exp \left( -\frac{1}{2} (x_0 - \bar{x}_0)^T P_0^{-1} (x_0 - \bar{x}_0) \right),
\]

\[
f_{A,Q}(x_t - A x_{t-1}) = \frac{1}{(2\pi)^{n_0/2}|Q|^{1/2}} \exp \left( -\frac{1}{2} (x_t - A x_{t-1})^T Q^{-1} (x_t - A x_{t-1}) \right),
\]

\[
f_{C,R}(y_t - C x_t) = \frac{1}{(2\pi)^{n_0/2}|R|^{1/2}} \exp \left( -\frac{1}{2} (y_t - C x_t)^T R^{-1} (y_t - C x_t) \right).
\]

Therefore, the complete likelihood \( L_{X_N,Y_N}(\theta) \) and the log-likelihood are defined by

\[
L_{X_N,Y_N}(\theta) = f_\theta(X_N, Y_N), \tag{2.19}
\]

\[
l_{X_N,Y_N}(\theta) = \log L_{X_N,Y_N}(\theta) = -\frac{1}{2} [l_1(\bar{x}_0, P_0) + l_2(A, Q) + l_3(C, R)], \tag{2.20}
\]

where, ignoring constants,

\[
l_1(\bar{x}_0, P_0) = \log |P_0| + (x_0 - \bar{x}_0)^T P_0^{-1} (x_0 - \bar{x}_0), \tag{2.21}
\]

\[
l_2(A, Q) = N \log |Q| + \sum_{t=1}^{N} (x_t - A x_{t-1})^T Q^{-1} (x_t - A x_{t-1}), \tag{2.22}
\]

\[
l_3(C, R) = N \log |R| + \sum_{t=1}^{N} (y_t - C x_t)^T R^{-1} (y_t - C x_t). \tag{2.23}
\]

The method proposes to compute the expected value of Eq. \( \text{(2.20)} \). The following property shows the obtained expressions:

**Property 2.1.** Given \( Y_N = \{y_1, y_2, \ldots, y_N\} \) and a value for the parameters \( \theta_j \), the expected value of Eq. \( \text{(2.20)} \) is

\[
E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] = E[l_1(\bar{x}_0, P_0)|Y_N, \theta_j] + E[l_2(A, Q)|Y_N, \theta_j] + E[l_3(C, R)|Y_N, \theta_j],
\]

with

\[
E[l_1(\bar{x}_0, P_0)|Y_N, \theta_j] = \log |P_0| + \text{tr} \left( P_0^{-1} \left[ P_0^N + (x_0^N - \bar{x}_0)(x_0^N - \bar{x}_0)^T \right] \right), \tag{2.24}
\]

\[
E[l_2(A, Q)|Y_N, \theta_j] = N \log |Q| + \text{tr} \left( Q^{-1} \left[ S_{xx} - S_{xb}A^T - AS_{bx} + AS_{bx}A^T \right] \right), \tag{2.25}
\]

\[
E[l_3(C, R)|Y_N, \theta_j] = N \log |R| + \text{tr} \left( R^{-1} \left[ S_{yy} - S_{yx}C^T - CS_{yx} + CS_{xx}C^T \right] \right), \tag{2.26}
\]

where \( \text{tr}(\square) \) is the trace of the matrix, and

\[
S_{xx} = \sum_{t=1}^{N} \left( P_t^N + x_t^N (x_t^N)^T \right), \tag{2.27}
\]
2.4 Maximum likelihood estimation using the EM algorithm

\[ S_{bb} = \sum_{t=1}^{N} (P_{t-1}^N + x_{t-1}^N (x_{t-1}^N)^T), \quad (2.28) \]

\[ S_{xb} = \sum_{t=1}^{N} (P_{t-1}^N + x_t^N (x_{t-1}^N)^T), \quad S_{bx} = S_{xb}^T, \quad (2.29) \]

\[ S_{yx} = \sum_{t=1}^{N} (y_t (x_t^N)^T), \quad S_{xy} = S_{yx}^T, \quad (2.30) \]

\[ S_{yy} = \sum_{t=1}^{N} (y_t y_t^T). \quad (2.31) \]

*Proof.* The proof of this property and the meaning of the symbols are given in Section 2.10.

### 2.4.2 Maximization step

Maximizing \( E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] \) with respect to the parameters \( \theta \), constitutes the M-step. This is the strong point of the EM algorithm because the maximum values, obtained equating to zero the corresponding derivatives of the expectations \( (2.24)-(2.26) \), are obtained from explicit formulas.

**Property 2.2.** The maximum of \( E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] \) is attained at

\[ \bar{x}_0 = x_0^N, \quad (2.32) \]

\[ P_0 = P_0^N, \quad (2.33) \]

\[ A = S_{xb}S_{bb}^{-1}, \quad (2.34) \]

\[ Q = \frac{1}{N} (S_{xx} - S_{xb}A^T - AS_{bx} + AS_{xb}A^T). \quad (2.35) \]

\[ C = S_{yx}S_{xx}^{-1}, \quad (2.36) \]

\[ R = \frac{1}{N} (S_{yy} - S_{yx}C^T - CS_{xy} + CS_{yx}C^T). \quad (2.37) \]

*Proof.* The proof of this property is given in Section 2.10.

### 2.4.3 EM algorithm

Steps E and M have to be repeated in an iterative way until the likelihood is maximized. The overall method can be summarized as follows:

- Initialize the procedure \((j = 0)\) selecting starting values for the parameters \( \theta_0 \) and a stop tolerance \( \delta_{adm} \).
- Repeat
  1. Perform the E-Step. Apply the Kalman filter (Properties A.4, A.5 and A.6) to obtain the expected values \( x_t^N, P_t^N \), and \( P_{t,t-1}^N \) with \( \theta_j \) as data. Use them to compute the matrices \( S_{xb}, S_{bb} \) and \( S_{xx} \) given by \((2.27)-(2.31)\).
2. Perform the M-Step. Compute the updated value of the parameters $\theta_{j+1}$ using (2.32)-(2.37).

3. Compute the likelihood $l_Y(\theta_{j+1})$ with Equation (2.13).

4. Compute the actual tolerance

$$\delta = \left| \frac{l_Y(\theta_{j+1}) - l_Y(\theta_j)}{l_Y(\theta_j)} \right|$$

(a) If $\delta > \delta_{adm}$, perform a new iteration with $\theta_{j+1}$ as the value of the parameters.

(b) If $\delta \leq \delta_{adm}$, stop the iterations. The estimate is $\theta_{j+1}$.

2.5 Choosing starting values for the EM algorithm

The initial step for the EM algorithm is to choose a starting value for the parameters $\theta$. This is a crucial step because, like in other iterative procedures, depending on the starting point, a local maximum can be reached instead of the global one (see for example [10], [59]). In this section we present the approaches that we have considered:

1. Using other estimation methods. A natural choice is to begin with estimates obtained by other identification methods. There are several techniques to realize system identification with the state space model (see for example [63]), but among them, we have chosen SSI because it is a well known method in the field of structural system identification. This method is based on the solution of the stochastic realization problem [2] and identifies state space models from (input and) output data by applying robust numerical techniques such as QR factorization, SVD and least squares. A complete overview of data-driven subspace identification (both deterministic and stochastic) is provided in [79].

As we shall discuss in the numerical simulations, the likelihood of the solution given by SSI can be improved using the EM algorithm. So the SSI method provides a good starting point for the EM, or from another point of view, the EM can be used to refine the solution given by SSI. This is not new, and for example, Ljung ([63], section 7.3) propose to improve the quality of SSI by using the estimated model as an initial estimate for the prediction error method (PEM).

However, using only one starting point could lead to a sub-optimal solution: if the solution given by SSI method is near to a local maximum of the likelihood function, the EM algorithm may be not able to leave this area.

2. Random initialization. This is probably the most employed way of initializing the EM algorithm. Matrices $A, C, Q, R$ are built using random numbers. Especial attention must be paid to matrices $Q$ and $R$: they are symmetrical definite positive matrices.

Random initialization is simple to apply in low-dimensional problems. However, we have found some difficulties when using completely random starting values in state space models of high order (what is our case), leading to unstable systems and inaccurate solutions in some cases. Moreover, with random values we do not take advantage of the available information of the problem: measured outputs, the range of frequencies analysed, typical values for the damping ratios, and so on.
For this reason, we propose in this work a procedure to build random starting values that overcomes these cited problems. Our technique use random frequencies, random damping ratios and random mode shapes as the random parameters required to build the initial values. The next section is devoted to define how to build this random initial values for the EM algorithm.

2.5.1 Procedure to build random starting values

In this section we describe a procedure to build random initial values for the parameters \( \theta_0 = (\tilde{A}, \tilde{C}, \tilde{Q}, \tilde{R}, \tilde{x}_0, \tilde{P}_0) \) in the context of operational modal analysis (\( \square \) means random initial values):

1. Given the order \( n_s \) for the state space model \( (2.2) \), generate \( n_s/2 \) random values for the natural frequencies \( \tilde{\omega}_j \), \( n_s/2 \) random values for the damping ratios \( \tilde{\zeta}_j \), and a random full matrix \( \tilde{\Phi} \in \mathbb{R}^{n_s \times n_s} \). We must take into account that admissible values for natural frequencies go from 0 to Nyquist frequency, and the damping ratios are comprised between 0 and 1 (this ensures the generated random starting point lies inside the unit circle). In essence, we are defining an imaginary structure with modal parameters equal to \( \tilde{\omega}_j, \tilde{\zeta}_j \) and \( \tilde{\Phi} \) (a proportionally damped structure).

2. Using the above random values, build the matrices \( \tilde{A} \) and \( \tilde{C} \) as

\[
\tilde{A} = \exp \left( \Delta t \begin{bmatrix} 0 & I \\ -\tilde{\Omega}^2 & -2\tilde{\Omega} \tilde{Z} \end{bmatrix} \right), \quad \tilde{C} = C_a \tilde{\Phi} \begin{bmatrix} -\tilde{\Omega}^2 & -2\tilde{\Omega} \tilde{Z} \end{bmatrix} \tag{2.39}
\]

where \( \tilde{\Omega} \) and \( \tilde{Z} \) are diagonal matrices formed by the random natural frequencies and the random damping ratios, \( \tilde{\Omega}_{ij} = \tilde{\omega}_j \delta_{ij} \) and \( \tilde{Z}_{ij} = \tilde{\zeta}_j \delta_{ij} \) (\( \delta_{ij} \) being the Kronecker delta).

\( C_a \) is the measurement location matrix corresponding to the acceleration responses of the structural system; we take \( C_a = I_{n_o \times n_s^2} \) (a \( n_o \times n_s^2 \) matrix with ones on the diagonal and zeros elsewhere).

The justification of Eqs. \( (2.39) \) is the special form that the state vector and matrices \( A \) and \( C \) have for a structural system with mass matrix \( M \in \mathbb{R}^{n_s^2 \times n_s^2} \), stiffness matrix \( K \in \mathbb{R}^{n_s^2 \times n_s^2} \) and viscous damping matrix \( C \in \mathbb{R}^{n_s^2 \times n_s^2} \) (see Prop. B.7):

\[
x(t) = \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix}; \quad A_e = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix}; \quad C_e = \begin{bmatrix} -C_aM^{-1}K & -C_aM^{-1}C \end{bmatrix};
\]

\( q(t) \in \mathbb{R}^{n_s \times 1} \) is the vector containing the solution for the considered DOFs (the dots mean derivative with respect to time). Defining a new state vector \( z(t) \) by mean of the similarity transformation \( T_1 \)

\[
x(t) = T_1z(t), \quad T_1 = \begin{bmatrix} \Phi & 0 \\ 0 & \Phi \end{bmatrix}, \tag{2.40}
\]

where \( \Phi \) is the eigenvector matrix of \( M^{-1}K \) matrix, we obtain

\[
A_{e1} = T_1^{-1}A_eT_1, \quad C_{e1} = C_eT_1 \Rightarrow
\]

\[
A_{e1} = \begin{bmatrix} 0 & I \\ -\Phi^{-1}M^{-1}K\Phi & -\Phi^{-1}M^{-1}C\Phi \end{bmatrix},
\]
\[ C_{i1} = C_0 \Phi [-\Phi^{-1} \mathcal{M}^{-1} \mathcal{K} \Phi - \Phi^{-1} \mathcal{M}^{-1} \mathcal{C} \Phi]. \]

(see Section 4 for the properties of the similarity transformations). Taking into account the orthogonal properties of \( \Phi \) (see Eqs. (3.32) and (3.33)) and the matrix inverse properties

\[
\Phi^{-1} \mathcal{M}^{-1} \mathcal{K} \Phi = \Phi^{-1} \mathcal{M}^{-1} (\Phi^T)^{-1} \Phi^T \mathcal{K} \Phi = (\mathcal{M} \Phi)^{-1} (\Phi^T)^{-1} \Phi^T \mathcal{K} \Phi = (\Phi^T \mathcal{M} \Phi)^{-1} \Phi^T \mathcal{K} \Phi = \mathcal{M}_m^{-1} \mathcal{K}_m = \Omega^2,
\]

where \( \Omega \) is a diagonal matrix formed with the natural frequencies, \([\Omega]_{ij} = \omega_j \delta_{ij}\). Applying the same procedure to the other component of the matrices, and assuming proportional damping

\[
\Phi^{-1} \mathcal{M}^{-1} \mathcal{C} \Phi = 2\Omega Z,
\]

where \( Z \) is a diagonal matrix formed with the damping ratios of each vibrational mode, \([Z]_{ij} = \zeta_j \delta_{ij}\). Hence, substituting both conditions into \( A_{c1} \) and \( C_{c1} \) results

\[
A_{c1} = \begin{bmatrix} 0 & I \\ -\Omega^2 & -2\Omega Z \end{bmatrix}, \quad C_{c1} = C_0 \Phi \begin{bmatrix} -\Omega^2 & -2\Omega Z \end{bmatrix}.
\]

The last step is to find the discrete version of \( A_{c1} \) and \( C_{c1} \) using Eqs. (A.13) and (A.14), what gives (2.39).

3. Matrices \( \hat{R} \) and \( \hat{Q} \) can be calculated after compute the discrete states associated with matrices \( \hat{A} \) and \( \hat{C} \) (Equation (2.39)). Taking into account Eq. (2.40)

\[
\tilde{z}_t = \begin{bmatrix} \Phi^{-1} & 0 \\ 0 & \Phi^{-1} \end{bmatrix} \begin{bmatrix} I \bar{y}_t \\ I \bar{v}_t \end{bmatrix}, \quad (4.1)
\]

where \( y_t \) and \( v_t \) are the discrete displacement and velocity respectively, and can be obtained from the measured accelerations \( y_t \) using numerical integration (Simpson’s rule, for example). \( I_0 \) is a \( \frac{N}{2} \times n_o \) matrix with ones on the diagonal and zeros elsewhere, and it is used for matrix order compatibility; we take \( I_0 = C^T_0 \).

(a) The value of \( \hat{R} \) is obtained computing the covariance matrix of the noise vector calculated from the observation equation (2.2b)

\[
y_t = C z_t + v_t \Rightarrow \tilde{v}_t = y_t - \hat{C} \tilde{z}_t \Rightarrow \hat{R} = \frac{1}{N} \sum_{t=1}^{N} \tilde{v}_t \tilde{v}_t^T. \quad (4.2)
\]

(b) In the same way, the value of \( \hat{Q} \) is obtained using the noise calculated from the state equation (2.2a)

\[
z_{t+1} = A z_t + w_t \Rightarrow \tilde{w}_t = \tilde{z}_{t+1} - \hat{A} \tilde{z}_t \Rightarrow \hat{Q} = \frac{1}{N} \sum_{t=1}^{N} \tilde{w}_t \tilde{w}_t^T. \quad (4.3)
\]

4. For numerical reasons, it is preferred that the elements of matrix \( Q \) were close to one. Therefore, we use the state transformation \( \tilde{z}_t = T \hat{a}_t \), being \( T \) a diagonal matrix with diagonal elements equal to the standard deviation \( s_i \) of the \( i \)th component of the noise \( \hat{w}_t \). Taking into account this last transformation, the random initial values for the EM algorithm are computed by

\[
\tilde{A}_0 = T \hat{A}^{-1}, \quad \tilde{C}_0 = \hat{C} T^{-1}, \quad \tilde{Q}_0 = T \hat{Q} T^T, \quad \tilde{R}_0 = \hat{R}. \quad (4.4)
\]
5. $\tilde{x}_0$, $\tilde{\Sigma}_0$ will be zero. The steady solution of the dynamic system is taken as initial condition.

Other random initial values can be used. However, most of tests we have made led to singular matrices in some steps of the iterations and the algorithm was interrupted unexpectedly (especially when we used experimental data). This means that not any matrices $A$, $C$, $Q$ and $R$ can be used to start the algorithm because they are all interrelated by means of equations (2.2a) and (2.2b).

The procedure explained above is easily implemented with low computational cost and gets very good results.

2.6 Proposed methods for Operational Modal Analysis with the EM

In essence, we use two different procedures to perform Operational Modal Analysis of structures with the EM algorithm. The first one is the combination of the SSI method and the EM algorithm. This is probably the simplest approach and can be summarized as:

1. Estimate matrices $A$, $C$, $Q$ and $R$ using the SSI algorithm.
2. Take these matrices like the starting point for the EM algorithm ($\tilde{x}_0$ and $P_0$ can be taken equal to zero).
3. Run the EM algorithm until convergence.
4. Compute the modal parameters from this solution.

In the presence of multiple local minima, using only one starting point for the EM algorithm can lead to sub-optimal solutions because the solution can depend on the initial point (this is a drawback of all iterative methods, and of course, of the EM too). Therefore, the EM solution would be conditioned by SSI solution. This difficulty can be overcome by generating different starting values with the procedure described in 2.5.1. The problem arising now is that we need to choose a final solution between all the solutions obtained with each initial position. A first option is to choose the solution with the highest log-likelihood (equation (2.13)). A similar strategy has been used for the EM algorithm in mixture models ([10]) and can be described as

1. Generate $p$ initial positions (section 2.5.1).
2. Run the EM algorithm at each initial position with a fixed number of iterations.
3. Select the solution providing best likelihood among the $p$ trials.
4. Compute the modal parameters from this solution.

The solution obtained this way will have the highest likelihood but we do not take advantage of the information given by the rest of starting points. Moreover, due to EM algorithm always increase the likelihood at each iteration, the final likelihoods computed for each initial value are quite similar. For these reasons we propose an alternative strategy:
1. Generate $p$ initial positions (section 2.5.1).

2. Run the EM algorithm at each initial position (so we have $p$ different solutions).

3. Compute the modal parameters from each solution.

4. Select the parameters that are present in more solutions.

The latter is not a trivial step. We recommend to use the following criteria to decide if two estimates correspond to the same mode:

\[
\left| \frac{\omega_{pi} - \omega_{qj}}{\omega_{pi}} \right| \leq \varepsilon_{\omega}, \tag{2.45}
\]

\[
|\zeta_{pi} - \zeta_{qj}| \leq \varepsilon_{\zeta}, \tag{2.46}
\]

\[
1 - MAC(\psi_{pi}, \psi_{qj}) \leq \varepsilon_{MAC}, \tag{2.47}
\]

where $\varepsilon_{\omega}$, $\varepsilon_{\zeta}$, $\varepsilon_{MAC}$ are tolerance limits to decide if mode $i$ estimated from initial point $p$ is the same that mode $j$ estimated from initial point $q$. MAC (modal assurance criterion) shows the degree of correlation between two vectors and it is computed as

\[
MAC(v_1, v_2) = \frac{|v_1^H v_2|^2}{(v_1^H v_1) (v_2^H v_2)}, \tag{2.48}
\]

where $v^H$ means Hermitian operator.

$\varepsilon_{\zeta}$ could be defined in a relative way, like $\varepsilon_{\omega}$, but it is well known that the damping is estimated with greater variability than the natural frequencies ([87]), so we prefer to use $\varepsilon_{\zeta}$ in absolute terms. This means that if two modes identified from different starting point have similar natural frequencies and similar mode shapes, but with damping ratios differing in, for instance, 0.02, we assume they correspond to the same mode.

### 2.7 Example 1: IASC-ASCE benchmark problem (simulated data)

We have evaluated the performance of the EM method for Operational Modal Analysis using the data provided by the ASCE benchmark problem for structural health monitoring [43]. This benchmark studies consist in Phase I and Phase II, and each phase is divided in simulated and experimental problem. The benchmark structure is a four-story, two-bay by two-bay steel-frame scale model structure built in the Earthquake Engineering Research Laboratory at the University of British Columbia, Canada (Figs. 2.1 and 2.6). The January 2004 issue of the Journal of Engineering Mechanics contains the results of six different studies of the Phase I simulated benchmark problems, together with a definition and overview paper [56].

We have selected this example because it has analytical and experimental part, Matlab subroutines for simulations and experimental data are available on the internet ([43]), and results can be compared with other researchers.

MATLAB-based finite element analysis code obtained from the IASC-ASCE SHM Task Group web site [43] has been used to simulate the dynamic response of the prototype structure. Two finite-element models based on the actual test structure were developed by the Task Group to generate the simulated structural response data: a 12DOF shear
2.7 Example 1: IASC-ASCE benchmark problem (simulated data)

Figure 2.1: Model of the benchmark structure and location of the 16 measured nodes.

building model and a more complex 120DOF model. We have used the former because of its simplicity (there are 3 DOF per floor): the floors move as rigid bodies, with translation in the $x$ and $y$ directions and rotation $\theta$ about the center column. The natural frequencies and the mode shapes of this model are given in Fig. 2.2.

In this work we have used the horizontal acceleration of 16 nodes of the structure. These nodes are located at the center of each side of the structure, two in the $x$ and two in the $y$ directions per floor (called $\ddot{y}_1$, $\ddot{y}_2$, ..., $\ddot{y}_{16}$ in Fig. 2.1).

The structure response has been generated by exciting the model with broadband ambient inputs applied in the $x$ and $y$ directions, using a sampling frequency equal to 1000 Hz and 20 seconds of total duration. 1% modal damping has been assigned to each mode. Finally, the observed value is the sum of the structure response and Gaussian noise. The root mean square (RMS) of this added noise is equal to the 30% of the largest structure response RMS.

2.7.1 Results using one simulated time history response.

First, we are going to perform modal identification using one simulated time history response. Figure 2.3 shows the power spectral density function (Welch method) for this simulated case and for eight nodes of the structure: four in $x$ direction (nodes 1, 5, 9 and 13 in Fig. 3.1) and four in $y$ direction (nodes 4, 8, 12 and 16). The theoretical natural frequencies are also plotted in dashed lines, and it is observed that most of this theoretical frequencies are coincident with a maximum of the spectrum, although not all peaks are present in all channels.

The first value to determine in an OMA problem is the order of the state space
1. $f = 9.41$ Hz
2. $f = 11.79$ Hz
3. $f = 16.38$ Hz
4. $f = 25.54$ Hz
5. $f = 32.01$ Hz
6. $f = 38.66$ Hz
7. $f = 44.64$ Hz
8. $f = 48.01$ Hz
9. $f = 48.44$ Hz
10. $f = 60.15$ Hz
11. $f = 67.48$ Hz
12. $f = 83.62$ Hz

Figure 2.2: Natural frequencies and mode shapes corresponding to the 12 DOF shear model (the circles are the points where measurements are recorded).

Figure 2.3: Power spectral density of accelerations in different nodes of the structure. The analytical natural frequencies are also plot in dashed lines.
2.7 Example 1: IASC-ASCE benchmark problem (simulated data)

There are different techniques to evaluate this parameter: for example, the number of singular values different to zero of the projection matrix in SSI \([79]\), the Akaike information criterion AIC for time series \([54]\), the contribution of the identified modes \([23]\), by inspection in stabilization diagrams,... In simulated structures the theoretical state space order equals two times the number of modes that are present in the data; we know there are 12 modes in the simulated case, so the theoretical order of the state space model is 24. For this reason, we decided to analyse the behaviour of the EM under the exact order.

Modal parameters have been obtained using three different approaches:

1. SS thankful method. We have used the SSI-DATA algorithms detailed in \([79]\).

2. EM method using SSI to obtain the initial value (EM1 in the following). We have applied the EM method as described in section 2.4.3. The value we have used for the initial parameters \(\theta_0 = (A_0, C_0, Q_0, R_0, \bar{x}_0^0, P_0^0)\) has been: matrices \((A, C, Q, R)\) estimated by SSI method, and zero matrices for \((\bar{x}_0^0, P_0^0)\).

3. EM method using random starting values (EM2 in the following). In this case, the initial matrices \(\theta_0 = (A_0, C_0, Q_0, R_0, \bar{x}_0^0, P_0^0)\) have been generated as is described in section 2.5.1.

The results of the three methods for the simulated case are shown in Table 2.1 where theoretical modal parameters are denoted by subscript \(th\), and identified parameters with subscript \(id\). We have considered that a parameter has been identified if it is verified at once that

\[
\left| \omega_{th} - \omega_{id} \right| \leq \epsilon_\omega = 0.02, \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad
<table>
<thead>
<tr>
<th>mode</th>
<th>$\omega_{id}(Hz)$</th>
<th>$\zeta_{id}$(%)</th>
<th>MAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.41</td>
<td>0.70</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>11.79</td>
<td>1.31</td>
<td>0.39</td>
</tr>
<tr>
<td>3</td>
<td>16.38</td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>25.54</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>5</td>
<td>32.01</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>6</td>
<td>38.66</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>7</td>
<td>48.01</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>8</td>
<td>57.98</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>9</td>
<td>67.48</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>10</td>
<td>83.62</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>11</td>
<td>94.44</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>12</td>
<td>10.15</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>13</td>
<td>50.86</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>14</td>
<td>67.48</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>15</td>
<td>83.62</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>16</td>
<td>94.44</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>17</td>
<td>10.15</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>18</td>
<td>50.86</td>
<td>1.00</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 2.1: Modal parameters identified from one simulated time history response using SSI method and EM method. In SSI and EM1, values in light gray do not verify \((2.49)-(2.51)\). In EM2, \(n\) is the number of times the parameter has been identified over 100 starting random values.
2.7 Example 1: IASC-ASCE benchmark problem (simulated data) 23

Figure 2.4: (a) Theoretical eigenvalues (□), and evolution from eigenvalues of matrix A identified using SSI (○) and eigenvalues of matrix A identified using EM (+). (b) Evolution from random starting eigenvalues (○) to EM eigenvalues (+). (c) Likelihood of SSI, EM1 and EM2 estimates shown in (a) and (b). The likelihood of two additional random starting values has been also plotted.

- The EM iterations correct the failed SSI eigenvalues A and B towards the theoretical values 6 and 8. It also joins two real poles in a complex pole that grows from the real axis towards mode 5.

- Mode 4 is not identified by any of the methods.

Looking at Table 2.1 and Figure 2.4 we can conclude that the EM algorithm improves the solution given by the SSI method. Not only the eigenvalues, even the eigenvectors are also improved: for example, the eigenvalue of mode 12 has been well estimated using SSI, but not so its eigenvector ($MAC = 0.46$); however, the eigenvector identified with EM1 method is very close to the theoretical one ($MAC = 0.99$). Similar behaviour can be also seen for modes 6, 8 and 10.

For the same simulated time history response we have generated 100 different random starting points within the unit circle by mean of the procedure described in section 2.5.1. The results obtained from these random starting points using the EM2 method have been included in Table 2.1. We have also included the number of times each mode has been identified over the total of the 100 starting points, called n.

Figure 2.4(b) shows the convergence of eigenvalues of matrix A to the theoretical
Operational Modal Analysis using the EM algorithm

Figure 2.5: Boxplots of SSI, EM1 and EM2 results for the 100 simulated cases.

eigenvalues for one of the 100 random starting points. In this case, only the mode 8 remains unidentified and it has been obtained the eigenvalue \(a\) in place. This mode is hard to find, and in fact, in none of the 100 starting points has been identified (see Table 2.1).

Comparing Figure 2.4(a) and 2.4(b), we see how the solution obtained by the expectation maximization algorithm depends on the chosen starting point.

Figure 2.4(c) shows the likelihood of the SSI solution (the circles (◦) in Figure 2.4(a)), and how the EM algorithm increase the likelihood with the iterations. The likelihood of the solution estimated by EM1 method (the crosses (+) in Figure 2.4(a)) is indicated as EM1 in Figure 2.4(c).

On the other hand, the likelihood of the solution estimated by EM2 method (the crosses (+) in Figure 2.4(b)) is called EM2a in the Figure 2.4(c). Two additional random starting points (from the total of 100 studied) have been also included, points EM2b and EM2c. From the point of view of the MLE method, we should choose the solution with higher likelihood, that is, the EM2c. However, we think we are wasting the information provided by the remaining 99 starting values. For this reason, we proposed as modal parameters those that are more repeated in all the starting points (the so-called EM2 method).
2.7 Example 1: IASC-ASCE benchmark problem (simulated data)

2.7.2 Results using 100 simulated time history response.

We are going now to analyse 100 simulated cases. Summarizing the results of these 100 simulations is not easy, but it would give us a clear comparison of the differences of the three methods. We have chosen the box-plot to present simultaneously the estimated value for the modal parameters (Figure 2.5).

The box-plot is a graphical display based on the order statistics summaries of median and quartiles. A box is drawn from the first quartile to the third. The distributional center is indicated by a line at the median, within each box. The result spread or variation is shown by the box’s height, which is the interquartile range. The plot is completed by a line at each side of the box that shown the “acceptable” values according to interquartile range and finally, the outliers which are extreme or discordant estimations (plotted as crosses in the figure).

The upper row of Figure 2.5 shows the box-plots for the modal parameters estimated with the SSI algorithm, the second row corresponds to the EM1 method, and the third row corresponds to the EM2 method. By columns, the first one presents the box-plots for the natural frequencies, the second one shows the results for the damping ratio and the last column is reserved for the MAC value.

Considering first the SSI method, the estimated frequencies of modes 1, 2, 3, 11 and 12 are very similar and are close to the actual value (note that the difference is close to zero). The estimated frequencies for modes 4, 5, 8 and 10, however, present a huge variability. The frequencies of the modes 6, 7 and 9 show intermediate variability. This behaviour is also observed in the plot corresponding to the damping ratio and the MAC value. The only difference is that the MAC obtained for mode 12 are about 0.6 and very far from the values that we consider acceptable (greater than 0.9).

Taking into account the results obtained with the EM1 method (second row in Figure 2.5), we can notice the improvement achieved in the natural frequencies, damping ratios and MAC values corresponding to all the modes except mode 8. Excluding this mode, most of the estimated frequencies correspond to the theoretical values (although with some outliers), most of damping ratios are close to 0.01, and most of MAC values are close to one.

This improvement is even more evident when using the algorithm EM2, except for mode 8 and for some very few outliers, the results are very close in most of the simulations to the theoretical values.

Figure 2.5 shows the whole estimated modal parameters for the 100 simulated responses. From these results we have extracted the modes that verify the criteria (2.49)-(2.51). We consider that these modes have been properly identified. Table 2.2 shows the average values and standard deviations of the modal frequencies, damping ratios and mode shapes. The table also includes the number of simulations, called N (over a total of 100), each mode has been identified. We have observed that:

- Using SSI method, we have identified four modes in the 100% of the simulations (modes 1, 2, 3 and 11) and two modes not so good (modes 7 and 9). The rest of the modes have not been identified. However, the modes estimated with this method have the best mean and standard deviation values of the three methods.

- Using method EM1, the results are improved. We can say that all modes have been identified except mode 8. We have observed in the simulations that, in general,
parameters well estimated by SSI method are not modified when using these parameters as starting point for the EM algorithm; but sometimes it can happen that, as the likelihood is maximized, some of them deteriorate lightly (as can be seen for modes 1 and 2) while the rest are improved. Usually the estimated values by the EM algorithm present bias estimation with $\omega_{th} < \omega_{id}$ and $\zeta_{th} < \zeta_{id}$, something known when the poles of matrix A are close to the unit circle \[93\].

- The method EM2 provides the best results of the three methods considered in this work (in the sense of modal parameters).

- A remarkable fact in the three methods is that significantly larger values of standard deviation are observed for damping ratios than for modal frequencies in most cases. For this reason, we have preferred to take $\varepsilon_{\zeta}$ in absolute value while $\varepsilon_{\omega}$ is chosen in a relative sense in equations (2.49) and (2.50).

- It is important to note that mode 8 is not identified by none of the three methods. We think this is because this mode could not be excited: even if we consider higher state space models, this mode is hard to find.

- In general terms, we have found that the estimates of modal parameters are improved after applying the EM algorithm to the SSI estimates. But sometimes it happens that the estimates of the first three modes become worse, although the likelihood is maximized. In principle, the properties of maximum likelihood estimation ensures that the estimates have asymptotically minimum variance under certain conditions, but this does not necessarily imply that for finite samples this property is maintained. Moreover, the proposed algorithm tries to minimize the errors as a whole, but reducing the variance of the whole can increase the variance of one parameter considered individually. This fact can be corrected in the algorithm, for example, using a weighted maximization of the likelihood, but then the computed solution is not the maximum likelihood solution.

### 2.8 Example 2: IASC-ASCE benchmark problem (experimental data)

The experimental phase of the benchmark problem was carried out in August 2002: the model structure was instrumented with three uniaxial accelerometers at each story level including the base. Nine configurations were investigated in order to simulate damage within the test structure by removing bracing and loosening beam-column connections (the configuration 1 was the reference -undamaged- case). For each configuration, experimental data were generated by three types of excitation: impacts of a sledge hammer, ambient vibration and electrodynamic shaker on the roof. We have analysed the time history response corresponding to configuration 1 and ambient vibration, named “shm01a” in the data set obtained at the task group website \[43\]. A detailed description of the test structure and experimental procedure can be found on this web site.

The duration of the acceleration data is 300 seconds with sampling frequency equal to 200 Hz, which have been filtered by a Butterworth high-pass filter with cut-off frequency equal to 0.1 Hz to eliminate the mean and drift. The base accelerations are excluded so the total number of channel is 12.
### Table 2.2: Modal parameters identified from 100 simulated time history responses using SSI method and EM method.

<table>
<thead>
<tr>
<th>mode</th>
<th>exact</th>
<th>SSI mean</th>
<th>SSI std</th>
<th>EM1 mean</th>
<th>EM1 std</th>
<th>EM2 mean</th>
<th>EM2 std</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.41</td>
<td>9.44</td>
<td>0.03</td>
<td>9.48</td>
<td>0.17</td>
<td>9.53</td>
<td>0.28</td>
</tr>
<tr>
<td>2</td>
<td>11.79</td>
<td>11.81</td>
<td>0.03</td>
<td>11.86</td>
<td>0.37</td>
<td>11.95</td>
<td>0.48</td>
</tr>
<tr>
<td>3</td>
<td>16.38</td>
<td>16.41</td>
<td>0.05</td>
<td>16.42</td>
<td>0.24</td>
<td>16.42</td>
<td>0.45</td>
</tr>
<tr>
<td>4</td>
<td>25.54</td>
<td>25.88</td>
<td>0.47</td>
<td>25.62</td>
<td>0.30</td>
<td>25.60</td>
<td>0.29</td>
</tr>
<tr>
<td>5</td>
<td>32.01</td>
<td>—</td>
<td>—</td>
<td>32.08</td>
<td>0.37</td>
<td>32.08</td>
<td>0.35</td>
</tr>
<tr>
<td>6</td>
<td>38.66</td>
<td>38.30</td>
<td>1.53</td>
<td>38.63</td>
<td>0.36</td>
<td>38.63</td>
<td>0.50</td>
</tr>
<tr>
<td>7</td>
<td>44.64</td>
<td>44.66</td>
<td>2.01</td>
<td>44.87</td>
<td>0.57</td>
<td>44.89</td>
<td>0.57</td>
</tr>
<tr>
<td>8</td>
<td>48.08</td>
<td>—</td>
<td>—</td>
<td>48.15</td>
<td>0.34</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>48.44</td>
<td>48.22</td>
<td>1.54</td>
<td>48.42</td>
<td>0.46</td>
<td>48.38</td>
<td>0.78</td>
</tr>
<tr>
<td>10</td>
<td>60.15</td>
<td>—</td>
<td>—</td>
<td>60.32</td>
<td>0.50</td>
<td>60.35</td>
<td>0.97</td>
</tr>
<tr>
<td>11</td>
<td>67.48</td>
<td>67.49</td>
<td>0.37</td>
<td>67.45</td>
<td>0.53</td>
<td>67.42</td>
<td>1.06</td>
</tr>
<tr>
<td>12</td>
<td>83.62</td>
<td>—</td>
<td>—</td>
<td>83.62</td>
<td>0.61</td>
<td>83.70</td>
<td>0.90</td>
</tr>
</tbody>
</table>

(a) Natural frequencies (Hz). \(N\) is the number of times the parameter has been identified;  

<table>
<thead>
<tr>
<th>mode</th>
<th>exact</th>
<th>SSI mean</th>
<th>SSI std</th>
<th>EM1 mean</th>
<th>EM1 std</th>
<th>EM2 mean</th>
<th>EM2 std</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.91</td>
<td>0.03</td>
<td>1.41</td>
<td>0.37</td>
<td>1.31</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.22</td>
<td>0.07</td>
<td>2.13</td>
<td>0.73</td>
<td>1.27</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.75</td>
<td>0.06</td>
<td>1.07</td>
<td>0.73</td>
<td>1.54</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>2.11</td>
<td>0.23</td>
<td>0.89</td>
<td>0.14</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.01</td>
<td>0.12</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>1.29</td>
<td>0.23</td>
<td>1.06</td>
<td>0.11</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>1.74</td>
<td>0.41</td>
<td>1.11</td>
<td>0.08</td>
<td>1.31</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.22</td>
<td>0.03</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>1.21</td>
<td>0.21</td>
<td>0.99</td>
<td>0.11</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.17</td>
<td>0.46</td>
<td>2.37</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.00</td>
<td>1.11</td>
<td>0.08</td>
<td>0.96</td>
<td>0.22</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>0.60</td>
<td>0.13</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>

(b) Damping ratios (%);  

<table>
<thead>
<tr>
<th>mode</th>
<th>SSI mean</th>
<th>SSI std</th>
<th>EM1 mean</th>
<th>EM1 std</th>
<th>EM2 mean</th>
<th>EM2 std</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.998</td>
<td>0.76</td>
<td>0.95</td>
<td>1.08</td>
<td>0.983</td>
<td>2.25</td>
</tr>
<tr>
<td>2</td>
<td>0.999</td>
<td>0.01</td>
<td>0.999</td>
<td>0.10</td>
<td>0.996</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>0.999</td>
<td>0.00</td>
<td>0.999</td>
<td>0.10</td>
<td>0.999</td>
<td>0.21</td>
</tr>
<tr>
<td>4</td>
<td>0.957</td>
<td>1.07</td>
<td>0.972</td>
<td>2.30</td>
<td>0.969</td>
<td>2.26</td>
</tr>
<tr>
<td>5</td>
<td>—</td>
<td>—</td>
<td>0.989</td>
<td>1.35</td>
<td>0.987</td>
<td>1.28</td>
</tr>
<tr>
<td>6</td>
<td>0.947</td>
<td>2.98</td>
<td>0.982</td>
<td>1.97</td>
<td>0.983</td>
<td>1.97</td>
</tr>
<tr>
<td>7</td>
<td>0.949</td>
<td>1.96</td>
<td>0.976</td>
<td>1.38</td>
<td>0.971</td>
<td>1.83</td>
</tr>
<tr>
<td>8</td>
<td>—</td>
<td>—</td>
<td>0.955</td>
<td>2.47</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>0.970</td>
<td>1.61</td>
<td>0.984</td>
<td>2.52</td>
<td>0.968</td>
<td>2.62</td>
</tr>
<tr>
<td>10</td>
<td>—</td>
<td>—</td>
<td>0.979</td>
<td>1.50</td>
<td>0.983</td>
<td>1.34</td>
</tr>
<tr>
<td>11</td>
<td>0.996</td>
<td>0.18</td>
<td>0.998</td>
<td>0.24</td>
<td>0.994</td>
<td>0.41</td>
</tr>
<tr>
<td>12</td>
<td>—</td>
<td>—</td>
<td>0.978</td>
<td>0.75</td>
<td>0.983</td>
<td>0.98</td>
</tr>
</tbody>
</table>

(c) Mode shapes.
All parametric system identification methods based on the state space model require to know the model order $n_s$, which in theory, is twice the number of identified modes. In OMA it is customary to estimate the state space model for a wide range of orders and the eigenfrequencies obtained for all these orders are plotted in an eigenfrequency vs. model order diagram, called a stabilization diagram [51]. Experience on a very large range of problems shows that in such analysis, the eigenfrequencies corresponding to physical modes appear at most of the used model orders, while mathematical and spurious poles tend to scatter around the frequency range. Then, from such diagram it is possible to select the optimal system order, and for this order, the valid system poles.

The implementation of this strategy can be executed by initially choosing a sufficiently high order for the state space model, and then gradually reduce the order of the model. System identification is performed with every model order so this procedure yields a set
Figure 2.7: Stabilization diagram obtained with SSI method. The used symbols are: “⊕” for stable pole and “+” for unstable pole. A mode is stable if $\varepsilon_\omega = 0.02$, $\varepsilon_\zeta = 0.03$, $\varepsilon_{MAC} = 0.10$.

of modal parameters for each selected order. Parameters that belong to two different model orders are then compared according to some preset criteria such as

$$\frac{|\omega_{pi} - \omega_{qj}|}{\omega_{pi}} \leq \varepsilon_\omega,$$  \hspace{1cm} (2.52)

$$\frac{|\zeta_{pi} - \zeta_{qj}|}{\zeta_{pi}} \leq \varepsilon_\zeta,$$ \hspace{1cm} (2.53)

$$1 - MAC(\psi_{pi}, \psi_{qj}) \leq \varepsilon_{MAC},$$ \hspace{1cm} (2.54)

where $\varepsilon_\omega$, $\varepsilon_\zeta$, $\varepsilon_{MAC}$ are tolerance limits to decide if mode $i$ estimated from model order $p$ is the same that mode $j$ estimated from model order $q$ (it is usual to take $q = p + 1$, that is, to compare consecutive model orders). This procedure is repeated for all available sets of modal parameters identified at each order in a sequential manner, and finally, the frequencies are plotted against their corresponding model orders, distinguishing between stable and unstable modes. It is usual to add to the graphic the power spectral density plot of a selected channel, or the Averaged Normalized Power Spectral Density (ANPSD)

$$ANPSD(f) = \frac{1}{n_o} \sum_{i=1}^{n_o} \frac{PSD(f)_i}{\sum_{j=1}^{N} PSD(f_j)_i}$$

where $PSD(f)_i$ is the power spectral density of channel $i$. In this case, auto-spectra
Table 2.3: Resulting modal parameters for the experimental time history response using SSI, EM1 and EM2 methods. $n$ is the number of times that the parameter has been identified using EM2 from a total of 100. Results obtained in other studies are included.

<table>
<thead>
<tr>
<th>Mode</th>
<th>SSI $\omega$ (Hz) $\zeta$ (%)</th>
<th>EM1 $\omega$ (Hz) $\zeta$ (%)</th>
<th>EM2 $\omega$ (Hz) $\zeta$ (%)</th>
<th>N</th>
<th>Other studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.52 1.04</td>
<td>7.55 1.46</td>
<td>7.55 2.12</td>
<td>36</td>
<td>7.48 7.49</td>
</tr>
<tr>
<td>2</td>
<td>7.77 0.60</td>
<td>7.81 1.53</td>
<td>7.83 2.10</td>
<td>85</td>
<td>7.76 7.76</td>
</tr>
<tr>
<td>3</td>
<td>14.53 0.44</td>
<td>14.54 0.78</td>
<td>14.90 4.32</td>
<td>48</td>
<td>14.48 14.49</td>
</tr>
<tr>
<td>4</td>
<td>19.89 0.01</td>
<td>19.89 0.01</td>
<td>19.89 0.02</td>
<td>100</td>
<td>19.89 19.89</td>
</tr>
<tr>
<td>5</td>
<td>21.01 0.04</td>
<td>21.08 0.52</td>
<td>21.23 1.11</td>
<td>100</td>
<td>21.01 21.01</td>
</tr>
<tr>
<td>6</td>
<td>23.49 85.42</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>22.69</td>
</tr>
<tr>
<td>7</td>
<td>25.49 0.09</td>
<td>25.62 0.47</td>
<td>25.88 2.69</td>
<td>67</td>
<td>25.49</td>
</tr>
<tr>
<td>8</td>
<td>28.16 2.12</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>28.31</td>
</tr>
<tr>
<td>9</td>
<td>28.26 0.09</td>
<td>28.37 0.32</td>
<td>28.40 0.49</td>
<td>87</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>28.87 0.41</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>—</td>
<td>29.04 2.12</td>
<td>28.96 3.74</td>
<td>37</td>
<td>—</td>
</tr>
<tr>
<td>12</td>
<td>29.36 0.20</td>
<td>29.15 0.49</td>
<td>29.20 0.78</td>
<td>88</td>
<td>—</td>
</tr>
<tr>
<td>13</td>
<td>—</td>
<td>30.44 4.34</td>
<td>31.50 6.45</td>
<td>26</td>
<td>—</td>
</tr>
<tr>
<td>14</td>
<td>32.03 1.51</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>15</td>
<td>—</td>
<td>32.14 1.85</td>
<td>32.19 2.33</td>
<td>52</td>
<td>—</td>
</tr>
<tr>
<td>16</td>
<td>—</td>
<td>34.03 0.74</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>17</td>
<td>34.04 0.11</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>33.31 2.45 36</td>
</tr>
<tr>
<td>18</td>
<td>—</td>
<td>35.24 8.40</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>19</td>
<td>49.51 2.94</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>20</td>
<td>—</td>
<td>52.22 3.03</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>21</td>
<td>59.99 0.03</td>
<td>59.99 0.02</td>
<td>59.84 0.37</td>
<td>69</td>
<td>—</td>
</tr>
<tr>
<td>22</td>
<td>—</td>
<td>—</td>
<td>81.67 1.39</td>
<td>19</td>
<td>—</td>
</tr>
</tbody>
</table>

functions are normalized and averaged to obtain an average spectra density function from all channels that, normally, shows all resonance frequencies of the system. The stable frequencies should be reflected by the peaks of the ANPSD plot.

The stabilization diagram of the experimental data is presented at Figure 2.7. This figure has been constructed using SSI method because is fast and accurate enough to build this type of plots. Modal parameters that belong to two consecutive model orders have been compared according to the criteria expressed in equations (2.49) to (2.51). Taking into account the stabilization diagram, we have chosen a model order equal to $n_s = 30$, because the principal stable modes are included for this order. Table 2.3 shows the experimental modal frequencies and damping ratios computed applying the same methods that in the simulated phase. If the values estimated with different methods are located in the same row of the table means that they correspond to the same mode, that is, verify the tolerances (2.49) to (2.51) between them.

The experimental vibration measurements of this structure have been also studied by various researchers applying different system identification techniques: for example, J. Ching and J. L. Beck [26] applied a two-stage Bayesian approach; and B. Alicioglu and M. Lus [3] used the SSI method too. The results of the two analysis are also presented in Table 2.3 for comparison purpose (in this case the frequencies have been placed in the table in ascending order and not following the criteria (2.49) to (2.51) because we don’t...
have their corresponding mode shapes and damping ratios.

Comprehensive analyses of Table 2.3 and Figure 2.7 show the following:

- The lower frequencies (first five modes) have been identified by the three algorithms, and they also appear in the two references mentioned above ([26] and [3]). Moreover, they are very clear modes in the spectrum and the stabilization diagram.

- Modes 7, 9, 12 and 21, which are higher frequency modes, have been identified by the three methods. None of them appears in [26] and two could coincide with the frequencies given in [3]. The stabilization diagram indicates the existence of stable modes in these frequency ranges, although it is difficult to specify more graphic information.

- In total, the three methods agree on the identification in 9 of the fifteen modes that the model order ($n_s = 30$) allowed us to estimate.

- Modes 6, 8, 10, 14 and 19 have been identified by the SSI method but they are not confirmed by any of the two versions of the EM method. The stabilization diagram is of little assistance. On the other hand, Table 2.3 shows the existence of close frequency values identified by the EM algorithms, but they differ in the damping or the eigenvector. We have checked the solutions obtained with higher orders (up to order 50) for the SSI method and the problem worsens: the new modes are still not mismatched with those of the EM algorithm and appear many new ones without correspondence.

- There are other three modes that have been identified for EM1 and EM2 jointly, modes 11, 13 and 15. In general terms, the modes estimated with EM2 for a larger number of starting points have been also identified either by one of the other methods or by both at the same time. This is a very valuable property of EM2 method.

- Modes 16, 18 and 20 have been only detected by the EM1 method. Finally, the natural frequency near to 80 Hz has been estimated only by the algorithm EM2. However, we can see in the stabilization diagram (Figure 2.7) that this mode exits and it is identified by the SSI method when using higher state space orders.

- The modal analysis of a structure is a complex problem and we should not take the results provided by only one method as the exact solution. We can have errors of two types: to include spurious modes and to forget others which really exist. This analysis shows that having multiple solutions is very useful for a better identification of the vibration modes of a structure.

2.9 Conclusions

The aim of this Chapter was to present the application of a time-domain stochastic system identification method based on the EM algorithm to operational modal analysis of structures. The EM algorithm is a well-known tool for iterative maximum likelihood estimation which for state space models has a particularly neat form. Due to the EM algorithm is iterative, a starting value is needed. When dealing with multivariate problems, like state space equations, the construction of a starting value is not trivial. Not
any starting value works, and most of them failed in the iterative process. In our opinion, the easiest approach is to begin with the estimates of another identification method. We have chosen SSI because it is very used in output-only problems, and it is fast and robust. On the other hand, we have developed a procedure to generate feasible random starting points as well.

The proposed method has been evaluated through a numerical and an experimental study in the context of the ASCE benchmark problem. The numerical results show that the proposed method estimates natural frequencies, damping ratios and mode shapes reasonably well in the presence of measurement noises even.

From our experiments, the following comments can be made:

- SSI is a powerful identification method for output only analysis and has numerous advantages. In fact is one of the most currently used method to perform operational modal analysis of structures. However, maximizing the likelihood in the state space model is equivalent to minimizing the error, that is, the state space model fitted by the EM algorithm to the data has a lower error than the SSI one. So we expect that the modal parameters (which are derived from the state space model) have a lower error as well. And the results of the examples seem to support this fact.

- But using a single run of the EM can lead to suboptimal solutions. For this reason we have proposed a procedure for building random initial values for the EM. This is not trivial because we are working with multivariate systems and any value taken at random does not work. Generating different starting point we can pick the optimal solution between them (we choose the parameters that appear again and again at different starting points). From the point of view of modal parameters, we have obtained the best results using the EM2 method.

- Although not crucial, the random stating values allows us to use the EM method independently, without recourse to another identification method.

This Chapter shows that the EM algorithm can be used to compute the maximum likelihood estimate of the modal parameters of a structure using output-only measurements. This is important because the maximum likelihood parameters have well-known statistical properties.

### 2.10 Appendix

**Proof of Property 2.1**

The process of computing the conditional expectations (2.24), (2.25) and (2.26) is reduced, by single algebraic manipulations, to compute the conditional mean \(\mathbb{E}[x_t|Y_N, \theta_j]\), and covariance functions \(\text{Cov}[x_t, x_{t-1}|Y_N, \theta_j]\). It is important to note that the vector \(\mathbb{E}[x_t|Y_t, \theta_j]\) is the usual Kalman filter estimator \(x_t^f\) (Eq. (A.51)), whereas \(\mathbb{E}[x_t|Y_N, \theta_j]\), \(t = 0, 1, \ldots, N-1\) is the smoother estimator of \(x_t\) based on all the observed data (Eq. (A.56)). Therefore, we take

\[
\mathbb{E}[x_t|Y_N, \theta_j] = x_t^N,
\]

and covariance functions

\[
\text{Cov}[x_t, x_t|Y_N, \theta_j] = \text{Var}[x_t|Y_N, \theta_j] = P_t^N,
\]
\[ \text{Cov}[x_t, x_{t-1}|Y_N, \theta_j] = P_{t,t-1}^N. \]

where \( x_t^N, P_t^N \) and \( P_{t,t-1}^N \) are computed using the Kalman filtering results according to Properties \( \text{A.1} \), \( \text{A.2} \), \( \text{A.3} \) and \( \text{A.6} \).

- **Proof of Equation (2.24)**

\[
E[\log |P_0|] = E[\log |P_0| + (x_0 - \bar{x}_0)^T P_0^{-1}(x_0 - \bar{x}_0)|Y_N, \theta_j] \\
= E[\log |P_0| + x_0^T P_0^{-1} x_0 - \bar{x}_0^T P_0^{-1} \bar{x}_0 + \bar{x}_0^T P_0^{-1} \bar{x}_0] \\
\overset{(C.20)}{=} \log |P_0| + \text{tr}(P_0^{-1} P_0^N) + (x_0^N)^T P_0^{-1} x_0^N - \bar{x}_0^T P_0^{-1} \bar{x}_0 + \bar{x}_0^T P_0^{-1} \bar{x}_0 \\
\overset{(C.22)}{=} \log |P_0| + \text{tr}(P_0^{-1}[P_0^N + (x_0^N - \bar{x}_0)(x_0^N - \bar{x}_0)^T]).
\]

- **Proof of Equation (2.25)**

\[
E[\log |Q| + \sum_{t=1}^{N} (x_t - Ax_{t-1})^T Q^{-1}(x_t - Ax_{t-1})|Y_N, \theta_j] \\
= N \log |Q| + \\
+ \sum_{k=1}^{N} E[(x_t^T Q^{-1} x_t - x_{t-1}^T Q^{-1} A x_{t-1} - x_{t-1}^T A^T Q^{-1} x_t + x_{t-1}^T A^T Q^{-1} A x_{t-1})|Y_N, \theta_j]
\]

\[
E[(x_t^T Q^{-1} x_t)|Y_N, \theta_j] \overset{(C.22)}{=} \text{tr}(Q^{-1} P_t^N) + (x_t^N)^T Q^{-1} x_t^N \\
\overset{(C.20)}{=} \text{tr}(Q^{-1}(P_t^N + x_t^N (x_t^N)^T));
\]

\[
E[(x_t^T Q^{-1} A x_{t-1})|Y_N, \theta_j] \overset{(C.22)}{=} \text{tr}(Q^{-1} A P_{t,t-1}^N) + (x_t^N)^T Q^{-1} A x_{t-1} \\
\overset{(C.20)}{=} \text{tr}(Q^{-1} A (P_{t,t-1}^N + x_{t-1}^N (x_{t-1}^N)^T));
\]

\[
E[(x_{t-1}^T A^T Q^{-1} x_t)|Y_N, \theta_j] \overset{(C.22)}{=} \text{tr}(Q^{-1} P_{t,t-1}^N A^T) + (x_{t-1}^N)^T Q^{-1} A^T x_t \\
\overset{(C.20)}{=} \text{tr}(Q^{-1} (P_{t,t-1}^N + x_t^N (x_{t-1}^N)^T) A^T);
\]

\[
E[(x_{t-1}^T A^T Q^{-1} A x_{t-1})|Y_N, \theta_j] \overset{(C.22)}{=} \text{tr}(Q^{-1} A P_{t,t-1}^N A^T) + (x_{t-1}^N)^T A^T Q^{-1} A^T x_{t-1} \\
\overset{(C.20)}{=} \text{tr}(Q^{-1} A (P_{t,t-1}^N + x_{t-1}^N (x_{t-1}^N)^T) A^T);
\]

- **Proof of Equation (2.26)**

\[
E[\log |R| + \sum_{t=1}^{N} (y_t - C x_t)^T R^{-1}(y_t - C x_t)|Y_N, \theta_j] \\
= N \log |R| + \\
+ \sum_{k=1}^{N} E[(y_t^T R^{-1} y_t - y_t^T R^{-1} C x_t - x_t^T C^T R^{-1} y_t + x_t^T C^T R^{-1} C x_t)|Y_N, \theta_j];
\]
Deriving and equating to zero:

\[ \begin{align*}
\mathbb{E}[(y_t^T R^{-1} y_t | Y_N, \theta_j)] &= y_t^T R^{-1} y_t \quad \text{(8.21)} \\
\mathbb{E}[(y_t^T R^{-1} C x_t | Y_N, \theta_j)] &= y_t^T R^{-1} C x_t \quad \text{(C.20)} \\
\mathbb{E}[(x_t^T C^T R^{-1} y_t | Y_N, \theta_j)] &= (x_t^N)^T C^T R^{-1} y_t \quad \text{(C.20)} \\
\mathbb{E}[(x_t^T C^T R^{-1} C x_t | Y_N, \theta_j)] &= \text{tr}(R^{-1} C(P_t^N C^T) + (x_t^N)^T C R^{-1} C^T x_t^N) \quad \text{(C.22)} \\
&= \text{tr}(R^{-1} C(P_t^N + x_t^N (x_t^N)^T C^T)).
\end{align*} \]

**Proof of Property 2.2**

Deriving and equating to zero:

- \[ \frac{\partial}{\partial x_0} \mathbb{E}[l_{X,Y}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial x_0} \mathbb{E}[l_1(\bar{x}_0, P_0)|Y_N, \theta_j] = 0; \]

\[ \begin{align*}
\mathbb{E}[l_1(\bar{x}_0, P_0)|Y_N, \theta_j] &= \log |P_0| + \text{tr} \left( P_0^{-1} \left[ P_0^N + (x_0^N - \bar{x}_0)(x_0^N - \bar{x}_0)^T \right] \right); \\
\frac{\partial}{\partial x_0} (\log |P_0|) &= 0; \\
\frac{\partial}{\partial x_0} (\text{tr}(P_0^{-1} P_0^N)) &= 0; \\
\frac{\partial}{\partial x_0} (\text{tr}(P_0^{-1} x_0^N (x_0^N)^T)) &= 0; \\
\frac{\partial}{\partial x_0} (\text{tr}(-P_0^{-1} x_0^N x_0^T)) &= -P_0^{-1} x_0^N; \\
\frac{\partial}{\partial x_0} (\text{tr}(-P_0^{-1}\bar{x}_0(x_0^N)^T)) &= -(P_0^{-1})^T \bar{x}_0; \\
\frac{\partial}{\partial x_0} (\text{tr}(P_0^{-1} x_0\bar{x}_0^T)) &= (P_0^{-1})^T \bar{x}_0 + P_0^{-1} \bar{x}_0;
\end{align*} \]
Equating to zero

\[-P_0^{-1} x_0^N - (P_0^{-1})^T x_0^N + (P_0^{-1})^T \bar{x}_0 + P_0^{-1} \bar{x}_0 = 0 \Rightarrow \bar{x}_0 = x_0^N.\]

- \[ \frac{\partial}{\partial P_0} \mathbb{E}[l_{X,Y}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial P_0} \mathbb{E}[l_1(\bar{x}_0, P_0)|Y_N, \theta_j] = 0; \]

\[ \begin{align*}
\mathbb{E}[l_1(\bar{x}_0, P_0)|Y_N, \theta_j] &= \log |P_0| + \text{tr} \left( P_0^{-1} \left[ P_0^N + (x_0^N - \bar{x}_0)(x_0^N - \bar{x}_0)^T \right] \right); \\
\frac{\partial}{\partial P_0} (\log |P_0|) &= \frac{\partial}{\partial P_0} (P_0^{-1})^T; \\
\frac{\partial}{\partial P_0} (\text{tr}(P_0^{-1} P_0^N)) &= -(P_0^{-1})^T P_0^N P_0^{-1};
\end{align*} \]
Equating to zero

\[(P_0^{-1})^T - (P_0^{-1} P_0^N P_0^{-1})^T = 0 \Rightarrow P_0 = P_0^N.\]
\[ \frac{\partial}{\partial A} \text{E}[l_{2}(A, Q)|Y_{N}, \theta_{j}] = 0 \Rightarrow \frac{\partial}{\partial A} \text{E}[l_{2}(A, Q)|Y_{N}, \theta_{j}] = 0; \]
\[ \text{E}[l_{2}(A, Q)|Y_{N}, \theta_{j}] = N \log |Q| + \text{tr} \left( Q^{-1} \left[ S_{xx} - S_{xb}A^{T} - AS_{bx} + AS_{bb}A^{T} \right] \right); \]
\[ \frac{\partial}{\partial A} (N \log |Q|) = 0; \]
\[ \frac{\partial}{\partial A} (\text{tr}(Q^{-1}S_{xx})) = 0; \]
\[ \frac{\partial}{\partial A} (\text{tr}(-Q^{-1}S_{xb}A^{T})) \overset{(C.23)}{=} -Q^{-1}S_{xb}; \]
\[ \frac{\partial}{\partial A} (\text{tr}(-Q^{-1}AS_{bx})) \overset{(C.24)}{=} -(Q^{-1})^{T}S_{bx}^{T}; \]
\[ \frac{\partial}{\partial A} (\text{tr}(Q^{-1}AS_{bb}A^{T})) \overset{(C.27)}{=} (Q^{-1})^{T}AS_{bb}^{T} + Q^{-1}AS_{bb}; \]

Equating to zero

\[ -Q^{-1}S_{xb} - (Q^{-1})^{T}S_{xb} + (Q^{-1})^{T}AS_{bb} + Q^{-1}AS_{bb} = 0 \Rightarrow A = S_{xb}S_{bb}^{-1}. \]

- \[ \frac{\partial}{\partial Q} \text{E}[l_{2}(X, Y_{N}, \theta)|Y_{N}, \theta_{j}] = 0 \Rightarrow \frac{\partial}{\partial Q} \text{E}[l_{2}(A, Q)|Y_{N}, \theta_{j}] = 0; \]
\[ \text{E}[l_{2}(A, Q)|Y_{N}, \theta_{j}] = N \log |Q| + \text{tr} \left( Q^{-1} \left[ S_{xx} - S_{xb}A^{T} - AS_{bx} + AS_{bb}A^{T} \right] \right); \]
\[ \frac{\partial}{\partial Q} (N \log |Q|) \overset{(A.28)}{=} N(Q^{-1})^{T}; \]
\[ \frac{\partial}{\partial Q} (\text{tr} \left( Q^{-1} \left[ S_{xx} - S_{xb}A^{T} - AS_{bx} + AS_{bb}A^{T} \right] \right)) \overset{(C.20)}{=} - (Q^{-1})^{T} \left[ S_{xx} - S_{xb}A^{T} - AS_{bx} + AS_{bb}A^{T} \right]^{-1}Q^{-1}; \]

Equating to zero

\[ N(Q^{-1})^{T} - (Q^{-1})^{T} \left[ S_{xx} - S_{xb}A^{T} - AS_{bx} + AS_{bb}A^{T} \right]^{-1}Q^{-1} = 0 \]
\[ \Rightarrow Q = \frac{1}{N} \left( S_{xx} - S_{xb}A^{T} - AS_{bx} + AS_{bb}A^{T} \right). \]

- \[ \frac{\partial}{\partial C} \text{E}[l_{3}(X, Y_{N}, \theta)|Y_{N}, \theta_{j}] = 0 \Rightarrow \frac{\partial}{\partial C} \text{E}[l_{3}(C, R)|Y_{N}, \theta_{j}] = 0; \]
\[ \text{E}[l_{3}(C, R)|Y_{N}, \theta_{j}] = N \log |R| + \text{tr} \left( R^{-1} \left[ S_{xy} - S_{yx}C^{T} - CS_{xy} + CS_{xx}C^{T} \right] \right); \]
\[ \frac{\partial}{\partial C} (N \log |R|) = 0; \]
\[ \frac{\partial}{\partial C} (\text{tr}(R^{-1}S_{yy})) = 0; \]
\[ \frac{\partial}{\partial C} (\text{tr}(-R^{-1}S_{yx}C^{T})) \overset{(C.26)}{=} -R^{-1}S_{yx}; \]
\[ \frac{\partial}{\partial C} (\text{tr}(-R^{-1}CS_{xy})) \overset{(C.21)}{=} -(R^{-1})^{T}S_{xy}; \]
\[ \frac{\partial}{\partial C} (\text{tr}(R^{-1}CS_{xx}C^{T})) \overset{(C.27)}{=} (R^{-1})^{T}CS_{xx}^{T} + R^{-1}CS_{xx}; \]

Equating to zero

\[ -R^{-1}S_{yx} - (R^{-1})^{T}S_{xy} + (R^{-1})^{T}CS_{xx}^{T} + R^{-1}CS_{xx} = 0 \Rightarrow C = S_{yx}S_{xx}^{-1}. \]
\[ \frac{\partial}{\partial R} E[l_{X_N, Y_N}(\theta) | Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial R} E[l_3(C, R) | Y_N, \theta_j] = 0; \]

\[ E[l_3(C, R) | Y_N, \theta_j] = N \log |R| + \text{tr} \left( R^{-1} \left[ S_{yy} - S_{yx} C^T - C S_{xy} + C S_{xx} C^T \right] \right); \]

\[ \frac{\partial}{\partial R} (N \log |R|) \overset{\text{C.23}}{=} N(R^{-1})^T; \]

\[ \frac{\partial}{\partial R} (\text{tr} \left( R^{-1} \left[ S_{yy} - S_{yx} C^T - C S_{xy} + C S_{xx} C^T \right] \right)) \overset{\text{C.26}}{=} \]

\[ = - \left( R^{-1} \left[ S_{yy} - S_{yx} C^T - C S_{xy} + C S_{xx} C^T \right] R^{-1} \right)^T; \]

Equating to zero

\[ N(R^{-1})^T - \left( R^{-1} \left[ S_{yy} - S_{yx} C^T - C S_{xy} + C S_{xx} C^T \right] R^{-1} \right)^T = 0 \]

\[ \Rightarrow R = \frac{1}{N} \left( S_{yy} - S_{yx} C^T - C S_{xy} + C S_{xx} C^T \right). \]
Chapter 3

Joint analysis of multiple records

3.1 Introduction

The dynamic data measured in a structural system is used, among others, for the estimation of the modal parameters. In general terms, one set of dynamic data recorded at a structure (or record) results in one estimate of the modal parameters.

Sometimes the data acquisition process is repeated many times, so the analyst has several similar records for the modal analysis of the structure that have been obtained in different experiments (multiple records). The first approach to address this situation consists on performing a separate analysis for each experiment. Then, the final solution should be the average of the individual estimates: the most important modes of the structure tend to appear in most of the records, while the spurious and/or weakly excited ones are only detected in individual records. However, the solutions obtained vary from one record to another, sometimes considerably, and finding the estimate of the same mode using data from different experiments is not simple, due to:

1. Not all modes of vibration appear in all records: the unmeasured excitations are different in each record, and the modes are not excited in the same way in all the records.

2. The statistical errors inherent to the estimation provide different results for the same mode in different records, and these errors are especially important in the damping.

3. The structures usually have modes of vibration with similar modal parameters and it is often difficult to differentiate among them.

4. Spurious modes are obtained due to numerical errors, over-specification of the system order, frequency components in the unknown inputs, . . .

The result is that the criteria for discarding modes and the method for combining the different true estimates are not clear and are almost always subjective decisions requiring additional information.

In this Chapter we propose the joint estimation of the modal parameters using the information provided by the multiple records. The joint estimation of the structure’s parameters combines information optimally: the modes that are repeated in different records are detected more clearly while modes specific to some records tend to blur in the
joint analysis. A specific state space model is proposed to achieve the joint estimation, and we propose to estimate this model using the EM algorithm.

### 3.2 State space model for the analysis of multiple records

Consider a linear, time invariant structural system. We put $n_o$ sensors in the system and measure $M$ different records for the system output (acceleration in this work). The measured data can be represented by

$$Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_{N}^{(r)}\}, \quad r = 1, 2, \ldots, M,$$

(3.1)

where $y_t^{(r)} \in \mathbb{R}^{n_o \times 1}$ is the observed vector for record $r$ and at time instant $t$. For simplicity, we consider all the records have the same length, $N$. Therefore, each record has $n_o \times N$ data ($n_o$ channels with $N$ data each), and in total, we have $n_o \times N \times M$ data. It is important to remark that the records are measured at different moments (so they are not simultaneously recorded) and the sensors do not change their position from one record to the next.

The objective is to estimate the modal parameters of the system using the $M$ records, that is, using all the available information.

#### 3.2.1 Classic state space model

The first approach is the most direct and naive one, that is, to estimate the state space model (2.2)

$$x_t = Ax_{t-1} + w_t, \quad w_t \sim N(0, Q),$$

(3.2a)

$$y_t = Cx_t + v_t, \quad v_t \sim N(0, R),$$

(3.2b)

for each record $r$. Because of there are $M$ different records, we obtain $M$ different estimates of the modal parameters:

$$\hat{\theta}^{(r)} = \{\hat{A}^{(r)}, \hat{C}^{(r)}, \hat{Q}^{(r)}, \hat{R}^{(r)}, \hat{x}_0^{(r)}, \hat{P}_0^{(r)}\}, \quad r = 1, 2, \ldots, M,$$

(3.3)

where $\hat{x}_0^{(r)}$ and $\hat{P}_0^{(r)}$ are the mean and variance of the initial states $x_0^{(r)}$ respectively (which are assumed to be normal distributed). Since the modal parameters are computed from matrices $A$ and $C$, and since there are $M$ different estimates of $A$ and $C$, $M$ different values for the modal parameters are obtained:

$$\hat{A}^{(r)}, \hat{C}^{(r)} \Rightarrow \hat{\omega}_j^{(r)}, \hat{\xi}_j^{(r)}, \hat{\psi}_j^{(r)}, \quad j = 1, 2, \ldots, n_s/2, \quad r = 1, 2, \ldots, M.$$

(3.4)

The most important modes of the structure should appear in most of the records. Then, modal parameters that seem to represent the same physical mode can be grouped, and mean values can be computed at the end. However, mode pairing between setups is difficult due to closely spaced modes, different excitation levels, spurious modes, etc.
3.2.2 Proposed joint state space model

As an alternative we propose to estimate the following state space model

\[ x^{(r)}_t = Ax^{(r)}_{t-1} + w^{(r)}_t, \quad w^{(r)}_t \sim N(0, Q^{(r)}), \quad (3.5a) \]

\[ y^{(r)}_t = Cx^{(r)}_t + v^{(r)}_t, \quad v^{(r)}_t \sim N(0, R^{(r)}), \quad (3.5b) \]

\[ r = 1, 2, \ldots, M. \]

Matrix \( A \) is the same for all the records because the system is time invariant, and matrix \( C \) is also constant for the same reason and because the sensors location do not change from one record to another. On the other hand, \( x^{(r)}_t, w^{(r)}_t \) and \( v^{(r)}_t \) are record-dependent because the observed system outputs \( y^{(r)}_t \) and the non-recorded system inputs are different from record to record.

The unknown parameters of this model are

\[ \theta = \{ A, C, Q^{(r)}, R^{(r)}, \bar{x}^{(r)}_0, P^{(r)}_0 \}, \quad r = 1, 2, \ldots, M. \quad (3.6) \]

The result of this model is a single estimation for matrices \( A \) and \( C \), so a single value for each modal parameter is obtained

\[ \hat{A}, \hat{C} \Rightarrow \hat{\omega}_j, \hat{\zeta}_j, \hat{\psi}_j, \quad j = 1, 2, \ldots, n_s/2. \quad (3.7) \]

In this case, paring modes among setups is not needed anymore. And, what is more important, matrices \( A \) and \( C \) (and therefore the modal parameters) are estimated using all the available information. That is why we say all the information is combined optimally.

3.3 Maximum likelihood estimation of the proposed model

We use Maximum Likelihood Estimation and the Expectation Maximization algorithm to estimate the state space models (3.2) and (3.5). For model (3.2), the algorithm was described in Chapter 2. Here we are going to extend it to estimate state space models for multiple records, Eq. (3.5).

First we need to compute the likelihood. Given \( M \) records with \( N \) observed outputs each, we call \( Y^{(r)}_N = \{ y^{(r)}_1, \ldots, y^{(r)}_t, \ldots, y^{(r)}_N \} \), \( r = 1, \ldots, M \). Considering the \( M \) records are independent, the likelihood is computed as

\[ L_Y(\theta) = \prod_{r=1}^{M} L_{Y^{(r)}_N}(\theta^{(r)}), \quad (3.8) \]

where

\[ L_{Y^{(r)}_N}(\theta^{(r)}) = \prod_{t=1}^{N} f_{\theta^{(r)}}(y^{(r)}_t | Y^{(r)}_{t-1}) \quad (3.9) \]

(see Eq. (2.10)). \( f_{\theta}(\cdot) \) denotes a generic density function with parameters represented by \( \theta \). For the state space model (3.5)

\[ f_{\theta^{(r)}}(y^{(r)}_t | Y^{(r)}_{t-1}) = \frac{1}{(2\pi)^{n_s/2} |\Sigma^{(r)}_t|^{1/2}} \exp \left( -\frac{1}{2} \left( \epsilon^{(r)}_t \right)^T \left( \Sigma^{(r)}_t \right)^{-1} \epsilon^{(r)}_t \right). \quad (3.10) \]
Thus, the likelihood (3.8) is computed as
\[
L_{Y_N}(\theta^{(r)}) = \prod_{r=1}^{M} \frac{1}{(2\pi)^{N_{\text{na}}/2}} \prod_{t=1}^{N} \left| \Sigma_t^{(r)}(\theta^{(r)}) \right|^{-1/2} \cdot \exp \left( -\frac{1}{2} \left( \epsilon_t^{(r)}(\theta^{(r)}) \right)^T \left( \Sigma_t^{(r)}(\theta^{(r)}) \right)^{-1} \epsilon_t^{(r)}(\theta^{(r)}) \right),
\]
where it has been emphasized the dependence of the innovations on the parameters \(\theta^{(r)}\). In practice we generally work with the logarithm of the likelihood:
\[
\log L_{Y_N}(\theta) = l_{Y_N}(\theta) = -\frac{n_{\text{a}}N}{2} \log 2\pi - \frac{1}{2} \sum_{r=1}^{M} \sum_{t=1}^{N} \log \left| \Sigma_t^{(r)}(\theta^{(r)}) \right| - \frac{1}{2} \sum_{r=1}^{M} \sum_{t=1}^{N} \left( \epsilon_t^{(r)}(\theta^{(r)}) \right)^T \left( \Sigma_t^{(r)}(\theta^{(r)}) \right)^{-1} \epsilon_t^{(r)}(\theta^{(r)}).
\]
(3.11)

We are going now to maximize the likelihood, Eq. (3.11), using the EM algorithm. Consider we know the observed values \(Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)}\}\) and also the states \(X_N^{(r)} = \{x_1^{(r)}, x_2^{(r)}, \ldots, x_N^{(r)}\}\). The joint density function for record \(r\) is then given by:
\[
f_{\theta^{(r)}}(X_N^{(r)}, Y_N^{(r)}) = f_{x_0^{(r)}, p_0^{(r)}}(x_0^{(r)}) \prod_{t=1}^{N} f_{A,Q^{(r)}}(x_t^{(r)}|x_{t-1}^{(r)}) \prod_{t=1}^{N} f_{C,R^{(r)}}(y_t^{(r)}|x_t^{(r)})
\]
(3.12)
(see Eq. (2.18)). Under Gaussian assumption
\[
f_{x_0^{(r)}, p_0^{(r)}}(x_0^{(r)}) = \frac{1}{(2\pi)^{n_{\text{a}}/2}|P_0^{(r)}|^{1/2}} \exp \left( -\frac{1}{2} \left( x_0^{(r)} - \bar{x}_0^{(r)} \right)^T \left( P_0^{(r)} \right)^{-1} \left( x_0^{(r)} - \bar{x}_0^{(r)} \right) \right),
\]
\[
f_{A,Q^{(r)}}(x_t^{(r)}|x_{t-1}^{(r)}) = \frac{1}{(2\pi)^{n_{\text{a}}/2}|Q^{(r)}|^{1/2}} \exp \left( -\frac{1}{2} \left( x_t^{(r)} - Ax_{t-1}^{(r)} \right)^T \left( Q^{(r)} \right)^{-1} \left( x_t^{(r)} - Ax_{t-1}^{(r)} \right) \right),
\]
\[
f_{C,R^{(r)}}(y_t^{(r)}|x_t^{(r)}) = \frac{1}{(2\pi)^{n_{\text{a}}/2}|R^{(r)}|^{1/2}} \exp \left( -\frac{1}{2} \left( y_t^{(r)} - Cx_t^{(r)} \right)^T \left( R^{(r)} \right)^{-1} \left( y_t^{(r)} - Cx_t^{(r)} \right) \right).
\]

Thus, if we consider \(M\) independent registers, the joint density function \(f_{\theta}(X_N, Y_N)\) will be the product of individual ones
\[
f_{\theta}(X_N, Y_N) = \prod_{r=1}^{M} f_{\theta^{(r)}}(X_N^{(r)}, Y_N^{(r)}).
\]
(3.13)

This is known as the complete data likelihood, \(L_{X_N,Y_N}(\theta) = f_{\theta}(X_N, Y_N)\). In practice we generally work with the log-likelihood, because information is combined by addition and it can be written as a sum of the log-likelihood of each individual record:
\[
\log L_{X_N,Y_N}(\theta) = \sum_{r=1}^{M} \log f_{\theta^{(r)}}(X_N^{(r)}, Y_N^{(r)}).
\]
(3.14)
3.3 Maximum likelihood estimation of the proposed model

The log-likelihood of record \(r\) can be written as the sum of three uncoupled functions

\[
I_{X_N,r}^{(r)}(\theta^{(r)}) = -\frac{1}{2} \left[ I_1(x_0^{(r)}, P_0^{(r)}) + I_2(A, Q^{(r)}) + I_3(C, R^{(r)}) \right],
\]

(3.15)

where, ignoring constants, are

\[
l_1(x_0^{(r)}, P_0^{(r)}) = \log |P_0^{(r)}| + \left( x_0^{(r)} - \bar{x}_0^{(r)} \right)^T \left( P_0^{(r)} \right)^{-1} \left( x_0^{(r)} - \bar{x}_0^{(r)} \right),
\]

(3.16)

\[
l_2(A, Q^{(r)}) = N \log |Q^{(r)}| + \sum_{t=1}^{N} \left( x_t^{(r)} - Ax_{t-1}^{(r)} \right)^T \left( Q^{(r)} \right)^{-1} \left( x_t^{(r)} - Ax_{t-1}^{(r)} \right),
\]

(3.17)

\[
l_3(C, R^{(r)}) = N \log |R^{(r)}| + \sum_{t=1}^{N} \left( y_t^{(r)} - Cx_t^{(r)} \right)^T \left( R^{(r)} \right)^{-1} \left( y_t^{(r)} - Cx_t^{(r)} \right).
\]

(3.18)

The objective now is to maximize the log-likelihood (3.14) using the Expectation Maximization algorithm. According to Section 2.3, the maximum of (3.14) will give us the maximum of Eq. (3.11), and therefore, the maximum likelihood estimation of the parameters.

3.3.1 E-step: expectation step

The log-likelihood (3.14) cannot be computed because the states \(X_N^{(r)}\) are unknown (in fact, the states are unobserved quantities). The method proposes to replace them with their expected values. First, we need the following definition

**Definition 3.1.** Given the output data \(Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)}\}\), \(r = 1, \ldots, M\), it is defined

\[
x_t^{N,(r)} = E \left[ x_t^{(r)} | Y_N^{(r)} \right]
\]

(3.19)

\[
P_t^{N,(r)} = E \left[ \left( x_t^{(r)} - x_t^{N,(r)} \right) \left( x_t^{(r)} - x_t^{N,(r)} \right)^T | Y_N^{(r)} \right]
\]

(3.20)

\[
P_{t,t-1}^{N,(r)} = E \left[ \left( x_t^{(r)} - x_t^{N,(r)} \right) \left( x_{t-1}^{(r)} - x_{t-1}^{N,(r)} \right)^T | Y_N^{(r)} \right]
\]

(3.21)

where \(E[\square]\) is the conditional expected operator.

Now we are able to compute the expected value of the log-likelihood:

**Property 3.1.** Given the observed vectors \(Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)}\}\), \(r = 1, \ldots, M\), and a value for the parameters \(\theta_j\), then

\[
E[l_{X_N,Y_N}(\theta) | Y_N, \theta_j] = \sum_{r=1}^{M} E[l_{X_N,Y_N}^{(r)}(\theta^{(r)}) | Y_N^{(r)}, \theta_j^{(r)}]
\]

(3.22)

\[
= \sum_{r=1}^{M} \left( E[l_1(x_0^{(r)}, P_0^{(r)}) | Y_N^{(r)}, \theta_j^{(r)}] + E[l_2(A, Q^{(r)}) | Y_N^{(r)}, \theta_j^{(r)}] + E[l_3(C, R^{(r)}) | Y_N^{(r)}, \theta_j^{(r)}] \right)
\]

with

\[
E \left[ l_1(x_0^{(r)}, P_0^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] = \\
= \log |P_0^{(r)}| + \text{tr} \left( \left( P_0^{(r)} \right)^{-1} \left( x_0^{N,(r)} - x_0^{(r)} \right) \left( x_0^{N,(r)} - x_0^{(r)} \right)^T \right),
\]

(3.23)
\[
E[l_2(A, Q^{(r)})|Y^{(r)}_N, \theta_j^{(r)}] = N \log |Q^{(r)}| + \text{tr} \left( (Q^{(r)})^{-1} \left[ S_{xx}^{(r)} - S_{xb}^{(r)} A^T - A S_{bb}^{(r)} A^T \right] \right),
\]

(3.24)

\[
E[l_3(C, R^{(r)})|Y^{(r)}_N, \theta_j^{(r)}] = N \log |R^{(r)}| + \text{tr} \left( (R^{(r)})^{-1} \left[ S_{yy}^{(r)} - S_{yx}^{(r)} C^T - C S_{xy}^{(r)} C^T \right] \right),
\]

(3.25)

where \( \text{tr}(\square) \) is the matrix trace operator and

\[
S_{xx}^{(r)} = \sum_{t=1}^{N} \left( P_t^{N,r} + x_t^{N,r} \left( x_t^{N,r} \right)^T \right),
\]

(3.26)

\[
S_{bb}^{(r)} = \sum_{t=1}^{N} \left( P_{t-1}^{N,r} + x_{t-1}^{N,r} \left( x_{t-1}^{N,r} \right)^T \right),
\]

(3.27)

\[
S_{xb}^{(r)} = \sum_{t=1}^{N} \left( P_{t-1}^{N,r} + x_{t-1}^{N,r} \left( x_{t-1}^{N,r} \right)^T \right),
\]

(3.28)

\[
S_{yx}^{(r)} = \sum_{t=1}^{N} y_t^{r} \left( x_t^{N,r} \right)^T, \quad S_{xy}^{(r)} = \left( S_{yx}^{(r)} \right)^T,
\]

(3.29)

\[
S_{yy}^{(r)} = \sum_{t=1}^{N} y_t^{r} \left( y_t^{r} \right)^T.
\]

(3.30)

Proof. The proof of the Eqs. (3.23), (3.24) and (3.25) is straightforward taking into account Prop. 2.2.

Finally, it is important to remark that \( x_t^{N,r}, P_t^{N,r} \) and \( P_{t-1}^{N,r} \) are computed applying the Kalman Filter and the Kalman Smoother (Properties A.4, A.5 and A.6) to each record \( r \).

3.3.2 M-step: maximization step

Maximizing \( E[l_{X,Y_N}(\theta)|Y_N, \theta_j] \) with respect to the parameters \( \theta \), constitutes the M-step.

Property 3.2. The maximum of \( E[l_{X,Y_N}(\theta)|Y_N, \theta_j] \) (Eq. (3.22)) is attained at

\[
\bar{x}_0^{(r)} = x_0^{N,r}, \quad r = 1, \ldots, M,
\]

(3.31)

\[
P_0^{(r)} = P_0^{N,r}, \quad r = 1, \ldots, M,
\]

(3.32)

\[
\text{vec}(A) = \left[ \sum_{r=1}^{M} \left( S_{bb}^{(r)} \otimes \left( Q^{(r)} \right)^{-1} \right) \right]^{-1} \sum_{r=1}^{M} \text{vec} \left( \left( Q^{(r)} \right)^{-1} S_{xb}^{(r)} \right),
\]

(3.33)

\[
Q^{(r)} = \frac{1}{N} \left( S_{xx}^{(r)} - S_{xb}^{(r)} A^T - A S_{bb}^{(r)} A^T \right), \quad r = 1, \ldots, M,
\]

(3.34)
3.4 Starting values for the EM algorithm

The selection of initial points for the EM algorithm in OMA is always a difficult problem. Besides, the more complex the model is, the more difficult the starting point is to build. In our opinion, any starting point needs to be:

1. fast and easy to build;
2. robust, so the EM algorithm does not have numerical problems in the iterations;
3. it has to take advantage the information provided by the measurements.

We think using SSI algorithm can help us to build the starting point for model (3.5) and meets all the above requirements. In this Chapter we propose to build the starting point \( \theta_0 = \left( A_0, C_0, Q_0^{(r)}, R_0^{(r)}, \bar{x}_0^{(r)}, P_0^{(r)} \right) \) as follows:
• First, we apply SSI to one of the \( M \) available records, let us say record \( s \). We call these parameters \( A_{SSI}^{(s)} \), \( C_{SSI}^{(s)} \), \( Q_{SSI}^{(s)} \), and \( R_{SSI}^{(s)} \).

• \( A_0 = A_{SSI}^{(s)} \). We take advantage of the information given by the eigenvalues of \( A_{SSI}^{(s)} \). They are useful to the EM algorithm to convergence to the right solution.

• \( C_0 = C_{SSI}^{(s)} \). The comments made for \( A_0 \) are also valid for \( C_0 \).

• \( Q_0^{(r)} = Q_{SSI}^{(s)} \) for all \( r \). The result is a positive definite matrix (this is important for the EM), and it is a good approximation to all covariance matrices.

• \( R_0^{(r)} = R_{SSI}^{(s)} \) for all \( r \). Another positive definite matrix and adequate for the outputs.

• \( \bar{x}_{0,0}^{(r)} \) and \( P_{0,0}^{(r)} \) are taken matrices of zeros.

3.5 Example 1: simulated data

An eight DOF simulated structure is used in this section to validate the proposed method. It consists in 8 masses and 9 springs and dash-pots (Figure 3.1(a)). The values chosen for the mass matrix \( M \), the damping matrix \( C \) and the stiffness matrix \( K \) are the following:

\[
K = \begin{bmatrix}
2400 & -1600 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1600 & 4000 & -2400 & 0 & 0 & 0 & 0 & 0 \\
0 & -2400 & 5600 & -3200 & 0 & 0 & 0 & 0 \\
0 & 0 & -3200 & 7200 & -4000 & 0 & 0 & 0 \\
0 & 0 & 0 & -4000 & 8800 & -4800 & 0 & 0 \\
0 & 0 & 0 & 0 & -4800 & 10400 & -5600 & 0 \\
0 & 0 & 0 & 0 & 0 & -5600 & 12000 & -6400 \\
0 & 0 & 0 & 0 & 0 & 0 & -6400 & 13600
\end{bmatrix} \left( \frac{N}{m} \right) ;
\]

(3.38)

\( M \) is equal to the identity matrix with eight rows and columns, and \( C \) is built assuming Rayleigh damping by mean of

\[
C = 0.680M + 1.743 \cdot 10^{-4}K \left( \frac{N \cdot s}{m} \right).
\]

(3.39)

The natural frequencies, mode shapes and damping ratios corresponding to matrices \( M \), \( C \) and \( K \), are presented in Figure 3.1(b).

The output signals have been obtained using:

• The acceleration of the eight DOFs has been generated four times, so we have four different records (\( M = 4 \)).
  - Record 1: The input is Gaussian white noise, \( u_t \sim N(0, \sigma_u^2 = 1) \), and has been applied to DOFs 1, 2, 3, 4.
  - Record 2: The input is Gaussian white noise, \( u_t \sim N(0, \sigma_u^2 = 4) \), and has been applied to DOFs 5, 6, 7, 8.
3.5 Example 1: simulated data

(a)

(b)

Figure 3.1: (a) Model of the simulated structure; (b) Natural frequencies, damping ratios and mode shapes of the simulated structure.

- Record 3: The input is Gaussian white noise, \( u_t \sim N(0, \sigma_u^2 = 9) \), and has been applied to DOFs 1, 3, 5, 7.
- Record 4: The input is Gaussian white noise, \( u_t \sim N(0, \sigma_u^2 = 16) \), and has been applied to DOFs 2, 4, 6, 8.

- Sampling frequency \( f_s = 50 \text{ Hz} \). Total duration of signals, 100 seconds (5000 time steps).
- The observed value is the sum of the structure response and a sensor Gaussian noise with variance equal to the 20% of the largest acceleration response variance.

We have analysed each record individually using the EM algorithm as shown in Chapter 2. In this case we have built the starting point for the EM using the parameters estimated by the SSI (first, we apply SSI to each record and then we apply EM using
Table 3.1: Natural frequencies, damping ratios and MAC values estimated from multiple records (simulated data) using the EM algorithm. The theoretical values are also included for comparison.

<table>
<thead>
<tr>
<th>Mode</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Natural frequencies (Hz)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td>2.94</td>
<td>5.87</td>
<td>8.60</td>
<td>11.19</td>
<td>13.78</td>
<td>16.52</td>
<td>19.54</td>
<td>23.12</td>
</tr>
<tr>
<td>Record 1</td>
<td>2.92</td>
<td>5.90</td>
<td>8.60</td>
<td>11.19</td>
<td>13.75</td>
<td>16.54</td>
<td>19.56</td>
<td>23.04</td>
</tr>
<tr>
<td>Record 2</td>
<td>2.93</td>
<td>5.90</td>
<td>8.62</td>
<td>11.16</td>
<td>13.79</td>
<td>17.25</td>
<td>19.59</td>
<td>23.04</td>
</tr>
<tr>
<td>Record 3</td>
<td>2.95</td>
<td>5.87</td>
<td>8.63</td>
<td>11.19</td>
<td>13.80</td>
<td>16.48</td>
<td>19.47</td>
<td>23.13</td>
</tr>
<tr>
<td>Record 4</td>
<td>2.94</td>
<td>5.85</td>
<td>8.59</td>
<td>11.19</td>
<td>13.75</td>
<td>16.46</td>
<td>19.54</td>
<td>23.10</td>
</tr>
<tr>
<td>Mean</td>
<td>2.96</td>
<td>5.90</td>
<td>8.62</td>
<td>11.20</td>
<td>13.78</td>
<td>16.53</td>
<td>19.55</td>
<td>23.10</td>
</tr>
<tr>
<td>Model (3.5)</td>
<td>2.96</td>
<td>5.90</td>
<td>8.62</td>
<td>11.20</td>
<td>13.78</td>
<td>16.53</td>
<td>19.55</td>
<td>23.10</td>
</tr>
</tbody>
</table>

| **Damping ratios (%)** | | | | | | | | |
| Theoretical | 2.00 | 1.24 | 1.10 | 1.10 | 1.15 | 1.23 | 1.35 | 1.50 |
| Record 1 | 2.36 | 1.41 | 1.13 | 1.25 | 1.32 | 1.46 | 1.50 | 1.50 |
| Record 2 | 2.66 | 1.33 | 0.79 | 1.22 | 1.22 | 1.34 | 1.54 | 1.52 |
| Record 3 | 2.86 | 0.90 | 1.03 | 1.10 | 1.13 | 1.22 | 1.51 | 1.34 |
| Record 4 | 2.00 | 1.16 | 1.00 | 1.21 | 1.48 | 1.45 | 1.33 | 1.74 |
| Mean | 2.08 | 1.75 | 1.14 | 1.07 | 1.18 | 1.13 | 1.35 | 1.37 |
| Model (3.5) | 2.24 | 1.73 | 1.14 | 1.09 | 1.18 | 1.12 | 1.34 | 1.35 |

| **MAC values between theoretical and estimated modes shapes** | | | | | | | | |
| Record 1 | 1.000 | 0.999 | 0.999 | 0.997 | 0.997 | 0.996 | 0.998 | 0.924 |
| Record 2 | 0.996 | 0.999 | 0.999 | 0.999 | 0.927 | 0.002 | 0.998 | 0.972 |
| Record 3 | 0.999 | 0.925 | 0.999 | 0.995 | 1.000 | 0.997 | 0.997 | 1.000 |
| Record 4 | 1.000 | 0.996 | 0.985 | 0.999 | 0.994 | 0.997 | 0.998 | 1.000 |
| Mean | 0.998 | 0.980 | 0.996 | 0.997 | 0.980 | 0.748 | 0.998 | 0.974 |
| Model (3.5) | 1.000 | 1.000 | 1.000 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 |

In general terms, the modal parameters have been identified very well in the four records. We see that the natural frequencies are the parameters estimated with lower bias and, on the contrary, the damping ratios are the parameters worst estimated. It is specially remarkable that mode 6 in record 2 has not been properly estimated.

Then we have estimated the proposed model (3.5) using the EM algorithm as described in Section 3.3. The starting point for the EM algorithm was built applying SSI to simulated record 1 (see Section 3.4). The obtained modal parameters are shown in Table 3.1. In the whole, these estimates are more accurate than the mean values computed from the individual estimates. Even the estimate of mode 6 from record 2 has been corrected.

Other important conclusion is that the results obtained using model (3.5) do not coincide with the mean values of the individual EM. This is logical because the result of model (3.5) is the most likelihood value taking into account the whole data (joint estimation).
3.6 Example 2: Tablate II Bridge

Figure 3.2 shows the Tablate II Bridge, located on a highway in the south of Granada (Spain). The structure of the bridge consists of a concrete deck resting on a metal arch with a span of 128 meters. The bridge has 3 accelerometers (longitudinal, transversal and vertical directions) permanently installed at the center of the arch, on the side walk (Figure 3.3). The system automatically saves the measured accelerations of the bridge when any of the three signals exceeds a certain threshold.

The described recording system was installed to monitor the bridge over time and produces around 20 different records a day. The data, mainly due to wind and traffic loading, were sampled at a rate of 250 Hz.

In order to illustrate the methodology described in this Chapter, we have chosen four different records obtained on April, 2009. Many researchers have reported the influence of temperature in modal parameters (for instance [65]). Because of model (3.5) assume the system is time invariant, we have chosen data recorded at the same temperature. We have also chosen data from different days to be sure the ambient loads are totally different. Table 3.2 summarizes the characteristics of the records we have chosen.

Since we only have data from three sensors, the mode shapes will be poorly estimated, so the analysis will focus only in the modal frequencies and the damping ratios. Table 3.3
Table 3.2: Multiple records from Tablate II Bridge used in the analysis.

<table>
<thead>
<tr>
<th>Record name</th>
<th>Date (yy-mm-dd)</th>
<th>Shot channel</th>
<th>Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record 1 DN187</td>
<td>2009-04-11</td>
<td>Vertical</td>
<td>24.5</td>
</tr>
<tr>
<td>Record 2 DN243</td>
<td>2009-04-19</td>
<td>Transversal</td>
<td>24.5</td>
</tr>
<tr>
<td>Record 3 DN313</td>
<td>2009-04-24</td>
<td>Transversal</td>
<td>24.5</td>
</tr>
<tr>
<td>Record 4 DN329</td>
<td>2009-04-25</td>
<td>Longitudinal</td>
<td>24.5</td>
</tr>
</tbody>
</table>

shows both the results obtained with the EM applied to each individual record and using the proposed joint model given by Eq. (3.5) (the starting point for each case has been built as commented in Section 3.4). We have estimated ten modes (the order of the state space models is 20). For the individual analysis, one can see that the frequencies and damping ratios estimated using record 1 are different to the ones estimated using records 2, 3 and 4. Besides, with real data the modes of vibration have close natural frequencies and it is complex to decide if similar estimates correspond to the same physical mode. Therefore, the mean values are not always possible to compute.

Six of the ten modes seem to appear in the four records: around 1.05 Hz, 1.10 Hz, 1.65 Hz, 2.80 Hz, 2.90 Hz, 4.85 Hz. Other estimates are present in three records: around 3.6 Hz, 4.2 Hz, and 4.4 Hz, and the rest are present in one or two records. In any case, this pairing process is far from being easy.

The results obtained using model (3.5) for state space model order equal to 20 are included in the last two columns of Table 3.3 (the corresponding stabilization diagram is plotted in Figure 3.4). Frequencies obtained with this model (the joint analysis) correspond to the values that are most repeated in the individual analysis. All modal frequencies appearing in three or more individual records also appear in the joint estimation. In general, when a frequency appears only in a single record, it does not appear in the joint estimation.

In the analysis of individual records, the estimated damping ratios present higher variability than the natural frequencies. Some values are higher than 5%, and could be considered over acceptable limits. However, in the joint estimate, damping values are lower and within the range of admissibility (except mode with frequency 2.874 Hz, with damping ratio of 9.33%). Another detail that catches the eye is that the overall damping value is significantly lower than the average values of the individual ones.

The principal conclusion is that joint estimation results do not match the individual mean values, neither for the frequencies nor for the damping. The joint analysis presented here combines information efficiently, giving more weight to those estimates that are more likely. The joint estimation of a mode includes the information from all records, even those that have not obtained the mode in question in their individual analysis.

3.7 Conclusions

In this Chapter we propose a complete methodology for the joint estimation of the model parameters. The joint estimation of the structure’s parameters combines information optimally: the main modes that are repeated in different records are detected more clearly while modes specific to some records tend to blur in the joint analysis.

Joint estimation results do not match the individual mean values, neither for the frequencies nor for the damping. The joint analysis presented here combines information
Figure 3.4: Stabilization diagram for Tablate Bridge using the joint state space model. The used symbols are: “⊕” for stable pole and “+” for unstable pole. A mode is stable if \( \varepsilon_\omega = 0.02, \varepsilon_\zeta = 0.03, \varepsilon_{MAC} = 0.10. \)

more efficiently, giving more weight to those estimates that are more likely. It is important to note that the joint estimation of a mode includes the information from all records, even those that have not obtained the mode in question in their individual analysis.

In conclusion, estimating parameters using all the recorded information and the EM algorithm, has numerous advantages in the modal analysis of a structure, and resolves the difficulties in combining the individual solutions coming from different records.

3.8 Appendix

Proof of Property 3.2

The maximum is attained deriving and equating to zero:

\[
\frac{\partial}{\partial \bar{x}_0} \text{E}[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \text{E}[l_1(\bar{x}_0, P_0)|Y_N^{(r)}, \theta_j^{(r)}] = 0;
\]

\[
\text{E} \left[ l_1(\bar{x}_0, P_0)|Y_N^{(r)}, \theta_j^{(r)} \right] = \log |P_0^{(r)}| + \text{tr} \left( \left( P_0^{(r)} \right)^{-1} \left[ P_0^{N,(r)} + (x_0^{N,(r)} - \bar{x}_0^{(r)}) (x_0^{N,(r)} - \bar{x}_0^{(r)})^T \right] \right);
\]
Table 3.3: Modal frequencies and damping ratios for the selected records: applying the EM to the individual records and model (3.5) for the joint estimation.

<table>
<thead>
<tr>
<th></th>
<th>Record 1</th>
<th>Record 2</th>
<th>Record 3</th>
<th>Record 4</th>
<th>Model (3.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>f (Hz)</td>
<td>ζ (%)</td>
<td>f (Hz)</td>
<td>ζ (%)</td>
<td>f (Hz)</td>
<td>ζ (%)</td>
</tr>
<tr>
<td>1,058</td>
<td>1,57</td>
<td>1,053</td>
<td>1,048</td>
<td>1,045</td>
<td>1,050</td>
</tr>
<tr>
<td>1,116</td>
<td>0,83</td>
<td>1,092</td>
<td>1,108</td>
<td>1,095</td>
<td>1,101</td>
</tr>
<tr>
<td>1,637</td>
<td>1,28</td>
<td>1,677</td>
<td>1,688</td>
<td>1,634</td>
<td>1,672</td>
</tr>
<tr>
<td>2,792</td>
<td>2,01</td>
<td>2,779</td>
<td>2,740</td>
<td>1,896</td>
<td>2,801</td>
</tr>
<tr>
<td>2,902</td>
<td>0,35</td>
<td>2,923</td>
<td>2,755</td>
<td>2,821</td>
<td>2,874</td>
</tr>
<tr>
<td>3,010</td>
<td>8,02</td>
<td>3,589</td>
<td>2,912</td>
<td>2,920</td>
<td>2,903</td>
</tr>
<tr>
<td>3,574</td>
<td>0,55</td>
<td>4,058</td>
<td>3,736</td>
<td>3,021</td>
<td>3,588</td>
</tr>
<tr>
<td>4,093</td>
<td>3,26</td>
<td>4,323</td>
<td>3,900</td>
<td>4,161</td>
<td>4,166</td>
</tr>
<tr>
<td>4,435</td>
<td>0,44</td>
<td>4,844</td>
<td>4,455</td>
<td>4,934</td>
<td>4,427</td>
</tr>
<tr>
<td>4,755</td>
<td>2,35</td>
<td>5,212</td>
<td>4,869</td>
<td>5,487</td>
<td>4,883</td>
</tr>
</tbody>
</table>

Joint analysis of multiple records

\[
\frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \log |P_0^{(r)}| = 0;
\]

\[
\frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \text{tr} \left( \left( P_0^{(r)} \right)^{-1} P_0^{N,(r)} \right) = 0;
\]

\[
\frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \text{tr} \left( \left( P_0^{(r)} \right)^{-1} x_0^{N,(r)} \left( x_0^{N,(r)} \right)^T \right) = 0;
\]

\[
\frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \text{tr} \left( - \left( P_0^{(r)} \right)^{-1} x_0^{N,(r)} \left( \bar{x}_0^{(r)} \right)^T \right) = - \left( P_0^{(r)} \right)^{-1} x_0^{N,(r)};
\]

\[
\frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \text{tr} \left( - \left( P_0^{(r)} \right)^{-1} \bar{x}_0^{(r)} \left( x_0^{N,(r)} \right)^T \right) = - \left( P_0^{(r)} \right)^{-T} x_0^{N,(r)};
\]

\[
\frac{\partial}{\partial \bar{x}_0} \sum_{r=1}^{M} \text{tr} \left( \left( P_0^{(r)} \right)^{-1} \bar{x}_0^{(r)} \left( \bar{x}_0^{(r)} \right)^T \right) = \left( P_0^{(r)} \right)^{-T} \bar{x}_0^{(r)} + \left( P_0^{(r)} \right)^{-1} \bar{x}_0^{(r)} = 0
\]

Equating to zero

\[- \left( P_0^{(r)} \right)^{-1} x_0^{N,(r)} - \left( P_0^{(r)} \right)^{-T} x_0^{N,(r)} + \left( P_0^{(r)} \right)^{-T} \bar{x}_0^{(r)} + \left( P_0^{(r)} \right)^{-1} \bar{x}_0^{(r)} = 0 \]

\[\Rightarrow \bar{x}_0^{(r)} = x_0^{N,(r)}.\]

**•** \( \frac{\partial}{\partial P_0^{(r)}} \mathbb{E}[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] = 0 \) \( \Rightarrow \frac{\partial}{\partial P_0^{(r)}} \sum_{r=1}^{M} \mathbb{E}[l_1(\bar{x}_0^{(r)}, P_0^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] = 0; \)

\[
\mathbb{E} \left[ l_1(\bar{x}_0^{(r)}, P_0^{(r)})|Y_N^{(r)}, \theta_j^{(r)} \right] = \log |P_0^{(r)}| + \text{tr} \left( \left( P_0^{(r)} \right)^{-1} \left[ P_0^{N,(r)} + \left( x_0^{N,(r)} - \bar{x}_0^{(r)} \right)(x_0^{N,(r)} - \bar{x}_0^{(r)})^T \right] \right).
\]
We have obtained that $\bar{x}_0^{(r)} = x_0^{N,(r)}$. Substituting above
\[
E \left[ l_1(\bar{x}_0^{(r)}, P_0^{(r)})|Y_0^{(r)}, \theta_j^{(r)} \right] = \log |P_0^{(r)}| + \text{tr} \left( (P_0^{(r)})^{-1} P_0^{N,(r)} \right);
\]
\[
\frac{\partial}{\partial P_0^{(r)}} \sum_{r=1}^{M} \log |P_0^{(r)}| = (P_0^{(r)})^{-1};
\]
\[
\frac{\partial}{\partial P_0^{(r)}} \sum_{r=1}^{M} \text{tr} \left( (P_0^{(r)})^{-1} P_0^{N,(r)} \right) = - \left( (P_0^{(r)})^{-1} P_0^{N,(r)} (P_0^{(r)})^{-1} \right)^T;
\]
Equating to zero
\[
\left( P_0^{(r)} \right)^{-1} - \left( (P_0^{(r)})^{-1} P_0^{N,(r)} (P_0^{(r)})^{-1} \right)^T = 0 \Rightarrow P_0^{(r)} = P_0^{N,(r)}.
\]

- $\frac{\partial}{\partial A} E[l_{X_N,Y_N}^{(r)}|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial A} \sum_{r=1}^{M} E[l_2(A,Q^{(r)}))]|Y_N, \theta_j^{(r)}] = 0$;
\[
E \left[ l_2(A,Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)} \right] = \sum_{r=1}^{M} \text{tr} \left( (Q^{(r)})^{-1} \left[ S_{xx}^{(r)} - S_{zb}^{(r)} A^T - A S_{bx}^{(r)} + A S_{bb}^{(r)} A^T \right] \right);
\]
\[
\frac{\partial}{\partial A} \sum_{r=1}^{M} N \log |Q^{(r)}| = 0;
\]
\[
\frac{\partial}{\partial A} \sum_{r=1}^{M} \text{tr} \left( (Q^{(r)})^{-1} S_{xx}^{(r)} \right) = 0;
\]
\[
\frac{\partial}{\partial A} \sum_{r=1}^{M} \text{tr} \left( - (Q^{(r)})^{-1} S_{zb}^{(r)} A^T \right) = - \sum_{r=1}^{M} (Q^{(r)})^{-1} S_{zb}^{(r)};
\]
\[
\frac{\partial}{\partial A} \sum_{r=1}^{M} \text{tr} \left( - (Q^{(r)})^{-1} A S_{bx}^{(r)} \right) = - \sum_{r=1}^{M} (Q^{(r)})^{-T} \left( S_{bx}^{(r)} \right)^T;
\]
\[
\frac{\partial}{\partial A} \sum_{r=1}^{M} \text{tr} \left( (Q^{(r)})^{-1} A S_{bb}^{(r)} A^T \right) = \sum_{r=1}^{M} (Q^{(r)})^{-T} A \left( S_{bb}^{(r)} \right)^T + \sum_{r=1}^{M} (Q^{(r)})^{-1} A S_{bb}^{(r)};
\]
Equating to zero
\[
\sum_{r=1}^{M} (Q^{(r)})^{-1} A S_{bb}^{(r)} = \sum_{r=1}^{M} (Q^{(r)})^{-1} S_{zb}^{(r)}
\]
(3.40)
and taking the vector operator of a matrix $\text{vec}(\bullet)$, we can write the linear system of equations to compute $A$:
\[
\sum_{r=1}^{M} \left( S_{bb}^{(r)} \otimes (Q^{(r)})^{-1} \right) \text{vec}(A) = \sum_{r=1}^{M} \text{vec} \left( (Q^{(r)})^{-1} S_{zb}^{(r)} \right).
\]
where $\otimes$ is the Kronecker product. Thus

\[
\text{vec}(A) = \left[ \sum_{r=1}^{M} \left( S_{bb}^{(r)} \otimes (Q^{(r)})^{-1} \right) \right]^{-1} \sum_{r=1}^{M} \text{vec} \left( (Q^{(r)})^{-1} S_{xb}^{(r)} \right).
\]

\[\bullet \quad \frac{\partial}{\partial Q^{(r)}} E[l_{2}(A, Q^{(r)})|Y_{N}^{(r)}, \theta_{j}^{(r)}] = 0 \Rightarrow \frac{\partial}{\partial Q^{(r)}} \sum_{r=1}^{M} E[l_{2}(A, Q^{(r)})|Y_{N}^{(r)}, \theta_{j}^{(r)}] = 0;\]

\[E \left[ l_{2}(A, Q^{(r)})|Y_{N}^{(r)}, \theta_{j}^{(r)} \right] = N \log |Q^{(r)}| + \text{tr} \left( (Q^{(r)})^{-1} \left[ S_{xx}^{(r)} - S_{xb}^{(r)} A^{T} - AS_{bx}^{(r)} + AS_{bb}^{(r)} A^{T} \right] \right) = - \left( (Q^{(r)})^{-1} \left[ S_{xx}^{(r)} - S_{xb}^{(r)} A^{T} - AS_{bx}^{(r)} + AS_{bb}^{(r)} A^{T} \right] \right)^{T} \left( Q^{(r)} \right)^{-1} T;\]

Equating to zero

\[N \left( Q^{(r)} \right)^{-T} - \left( (Q^{(r)})^{-1} \left[ S_{xx}^{(r)} - S_{xb}^{(r)} A^{T} - AS_{bx}^{(r)} + AS_{bb}^{(r)} A^{T} \right] \right)^{T} \left( Q^{(r)} \right)^{-1} = 0;\]

\[\Rightarrow Q^{(r)} = \frac{1}{N} \left( S_{xx}^{(r)} - S_{xb}^{(r)} A^{T} - AS_{bx}^{(r)} + AS_{bb}^{(r)} A^{T} \right).\]

\[\bullet \quad \frac{\partial}{\partial C} E[l_{2}(C, R)|Y_{N}, \theta_{j}] = 0 \Rightarrow \frac{\partial}{\partial C} \sum_{r=1}^{M} E[l_{2}(C, R)|Y_{N}, \theta_{j}] = 0;\]

\[E \left[ l_{2}(C, R)|Y_{N}^{(r)}, \theta_{j}^{(r)} \right] = N \log |R^{(r)}| + \text{tr} \left( (R^{(r)})^{-1} \left[ S_{yy}^{(r)} - S_{yx}^{(r)} C^{T} - CS_{xy}^{(r)} + CS_{xx}^{(r)} C^{T} \right] \right);\]

\[\frac{\partial}{\partial C} \sum_{r=1}^{M} N \log |R^{(r)}| = 0;\]

\[\frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} S_{yy}^{(r)} \right) = 0;\]

\[\frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} S_{yx}^{(r)} C^{T} \right) = - \sum_{r=1}^{M} (R^{(r)})^{-1} S_{yx}^{(r)};\]

\[\frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} CS_{xy}^{(r)} \right) = - \sum_{r=1}^{M} (R^{(r)})^{-T} (S_{yx}^{(r)})^{T}.\]
\[
\frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} C (S^{(r)}_{xx}) C^T \right) = \sum_{r=1}^{M} (R^{(r)})^{-T} C \left( S^{(r)}_{xx} \right)^T + \sum_{r=1}^{M} (R^{(r)})^{-1} CS^{(r)}_{xx}.
\]

Equating to zero
\[
\sum_{r=1}^{M} \left( S^{(r)}_{xx} \otimes (R^{(r)})^{-1} \right) \text{vec}(C) = \sum_{r=1}^{M} \text{vec} \left( (R^{(r)})^{-1} S^{(r)}_{yx} \right) \Rightarrow
\]
\[
\text{vec}(C) = \left[ \sum_{r=1}^{M} \left( S^{(r)}_{xx} \otimes (R^{(r)})^{-1} \right) \right]^{-1} \sum_{r=1}^{M} \text{vec} \left( (R^{(r)})^{-1} S^{(r)}_{yx} \right).
\]

\[
\frac{\partial}{\partial R^{(r)}} E[l_{X_N,Y_N}(\theta)|Y_N,\theta_j] = 0 \Rightarrow \frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} E[l_3(C, R^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] = 0
\]
\[
E \left[l_3(C, R^{(r)})|Y_N^{(r)}, \theta_j^{(r)} \right] =
\]
\[
= N \log |R^{(r)}| + \text{tr} \left( (R^{(r)})^{-1} \left[ S^{(r)}_{yy} - S^{(r)}_{yx} C^T - CS^{(r)}_{xy} + CS^{(r)}_{xx} C^T \right] \right);
\]
\[
\frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} N \log |R^{(r)}| = N \left( R^{(r)} \right)^{-T};
\]
\[
\frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} \left[ S^{(r)}_{yy} - S^{(r)}_{yx} C^T - CS^{(r)}_{xy} + CS^{(r)}_{xx} C^T \right] \right) =
\]
\[
= - \left( (R^{(r)})^{-1} \left[ S^{(r)}_{yy} - S^{(r)}_{yx} C^T - CS^{(r)}_{xy} + CS^{(r)}_{xx} C^T \right] \left( R^{(r)} \right)^{-1} \right)^T;
\]

Equating to zero
\[
N \left( R^{(r)} \right)^{-T} - \left( (R^{(r)})^{-1} \left[ S^{(r)}_{yy} - S^{(r)}_{yx} C^T - CS^{(r)}_{xy} + CS^{(r)}_{xx} C^T \right] \left( R^{(r)} \right)^{-1} \right)^T = 0,
\]
\[
R^{(r)} = \frac{1}{N} \left( S^{(r)}_{yy} - S^{(r)}_{yx} C^T - CS^{(r)}_{xy} + CS^{(r)}_{xx} C^T \right).
\]
Chapter 4

Analysis of multiple setups of sensors

4.1 Introduction

The modal parameters estimated from ambient vibration measurements comprise natural frequencies, damping ratios and mode shapes. Natural frequencies and damping ratios might be computed using the data recorded by a single sensor, but the mode shapes can be only estimated at those points where a sensor is placed (see Appendix B). When the number of available sensors is lower than the number of tested points (because the structure is large or because the resolution required in the mode shapes is high), it is common practice to perform non-simultaneous measurement setups changing the position of the sensors among setups [89]. Since the mode shapes estimated in each setup can not be scaled in an absolute sense (e.g. to unity modal mass), some sensors have permanent positions because these fixed or reference sensors are needed to glue or merge the mode shapes estimated at each setup (or partial mode shapes) into global mode shapes. The rest of the sensors change their position in the structure from one setup to the next, so that different parts of the global mode shapes can be estimated.

The algorithms available to estimate or identified the modal parameters from ambient vibration measurements (or OMA), were originally developed to process the data of one setup of sensors individually. That is why it is usual to process separately the data measured in each setup and then merge the partial mode shapes to obtain global ones: the reference sensors are used to combine the different parts of the mode shapes, while the eigenfrequencies and damping ratios are averaged. This way of computing global mode shapes, which can be named as the multi-step approach, is outlined in Figure 4.1. However, some drawbacks can be pointed out:

- If the number of setups is large, this approach is tiresome, especially if the modes of interest are not well excited and therefore difficult to extract at all the setups [7, 25, 69].
- Since the excitation can not be controlled or measured, differences in the excitations results in the extraction of spurious or unphysical modes in some setups [7, 25, 69].
- It is not easy to pair the corresponding mode at each setup, specially when there are modes with closely spaced natural frequencies or with similar mode shapes at the reference DOFs [7, 25, 69].
A lot of effort has to be done to properly merge the different parts of the mode shapes estimated in each setup, for example in a least square sense \[6\].

On the other hand, there is an increasing interest to process all setups at the same time because the global mode shapes are obtained automatically (the one-step approach, in contrast to the multi-step one, shown in Figure 4.2): in Ref. \[50\] a frequency-domain maximum likelihood identification technique was used for the modal parameter estimation. A comparison was made between a non-parametric and a parametric approach, where the unwanted non-stationary effects are removed respectively before and after the system identification step; in Refs. \[71\], \[34\] and \[35\] the Stochastic Subspace Identification (SSI) method was adapted to handle these kind of problems. However, it is necessary to normalize the data of the different setups before to apply the algorithms because the unmeasured background excitation of each setup might be different. Finally, some of these techniques were analysed and compared in \[36\], \[65\] and \[89\].

In this Chapter we propose a new method to manage multiple setups of sensors and ambient vibrations. Some ideas about how to face multiple records were introduced in Chapter 3. Here we extend the state space model \[3.5\] to take into account changes in the position of the sensors from one record to the next. We derive the equations for the estimation of this new model by mean of the EM algorithm and apply the proposed method to interesting examples.

### 4.2 State space model for the analysis of multiple setups of sensors

In Chapter 3 we used the state space model

\[
\begin{align*}
x_t^{(r)} &= Ax_{t-1}^{(r)} + w_t^{(r)}, \quad w_t^{(r)} \sim N(0, Q^{(r)}), \\
y_t^{(r)} &= Cx_t^{(r)} + u_t^{(r)}, \quad u_t^{(r)} \sim N(0, R^{(r)}),
\end{align*}
\]

to deal with multiple records. In the case of multiple setups of sensors, we have multiple records as well, but now the position of the sensors changes from one setup to the other.
Figure 4.2: Estimating the modal parameters processing $M$ setups at the same time (the one-step approach). MPE means modal parameter estimation.

To take into account this situation, the matrix $C$ has to change with the records, that is

$$x_t^{(r)} = Ax_{t-1}^{(r)} + w_t^{(r)}, \quad w_t^{(r)} \sim N(0, Q^{(r)}), \quad (4.1a)$$
$$y_t^{(r)} = C^{(r)} x_t^{(r)} + v_t^{(r)}, \quad v_t^{(r)} \sim N(0, R^{(r)}). \quad (4.1b)$$

Since we know the position of the sensors in each setup, it is more advantageous to write this model as

$$x_t^{(r)} = Ax_{t-1}^{(r)} + w_t^{(r)}, \quad w_t^{(r)} \sim N(0, Q^{(r)}), \quad (4.2a)$$
$$y_t^{(r)} = L^{(r)} C x_t^{(r)} + v_t^{(r)}, \quad v_t^{(r)} \sim N(0, R^{(r)}), \quad (4.2b)$$

where $C \in \mathbb{R}^{n_o \times n_s}$ is a global matrix taking into account all the measured positions; $n_o$ is the number of different DOFs where the output has been measured; $n_{or}$ is the number of DOFs measured in record $r$; $L^{(r)} \in \mathbb{R}^{n_{or} \times n_o}$ is a location or selection matrix formed by ones and zeros verifying $C^{(r)} = L^{(r)} C$; so if we define a vector with all the different DOFs that have been measured in the structure, called the global list of measured DOFs or $y_t^{(G)} \in \mathbb{R}^{n_o \times 1}$, the location matrices $L^{(r)}$ need to verify $y_t^{(r)} = L^{(r)} y_t^{(G)}$. The location matrices are known for each setup $r$.

The state space model given by Eq. (4.2) is the model we are going to consider in this Chapter. The unknown parameters of the model are

$$\theta = \{A, C, Q^{(r)}, R^{(r)}, \bar{x}_0^{(r)}, P_0^{(r)}\}, \quad r = 1, 2, \ldots, M \quad (4.3)$$

where $\bar{x}_0^{(r)}$ and $P_0^{(r)}$ are the mean and variance of the initial states $x_0^{(r)}$ respectively (which are assumed to be normal distributed). The main properties of this model are summarized here:

- It provides a single value for the natural frequency, a single value for the damping ratio and a single value for the global mode shape for each mode.
- Modes with closely spaced natural frequencies and similar mode shapes are easily estimated.
- It strengthens the parameters that are common to most of the setups and removes the parameters specific to individual setups. This is specially important when the excitation is non-white in given setups or quite different among setups.
• The model allows different unmeasured excitations from setup to setup because different \( Q \) and \( R \) matrices are estimated for each record.

• The same sensor position can be measured more than one time. All the recorded information is used by the model, and, the more information you record, the better estimate of the modes is obtained.

• The model can be used with the known scheme of permanent-moving sensors, where permanent sensors do not change their position from setup to setup. But the model allows to use setups without permanent sensors. The only requirement is that each setup shares the position of some sensors with at least another setup (the overlapped sensors). The number of overlapped sensors can be different from one setup to the next.

4.3 Maximum likelihood estimation: EM algorithm

In essence, model (4.2) is also a model for multiple records. The likelihood for these cases was derived in Chapter 3

\[
\log L_{Y_N}(\theta) = l_{Y_N}(\theta) = -\frac{n_pNM}{2} \log 2\pi - 2\sum_{r=1}^{M} \sum_{t=1}^{N} \log \left| \Sigma_t^{(r)}(\theta^{(r)}) \right| - \frac{1}{2} \sum_{r=1}^{M} \sum_{t=1}^{N} \left( \epsilon_t^{(r)}(\theta^{(r)}) \right)^T \left( \Sigma_t^{(r)}(\theta^{(r)}) \right)^{-1} \epsilon_t^{(r)}(\theta^{(r)}).
\]

We describe now how to maximize Eq. (4.4) using the Expectation-Maximization algorithm. Consider we know both the observed outputs \( Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)}\} \) and the states \( X_N^{(r)} = \{x_1^{(r)}, x_2^{(r)}, \ldots, x_N^{(r)}\} \). Then, the density function for one individual record is given by

\[
f_{\theta^{(r)}}(X_N^{(r)}, Y_N^{(r)}) = f_{x_0^{(r)}, p_0^{(r)}}(x_0^{(r)}) \prod_{t=1}^{N} f_{A,Q^{(r)}}(X_t^{(r)} | X_{t-1}^{(r)}) \prod_{t=1}^{N} f_{C,R^{(r)}}(Y_t^{(r)} | X_t^{(r)}),
\]

where under Gaussian assumption

\[
f_{x_0^{(r)}, p_0^{(r)}}(x_0^{(r)}) = \frac{1}{(2\pi)^{n_s/2} |P_0^{(r)}|^{1/2}} \exp \left( -\frac{1}{2} (x_0^{(r)} - \bar{x}_0^{(r)})^T \left( P_0^{(r)} \right)^{-1} (x_0^{(r)} - \bar{x}_0^{(r)}) \right),
\]

\[
f_{A,Q^{(r)}}(X_t^{(r)} | X_{t-1}^{(r)}) = \frac{1}{(2\pi)^{n_s/2} |Q^{(r)}|^{1/2}} \exp \left( -\frac{1}{2} (x_t^{(r)} - Ax_{t-1}^{(r)})^T \left( Q^{(r)} \right)^{-1} (x_t^{(r)} - Ax_{t-1}^{(r)}) \right),
\]

\[
f_{C,R^{(r)}}(Y_t^{(r)} | X_t^{(r)}) = \frac{1}{(2\pi)^{n_r/2} |R^{(r)}|^{1/2}} \exp \left( -\frac{1}{2} (y_t^{(r)} - L^{(r)}Cx_t^{(r)})^T \left( R^{(r)} \right)^{-1} (y_t^{(r)} - L^{(r)}Cx_t^{(r)}) \right),
\]
Thus, if we consider $M$ independent setups, the joint density function $f_\theta(X_N,Y_N)$ will be the product of individual ones

$$f_\theta(X_N,Y_N) = \prod_{r=1}^M f_{\theta(r)}(X_N^{(r)},Y_N^{(r)}). \quad (4.6)$$

The complete data likelihood is defined by $L_{X_N,Y_N}(\theta) = f_\theta(X_N,Y_N)$. In practice the log-likelihood is used, so information is combined by addition and it can be written as a sum of the log-likelihood of each individual record:

$$l_{X_N,Y_N}(\theta) = \log L_{X_N,Y_N}(\theta) = \sum_{r=1}^M l_{X_N^{(r)},Y_N^{(r)}}(\theta^{(r)}). \quad (4.7)$$

Finally, the log-likelihood of record $r$ can be written as the sum of three uncoupled functions

$$l_{X_N^{(r)},Y_N^{(r)}}(\theta^{(r)}) = -\frac{1}{2} [l_1(x_0^{(r)},P_0^{(r)}) + l_2(A,Q^{(r)}) + l_3(C,R^{(r)})], \quad (4.8)$$

where, ignoring constants, are

$$l_1(x_0^{(r)},P_0^{(r)}) = \log |P_0^{(r)}| + (x_0^{(r)} - \bar{x}_0^{(r)})^T (P_0^{(r)})^{-1} (x_0^{(r)} - \bar{x}_0^{(r)}), \quad (4.9)$$

$$l_2(A,Q^{(r)}) = N \log |Q^{(r)}| + \sum_{t=1}^N (z_t^{(r)} - Ax_t^{(r)})^T (Q^{(r)})^{-1} (z_t^{(r)} - Ax_t^{(r)}), \quad (4.10)$$

$$l_3(C,R^{(r)}) = N \log |R^{(r)}| + \sum_{t=1}^N (y_t^{(r)} - L^{(r)}Cz_t^{(r)})^T (R^{(r)})^{-1} (y_t^{(r)} - L^{(r)}Cz_t^{(r)}). \quad (4.11)$$

The MLE of the parameters $\theta$ (see Eq. (4.3)) is obtained by maximizing the log-likelihood given by Eq. (4.7) by mean of the EM algorithm in an iterative two steps procedure:

- **Expectation step.** To compute the expected value of the log-likelihood given by Eq. (4.7), that is, $\mathbb{E}[l_{X_N,Y_N}(\theta)|Y_N,\theta_j]$.
- **Maximization step.** To maximize $\mathbb{E}[l_{X_N,Y_N}(\theta)|Y_N,\theta_j]$ with respect to the parameters $\theta$.

### 4.3.1 E-step: expectation step

**Property 4.1.** Given the observed vectors $Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)}\}$ and the location matrices $L^{(r)}$ for $r = 1, \ldots, M$, and a value for the parameters $\theta_j$, then

$$\mathbb{E}[l_{X_N,Y_N}(\theta)|Y_N,\theta_j] = \sum_{r=1}^M \mathbb{E}[l_{X_N^{(r)},Y_N^{(r)}}(\theta^{(r)})|Y_N^{(r)},\theta_j^{(r)}]$$

$$= \sum_{r=1}^M \left( \mathbb{E}[l_1(x_0^{(r)},P_0^{(r)})|Y_N^{(r)},\theta_j^{(r)}] + \mathbb{E}[l_2(A,Q^{(r)})|Y_N^{(r)},\theta_j^{(r)}] + \mathbb{E}[l_3(C,R^{(r)})|Y_N^{(r)},\theta_j^{(r)}] \right)$$
Proof. We do not include the proof of this Property because it is easily obtained from the proof of Prop. 3.1.

\[ \mathbb{E} \left[ l_2(A, Q^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] = \begin{cases} \log |Q^{(r)}| + \text{tr} \left( \left( Q^{(r)} \right)^{-1} \left[ S_{xx}^{(r)} - S_{xb}^{(r)} A^T - A S_{bx}^{(r)} + A S_{bb}^{(r)} A^T \right] \right), & \text{if } r = 1, \ldots, M, \\ N \log |R^{(r)}| + \text{tr} \left( R^{(r)} \right)^{-1} \left[ S_{yy}^{(r)} - S_{yx}^{(r)} C^T \left( L^{(r)} \right)^T - L^{(r)} C S_{xy}^{(r)} + L^{(r)} C S_{xx}^{(r)} C^T \left( L^{(r)} \right)^T \right], & \text{otherwise} \end{cases} \]

The definition of \( x_t^{N,(r)}, P_t^{N,(r)}, P_{t,t-1}^{N,(r)}, S_{xx}^{(r)}, S_{bb}^{(r)}, S_{xb}^{(r)}, S_{yx}^{(r)} \) and \( S_{yy}^{(r)} \) can be found at Def. 3.1 and Prop. 3.1.

4.3.2 M-step: maximization step.

**Property 4.2.** The maximum of \( \mathbb{E}[l_{X,N,Y_N}(\theta) | Y_N, \theta_j] \) (Eq. (4.12)) is attained at

\[ \bar{x}_0^{(r)} = x_0^{N,(r)}, \quad r = 1, \ldots, M, \]

\[ P_0^{(r)} = P_0^{N,(r)}, \quad r = 1, \ldots, M, \]

\[ \text{vec}(A) = \sum_{r=1}^{M} \left( S_{bb}^{(r)} \otimes \left( Q^{(r)} \right)^{-1} \right)^{-1} \sum_{r=1}^{M} \text{vec} \left( \left( Q^{(r)} \right)^{-1} S_{xb}^{(r)} \right), \]

\[ Q^{(r)} = \frac{1}{N} \left( S_{xx}^{(r)} - S_{xb}^{(r)} A^T - A S_{bx}^{(r)} + A S_{bb}^{(r)} A^T \right), \quad r = 1, \ldots, M, \]

\[ \text{vec}(C) = \sum_{r=1}^{M} S_{xx}^{(r)} \otimes \left( L^{(r)} \right)^T \left( R^{(r)} \right)^{-1} L^{(r)} \right)^{-1} \sum_{r=1}^{M} \text{vec} \left( L^{(r)} \left( R^{(r)} \right)^{-1} S_{xy}^{(r)} \right), \]

\[ R^{(r)} = \frac{1}{N} \left( S_{yy}^{(r)} - S_{yx}^{(r)} C^T \left( L^{(r)} \right)^T - L^{(r)} C S_{xy}^{(r)} + L^{(r)} C S_{xx}^{(r)} C^T \left( L^{(r)} \right)^T \right). \]

**Proof.** The proof of this Property is given in Section 4.9. 

4.3.3 Overall procedure

The two steps, expectation and maximization, have to be repeated iteratively until the likelihood is maximized. The overall method can be summarized as follows:

- Initialize the procedure \((j = 0)\) selecting starting values for the parameters \(\theta_0\) and a stop tolerance \(\delta_{adm}\).

- Repeat

  1. Perform the E-Step. Apply the Kalman filter (Properties A.4, A.5 and A.6) to each record \(r\) to obtain the expected values \(x_t^{N,(r)}, P_t^{N,(r)}\), and \(P_{t,t-1}^{N,(r)}\) with \(\theta_j\) as data. Use them to compute the matrices \(S_{xb}^{(r)}, S_{bb}^{(r)}, S_{yx}^{(r)}\) and \(S_{xx}^{(r)}\) given by (3.26)-(3.29).

  2. Perform the M-Step. Compute the updated value of the parameters \(\theta_{j+1}\) using (4.16)-(4.21).

  3. Compute the likelihood \(l_{YN}(\theta_{j+1})\) with Equation (4.4).

  4. Compute the actual tolerance

     \[
     \delta = \frac{|l_{YN}(\theta_{j+1}) - l_{YN}(\theta_j)|}{|l_{YN}(\theta_j)|} \tag{4.22}
     \]

     (a) If \(\delta > \delta_{adm}\), perform a new iteration with \(\theta_{j+1}\) as the value of the parameters.

     (b) If \(\delta \leq \delta_{adm}\), stop the iterations. The estimate is \(\theta_{j+1}\).

4.4 Starting values for the EM algorithm

The selection of the starting point for the EM algorithm is not easy for models like (4.2). On the one hand, the starting point needs to be built fast and easy; and on the other hand, it is advisable to take advantage of the information provided by the data. We think a good alternative is to use the parameters estimated with another method (for example, SSI) to build the starting point. This approach can be used here applying SSI to one setup and then using the result as starting point for all the setups. We propose to build the starting point \(\theta_0 = \left( A_0, C_0, Q_0^{(r)}, R_0^{(r)}, \bar{x}_0^{(r)}, P_0^{(r)} \right)\) as follows:

- First, to apply the SSI algorithm to one of the setups, setup \(s\), obtaining \(A_{SSI}^{(s)}, C_{SSI}^{(s)}, Q_{SSI}^{(s)}\) and \(R_{SSI}^{(s)}\).

- \(A_0 = A_{SSI}^{(s)}\). We take advantage of the information given by the eigenvalues of \(A_{SSI}^{(s)}\). These eigenvalues may be influenced by the particular position of the sensors in setup \(s\), but their information is useful to the EM algorithm to convergence to the right solution.

- \(C_0\) is constructed so that the product \(L^{(r)}C_0\) are equal to the first \(n_{ax}\) rows of \(C_{SSI}^{(s)}\). This condition give us an idea of the relation between the states given by \(A_0\) and the outputs. Again, this is only valid for setup \(s\), but it is enough for a starting point.
Analysis of multiple setups of sensors

- \( Q_0^{(r)} = Q_{SSI}^{(s)} \) for all \( r \). The result is a positive definite matrix (this is important for the EM), and it is a good approximation to the rest of the covariance matrices.

- \( R_0^{(r)} \) = the first \( n_{or} \) rows and columns of \( R_{SSI}^{(s)} \). Another positive definite matrix and adequate for the outputs.

- \( \hat{x}_{0,0}^{(r)} \) and \( P_{0,0}^{(r)} \) are taken matrices of zeros.

Note that this approach to build the initial point needs SSI to be applied to the setup with the larger number of measured points (the larger \( n_{or} \)).

4.5 Example 1: simulated data

4.5.1 Description of the data

We have used the simulated structure presented in Section 3.5. The natural frequencies, mode shapes and damping ratios corresponding to this structure were presented in Figure 3.1(b).

In this Chapter, the acceleration of the eight DOFs has been generated three times in order to simulate three different tests:

- The input in each test is Gaussian white noise with zero mean and variances 1, 4 and 9 for each one (the excitation level is different for each setup). These inputs have been applied to all the DOFs. The objective is to excite all the modes, that is why the inputs are white noise and have been applied to all DOFs.

- Sampling frequency \( f_s = 50 \) Hz. Total duration of signals, 100 seconds (5000 time steps).

- The observed value is the sum of the structure response and a sensor Gaussian noise with variance equal to the 20% of the largest acceleration response variance.

4.5.2 Proposed method with permanent-moving sensors

First we are going to apply the method proposed in this Chapter considering permanent-moving sensors, that is, all the setups have a certain number of sensors located at the same place (the permanent sensors) and the rest are moving sensors. This strategy is widely used nowadays to estimate the modes of large structures (see for example [34, 35, 36, 37, 38]).

We have simulated the following three setups of sensors:

- Setup 1: Acceleration of masses 1, 2, 3, 4, 5 from the simulated test 1.

- Setup 2: Acceleration of masses 1, 2, 6, 7 from the simulated test 2.

- Setup 3: Acceleration of masses 1, 2, 8 from the simulated test 3.

That is, we have considered masses 1 and 2 are permanent pools, and the rest are moving pools. Besides, we have considered different number of moving sensors in each setup: three, two and one respectively.
The location matrices for this case are built now. First, we define the global list of measured DOFs, \( y^{(G)}_t \), and the measured DOF at each setup:

\[
y^{(G)}_t = \begin{bmatrix} 
\ddot{q}_{1,t} \\
\ddot{q}_{2,t} \\
\ddot{q}_{3,t} \\
\ddot{q}_{4,t} \\
\ddot{q}_{5,t} \\
\ddot{q}_{6,t} \\
\ddot{q}_{7,t} \\
\ddot{q}_{8,t}
\end{bmatrix} ; \quad y^{(1)}_t = \begin{bmatrix} 
\ddot{q}_{1,t} \\
\ddot{q}_{2,t} \\
\ddot{q}_{3,t} \\
\ddot{q}_{4,t} \\
\ddot{q}_{5,t} \\
\ddot{q}_{6,t} \\
\ddot{q}_{7,t} \\
\ddot{q}_{8,t}
\end{bmatrix} ; \quad y^{(2)}_t = \begin{bmatrix} 
\ddot{q}_{1,t} \\
\ddot{q}_{2,t} \\
\ddot{q}_{6,t} \\
\ddot{q}_{7,t}
\end{bmatrix} ; \quad y^{(3)}_t = \begin{bmatrix} 
\ddot{q}_{1,t} \\
\ddot{q}_{2,t} \\
\ddot{q}_{8,t}
\end{bmatrix} .
\] (4.23)

where \( \ddot{q}_{k,t} \) is the acceleration of mass \( k \) at time instant \( t \). The following relationships must be verified

\[
y^{(1)}_t = L^{(1)} y^{(G)}_t , \quad y^{(2)}_t = L^{(2)} y^{(G)}_t , \quad y^{(3)}_t = L^{(3)} y^{(G)}_t ;
\] (4.24)

Thus

\[
L^{(1)} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix} ; \quad L^{(2)} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix} ;
\]

\[
L^{(3)} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} .
\]

The proposed method has been applied to the simulated data obtaining the results shown in Figure 4.3 (the starting point was built applying SSI to record 1). The natural frequencies and mode shapes are estimated with precision; the damping ratios have more bias, even when using simulated data. This variability of damping ratios is not a drawback of the EM algorithm, but it is present in greater or lesser extent in all the identification methods (see [82] for a comparative analysis of damping ratios estimated using different approaches).

### 4.5.3 Multi-step approach

Figure 4.3 shows that the values obtained using the proposed method with the simulated data are good. An important issue is the comparison of these results with the traditional scheme of processing each setup separately and then to merge the partial mode shapes (the multi-step approach, in contrast to the proposed method which can be considered as a one-step approach). We have estimated the modal parameters from the setups 1, 2 and 3 defined in Section 4.5.2 using the SSI algorithm under a state space order \( n_s = 16 \) (the theoretic order). Since the natural frequencies are quite separated, it is easy to identify the corresponding modes in each setup. Then, mean frequencies and mean damping ratios of each mode can be computed, and the global mode shapes can be merged. Table 4.1 shows the obtained results.

In general terms, the natural frequencies estimated by SSI in the three setups are very similar to the theoretical ones. The main deviations are observed in mode number
8, although it is a small one: theoretical = 23.12 Hz, estimated in setup 1 = 24.08 Hz (5.25 %). On the contrary, the variability in the estimated damping ratios is higher: for mode 8 (estimated worse mode), the theoretical value is 1.50, and the values estimated for setup 1, 2 and 3 are 2.62, 1.51 and 1.86 respectively (mean value equal to 2.00). Finally, MAC values between the merged modal shapes and the theoretical ones are very good for modes 1 to 5, not so good for modes 6 and 7, and bad for mode 8.

The natural frequencies and damping ratios estimated by the proposed one-step method are very similar to the mean values of the SSI. However, the MAC values corresponding to the one-step approach are very good for all the modes, even for mode 8 (MAC=0.9817).

### 4.5.4 Proposed method with overlapped sensors

We show now that the method can be used with more general setups of sensors, for example with setups without permanent sensors. As we indicated in Section 4.2, the only requirement is that each setup shares the position of some sensors with, at least, another setup (the overlapped sensors). In this example we define

- Setup 1: Acceleration of masses 1, 2, 3, 4 from the simulated test 1.
4.5 Example 1: simulated data

<table>
<thead>
<tr>
<th>Mode number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>f (Hz)</td>
<td>2.94</td>
<td>5.87</td>
<td>8.60</td>
<td>11.19</td>
<td>13.78</td>
<td>16.52</td>
<td>19.54</td>
<td>23.12</td>
</tr>
<tr>
<td></td>
<td>2.95</td>
<td>5.86</td>
<td>8.60</td>
<td>11.25</td>
<td>13.75</td>
<td>16.54</td>
<td>19.54</td>
<td>24.08</td>
</tr>
<tr>
<td></td>
<td>2.96</td>
<td>5.87</td>
<td>8.63</td>
<td>11.19</td>
<td>13.80</td>
<td>16.52</td>
<td>19.72</td>
<td>22.88</td>
</tr>
<tr>
<td></td>
<td>2.96</td>
<td>5.85</td>
<td>8.60</td>
<td>11.18</td>
<td>13.79</td>
<td>16.55</td>
<td>19.63</td>
<td>23.14</td>
</tr>
<tr>
<td></td>
<td>2.96</td>
<td>5.86</td>
<td>8.61</td>
<td>11.21</td>
<td>13.78</td>
<td>16.54</td>
<td>19.63</td>
<td>23.37</td>
</tr>
<tr>
<td></td>
<td>2.97</td>
<td>5.87</td>
<td>8.60</td>
<td>11.20</td>
<td>13.78</td>
<td>16.51</td>
<td>19.55</td>
<td>23.07</td>
</tr>
<tr>
<td>ζ (%)</td>
<td>2.00</td>
<td>1.24</td>
<td>1.10</td>
<td>1.10</td>
<td>1.15</td>
<td>1.23</td>
<td>1.35</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>3.32</td>
<td>1.10</td>
<td>0.91</td>
<td>1.30</td>
<td>1.37</td>
<td>1.08</td>
<td>1.38</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>1.52</td>
<td>1.06</td>
<td>0.97</td>
<td>1.35</td>
<td>1.11</td>
<td>1.19</td>
<td>1.81</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>1.98</td>
<td>0.91</td>
<td>0.90</td>
<td>0.82</td>
<td>0.97</td>
<td>1.44</td>
<td>1.42</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>2.27</td>
<td>1.03</td>
<td>0.92</td>
<td>1.15</td>
<td>1.15</td>
<td>1.24</td>
<td>1.54</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>2.44</td>
<td>1.04</td>
<td>0.86</td>
<td>1.05</td>
<td>1.16</td>
<td>1.22</td>
<td>1.64</td>
<td>1.70</td>
</tr>
<tr>
<td>MAC</td>
<td>1,000</td>
<td>0.998</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.761</td>
<td>0.721</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.997</td>
<td>1,000</td>
<td>0.999</td>
<td>0.998</td>
<td>0.989</td>
<td>0.982</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Natural frequencies and damping ratios estimated from the simulated data; SSI(k) stands for the values estimated applying the SSI algorithm to setup k (\(n_s = 16\)); “SSI mean” is the mean of the values estimated with SSI; EM stands for the values obtained using the proposed method (permanent-moving sensors); MAC values are computed between estimated mode shapes and the theoretical ones.

- Setup 2: Acceleration of masses 3, 4, 5, 6 from the simulated test 2.
- Setup 3: Acceleration of masses 5, 6, 7, 8 from the simulated test 3.

That is, setup 1 and setup 2 have the overlapped sensors at masses 3 and 4; setup 2 and setup 3 share sensors at masses 5 and 6; finally, setups 1 and 3 do not have overlapped sensors. Therefore, DOFs 1, 2, 7 and 8 are measured one time, and DOFs 3, 4, 5, and 6 are measured two times. With the proposed model, all the DOFs can be measured in more than one setup, and all the measurements are used to estimate the mode shapes. This means that we could add extra tests so each DOF is measured the times we want. The proposed method would use all the information to estimate the modes, and in theory, the more data is analysed, the better estimate is obtained.

The location matrices that correspond to the three setups given above are defined by

\[
y_t^{(1)} = L^{(1)} y_t^{(G)}, \quad y_t^{(2)} = L^{(2)} y_t^{(G)}, \quad y_t^{(3)} = L^{(3)} y_t^{(G)},
\]

where

\[
L^{(1)} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\end{bmatrix}, \quad
L^{(2)} = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\end{bmatrix};
\]

\[
L^{(3)} = \begin{bmatrix}
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]
### Table 4.2: Natural frequencies and damping ratios of the modes estimated by the proposed method with permanent sensors and with overlapped sensors. MAC values are computed between estimated values and theoretical values.

<table>
<thead>
<tr>
<th>Frequencies (Hz)</th>
<th>Damping ratios (%)</th>
<th>MAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.94</td>
<td>2.97</td>
<td>2.96</td>
</tr>
<tr>
<td>5.87</td>
<td>5.87</td>
<td>5.86</td>
</tr>
<tr>
<td>8.60</td>
<td>8.60</td>
<td>8.61</td>
</tr>
<tr>
<td>11.19</td>
<td>11.20</td>
<td>11.19</td>
</tr>
<tr>
<td>13.78</td>
<td>13.78</td>
<td>13.78</td>
</tr>
<tr>
<td>16.52</td>
<td>16.51</td>
<td>16.52</td>
</tr>
<tr>
<td>19.54</td>
<td>19.55</td>
<td>19.57</td>
</tr>
<tr>
<td>23.12</td>
<td>23.07</td>
<td>23.07</td>
</tr>
</tbody>
</table>

\[ y^{(G)}_t = \begin{bmatrix} \ddot{q}_{1,t} \\ \ddot{q}_{2,t} \\ \ddot{q}_{3,t} \\ \ddot{q}_{4,t} \\ \ddot{q}_{5,t} \\ \ddot{q}_{6,t} \\ \ddot{q}_{7,t} \\ \ddot{q}_{8,t} \end{bmatrix} ; \quad y^{(1)}_t = \begin{bmatrix} \ddot{q}_{1,t} \\ \ddot{q}_{2,t} \\ \ddot{q}_{3,t} \\ \ddot{q}_{4,t} \\ \ddot{q}_{5,t} \\ \ddot{q}_{6,t} \\ \ddot{q}_{7,t} \end{bmatrix} ; \quad y^{(2)}_t = \begin{bmatrix} \ddot{q}_{3,t} \\ \ddot{q}_{4,t} \\ \ddot{q}_{5,t} \\ \ddot{q}_{6,t} \\ \ddot{q}_{7,t} \end{bmatrix} ; \quad y^{(3)}_t = \begin{bmatrix} \ddot{q}_{5,t} \\ \ddot{q}_{6,t} \\ \ddot{q}_{7,t} \end{bmatrix}. \]

and \( \ddot{q}_{k,t} \) is the acceleration of mass \( k \) at time instant \( t \).

The results are given in Table 4.2. This table also includes the results computed in Section 4.5.2, that is, using permanent-moving sensors. We can check that the results obtained with both approaches are very similar to the theoretical ones.

### 4.6 Example 2: Ninove footbridge

#### 4.6.1 Bridge description and ambient vibrations tests

The objective of this section is to apply the proposed method to a real structure using the scheme of permanent-moving sensors. The chosen structure is a footbridge over the Dender river, in Ninove (Belgium). The Ninove footbridge is a cable-stable bridge with two pylons and six pairs of cables. The pylons split the footbridge in two spans of 36.00 m and 22.50 m respectively. The bridge deck is a steel tubular truss. A picture of the footbridge has been included in Figure 4.4(a).

A series of vibration tests were conducted on April 17 and 18, 2011, by researchers from K.U.Leuven. The footbridge’s dynamic response was measured with seven stations equipped with triaxial accelerometers synchronized by GPS. The measurements were carried out in a total of 13 different setups that are shown in Figure 4.4(b): three stations served as references (permanently located at pools 7, 20 and 39), and the other four were used as mobile. For example, according to the figure, the first setup was composed by the pools 7, 20, 39, 1, 2, 3, 4; the second one was 7, 20, 39, 5, 6, 8, 9, and so on.

For each setup, time series of 10 minutes were collected with a sampling frequency of...
4.6 Example 2: Ninove footbridge

(a)

(b)

Figure 4.4: (a) Ninove footbridge. (b) Footbridge geometry and setups layout.

200 Hz. Since the modes of interest are in the range 0-15 Hz, the data were re-sampled so that the new sampling frequency was 40 Hz (we used the Matlab function decimate).

All the measured channels are taken into account in the analysis, that is, vertical, longitudinal and transversal accelerations. This is important because in structures like this one, the mode shapes involves components in various directions.

4.6.2 Proposed method

All parametric system identification methods based on the state space model require to know the model order \( n_s \), which in theory, is twice the number of identified modes. In modal testing applications it is customary to estimate the state space model for a wide range of orders and to build a stabilization diagram \([51]\). The proposed method has been used to compute the stabilization diagram from the multiple setups, and it has been plotted in Figure 4.5. The stabilization diagram can be built in the usual way because the proposed method estimate a single A and C matrices. The global mode shapes are now used to decide if the mode is stable or not.

Taking into account the stabilization diagram, we have chosen the modes estimated with a state space order equal to \( n_s = 58 \). This means that 29 modes should be estimated, that is, the A matrix should have 29 pairs of complex conjugate eigenvalues.
The results are included in Table 4.3. Not all the modes estimated correspond to physical modes: for example, modes 1 and 30 come from real eigenvalues; modes 7, 14 and 20 have a damping ratio too high to be a structural mode. In fact, these modes are non-stable in the stabilization diagram. Other modes have low damping ratio but are non-stable:

- Modes 27, 28, 29 may be affected by the filter applied in the process to reduce the sampling frequency from 200 Hz to 40 Hz. According to the Matlab help (decimate function), the filter applied is a Chebyshev Type I filter with a cut-off frequency of 0.8F, where F is the new Nyquist frequency (F=20 Hz). This means that modes with frequency higher than 16 Hz might be affected by the filter.

- Mode 21 are estimated at several orders but in a non-stable way.

- Modes 2, 4, 15, 24 and 25 are also non-stable.

The rest of the estimated modes (i.e., modes 3, 5, 6, 8, 9, 10, 11, 12, 13, 16, 17, 18, 19, 22, 23 and 26) seem to be physical modes: they appear as stable in the stabilization diagram, the damping ratios are about 1 - 3 %, and the mode shapes are feasible. These modes include vertical bending modes, lateral bending modes, longitudinal modes and also modes with coupled torsion. Figure 4.6 shows these identified modes. We have also drawn the position of the pylons and the cable anchorages because the modes are influenced by these elements.
4.6 Example 2: Ninove footbridge

Figure 4.6: Identified modes at Ninove footbridge using the proposed method with $n_s = 58$. ■ stands for footbridge pilons, and ● stands for cables anchorages.
### Table 4.3: Natural frequencies and damping ratios of the modes estimated by the proposed method at $n_s = 58$.

<table>
<thead>
<tr>
<th>No</th>
<th>$f$ (Hz)</th>
<th>$\zeta$ (%)</th>
<th>Stab. diag.</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.79</td>
<td>–</td>
<td>+</td>
<td>Real eigenvalue.</td>
</tr>
<tr>
<td>2</td>
<td>2.96</td>
<td>1.82</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.97</td>
<td>0.37</td>
<td>⊕</td>
<td>Vertical bending</td>
</tr>
<tr>
<td>4</td>
<td>3.02</td>
<td>5.81</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.07</td>
<td>1.02</td>
<td>⊕</td>
<td>Lateral bending long span coupled with torsion.</td>
</tr>
<tr>
<td>6</td>
<td>3.80</td>
<td>1.51</td>
<td>⊕</td>
<td>Lateral bending coupled with torsion.</td>
</tr>
<tr>
<td>7</td>
<td>4.15</td>
<td>33.72</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5.78</td>
<td>2.52</td>
<td>⊕</td>
<td>Vertical bending short span.</td>
</tr>
<tr>
<td>9</td>
<td>6.01</td>
<td>0.53</td>
<td>⊕</td>
<td>Vertical bending.</td>
</tr>
<tr>
<td>10</td>
<td>6.95</td>
<td>0.50</td>
<td>⊕</td>
<td>Torsion long span.</td>
</tr>
<tr>
<td>11</td>
<td>8.00</td>
<td>1.23</td>
<td>⊕</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.11</td>
<td>5.01</td>
<td>⊕</td>
<td>Longitudinal mode.</td>
</tr>
<tr>
<td>13</td>
<td>8.28</td>
<td>1.08</td>
<td>⊕</td>
<td>Torsion long span, cables zone</td>
</tr>
<tr>
<td>14</td>
<td>8.84</td>
<td>31.48</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>9.64</td>
<td>12.75</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>9.97</td>
<td>1.98</td>
<td>⊕</td>
<td>Torsion both spans.</td>
</tr>
<tr>
<td>17</td>
<td>10.85</td>
<td>1.39</td>
<td>⊕</td>
<td>Torsion short span.</td>
</tr>
<tr>
<td>18</td>
<td>11.86</td>
<td>2.65</td>
<td>⊕</td>
<td>Longitudinal mode.</td>
</tr>
<tr>
<td>19</td>
<td>12.78</td>
<td>2.78</td>
<td>⊕</td>
<td>Torsion.</td>
</tr>
<tr>
<td>20</td>
<td>12.93</td>
<td>47.66</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>13.39</td>
<td>2.19</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>13.65</td>
<td>0.84</td>
<td>⊕</td>
<td>Vertical bending long span.</td>
</tr>
<tr>
<td>23</td>
<td>14.09</td>
<td>1.04</td>
<td>⊕</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>14.09</td>
<td>6.60</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>14.18</td>
<td>6.60</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>14.71</td>
<td>0.40</td>
<td>⊕</td>
<td>Vertical bending short span.</td>
</tr>
<tr>
<td>27</td>
<td>16.47</td>
<td>4.23</td>
<td>+</td>
<td>Filter.</td>
</tr>
<tr>
<td>28</td>
<td>17.05</td>
<td>1.26</td>
<td>+</td>
<td>Filter.</td>
</tr>
<tr>
<td>29</td>
<td>17.29</td>
<td>0.13</td>
<td>+</td>
<td>Filter.</td>
</tr>
<tr>
<td>30</td>
<td>19.75</td>
<td>–</td>
<td>+</td>
<td>Real eigenvalue.</td>
</tr>
</tbody>
</table>

### 4.6.3 Multi-step approach using real measured data

Similar to the simulated case, we want to compare the results of the proposed method with the results obtained in a multi-step approach but using real measured data. We want to remark that the objective of this section is not to propose a procedure for the multi-step approach. On the contrary, we use the simplest and naive procedure (improved methods can be found in [6], [65], and [89]). The purposes of this section are:

- To show that the process of gluing modes from individual estimates is difficult and time consuming, specially with real measured data. The procedure also requires user interaction and user expertise to merge the correct modes.
- To show the advantages of using a one-step approach, like the proposed method.
• Finally, to verify that the modes obtained with the proposed method are also obtained using the multi-step approach, what gives us confidence on the method applied to experimental data.

The SSI algorithm have been applied to the thirteen setups. We have used $n_s = 58$, the order chosen in Table 4.3 and Figure 4.6. The estimated eigen-frequencies are given in Table 4.4. In the case of simulated data, each mode was clearly recognizable in each setup because the natural frequencies were chosen sufficiently spaced. In contrast, the natural frequencies obtained in real structures are difficult to organize: for example, in the range $(3 \text{ Hz} - 4 \text{ Hz})$ there are five values in setup 1, two values in setup 2, six in setup 3, four in setup 4, and so on; and there are very close frequencies: the mode with $3.03 \text{ Hz}$ in setup 1 probably corresponds to the mode with $3.04 \text{ Hz}$ in setup 2, but in setup 3 we find two modes with $3.03 \text{ Hz}$ and $3.05 \text{ Hz}$. In conclusion, the problem of matching modes among setups to obtain global mode shapes is difficult, specially in structures with a high number of setups.

We have obtained some merged modes from Table 4.4 comparing natural frequencies, damping ratios and mode shapes among setups. We assumed that mode $i$ identified in setup $p$ and mode $j$ from setup $q$ correspond to the same physical mode if they verified

$$\left|\frac{\omega_{pi} - \omega_{qj}}{\omega_{pi}}\right| \leq 0.10, \quad |\zeta_{pi} - \zeta_{qj}| \leq 0.10, \quad 1 - MAC(\tilde{\psi}_{pi}, \tilde{\psi}_{qj}) \leq 0.80. \quad (4.25)$$

where $\omega_j$, $\zeta_j$ and $\psi_j$ are the estimated natural frequencies, damping ratios and modal vectors (see by Eqs. (B.80), (B.81) and (B.82)). Note the MAC has to be computed using the components of the modal vector $\psi_j$ corresponding to the reference sensors, called here $\tilde{\psi}_j$.

These limit values are not too demanding, and of course other values can be tried. However, we only want to show some results, and to analyse the better limit values or the better procedure for the multi-step approach is out of the scope of this work.

We have found six modes that verified the above conditions in the thirteen setups. The eigenfrequencies of these modes are highlighted in Table 4.4 with the subscripts $(1), (2), (3), (4), (5)$ and $(6)$. We have computed the mean eigenfrequencies and the mean damping ratios for these modes, obtaining for the eigenfrequencies: $(1): 2.97 \text{ Hz}, (2): 3.08 \text{ Hz}, (3): 3.82 \text{ Hz}, (4): 6.02 \text{ Hz}, (5): 6.94 \text{ Hz}, (6): 7.99 \text{ Hz}$. Checking the modes obtained with the proposed method, Table 4.3, we find values very close to them: mode $3: 2.97 \text{ Hz}$, mode $5: 3.07 \text{ Hz}$, mode $6: 3.80 \text{ Hz}$, mode $9: 6.01 \text{ Hz}$, mode $10: 6.95 \text{ Hz}$, mode $11: 8.00 \text{ Hz}$. If we compute the MAC values using the mode shapes obtained with both approaches (the global mode shapes in this case), the results are very good, with the exception of mode around $3.08 \text{ Hz}$ ($\text{MAC}=0.66$). All these results are summarized in Table 4.5.

Apart from these modes, present in all the setups, we have indicated in Table 4.4 other modes that appear in almost all the setups: mode $(10)$ which come up in twelve setups, mode $(9)$ in eleven setups, and modes $(8)$ and $(11)$ in ten setups. They correspond for sure to vibrational modes, but the values estimated by SSI do not allow to obtain the global mode shapes. For example, let us analyse mode $(10)$, with frequency around $14.70 \text{ Hz}$. It is present in all the setups except in setup 2. In this setup, there are four possibles candidates: the modes with $14.12 \text{ Hz}$, $14.54 \text{ Hz}$, $14.77 \text{ Hz}$ and $15.22 \text{ Hz}$. The MAC obtained between these candidates and mode $(10)$ estimated in other setups are given in Table 4.6. We see that mode with $14.54 \text{ Hz}$ is the best candidate for mode $(10)$ in
setup 2, but the MAC with other setups is around 0.45. The conclusion is that this mode has not been properly estimated in setup 2. Maybe this mode can be improved choosing a higher order for the state space model, but this implies more modes to analyse.

On the contrary, the proposed method obtains directly the global mode shapes, their natural frequencies and their damping ratios. If a mode is worse estimated in certain setups, the method uses the information of the remaining setups to obtain the global mode shape.

Table 4.5 shows that the multi-step mode shapes have a high MAC with the obtained with proposed one-step method. Besides, the mean natural frequencies computed from the multi-step are very similar to the one-step natural frequencies. And, in lesser extent, this also happens with the damping ratios. Other modes where the multi-step approach fails, like mode (10), are estimated by the proposed method (mode 26, 14.71 Hz, 0.40%). Finally, in addition to the six complete modes and the four incomplete modes selected from the Table 4.4, the one-step method proposes some other with nice modal parameters (see Figure 4.6).

4.7 Example 3: gym building

In this section we have chosen measured data from a building placed at the Universidad Politécnica Madrid. Some pictures are shown in Figure 4.7(a)). The building geometry is outlined in Figure 4.7(b)). The ground floor is devoted to cars parking and it has no walls. In the first floor we can distinguish two parts: the first one contains some offices and the showers and it has two stories; the second one is the gym room with 12 meters wide, 22 meters long (29 meters if we take into account the angled end) and 6 meters high. With respect to the structure, the beams and columns are made of steel, and the floors and walls are made of concrete.

On September 29, 2010, the building was subjected to dynamic tests to investigate the modal parameters. The tests were conducted by researchers from the Universidad Politécnica Madrid. In total, five different setups were recorded, all vertical accelerations, with a duration of 300 seconds each one. Apart from the accelerometers, a 186 N shaker acting in the vertical direction was used to increase the excitation level of the structure using a chirp signal as input. The shaker position changed from one setup to the next to properly excite the structure. The accelerometers, as well as the shaker, were placed in the gym-room part.

On April 20, 2012, new setups were planned in order to measure additional points (also in the gym room). The shaker was used again to increase the excitation level. Nine setups were recorded, eight of which had the sensors at the same place while the shaker was changing position. The proposed model is able to manage these situations, and all the measurements are considered to estimate the parameters.

Table 4.7 summarizes the setups, the points measured at each setup and the shaker position. It is important to take into account the following remarks:

- The objective of this example is to show how the proposed model can be used to estimate the global mode shapes of a structure without using permanent sensors, that is, sensors common to all the setups. The example also shows that a sensor position can be measured more than one time, and all the measurements are used to compute the estimate.
Table 4.4: Natural frequencies estimated at \( n = 4.7 \) Example 3: gym building

<table>
<thead>
<tr>
<th>Setup number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.05</td>
<td>1.13</td>
<td>1.00</td>
<td>1.07</td>
<td>1.29</td>
<td>0.63</td>
<td>1.04</td>
<td>1.02</td>
<td>1.03</td>
<td>1.02</td>
<td>1.02</td>
<td>1.09</td>
<td>1.01</td>
</tr>
<tr>
<td>(1)</td>
<td>2.98</td>
<td>2.90</td>
<td>2.98</td>
<td>2.75</td>
<td>2.74</td>
<td>2.91</td>
<td>2.89</td>
<td>2.97</td>
<td>2.97</td>
<td>2.95</td>
<td>2.98</td>
<td>2.96</td>
<td>2.97</td>
</tr>
<tr>
<td>(2)</td>
<td>3.09</td>
<td>3.04</td>
<td>3.05</td>
<td>2.93</td>
<td>2.98</td>
<td>2.98</td>
<td>3.06</td>
<td>3.50</td>
<td>3.08</td>
<td>3.10</td>
<td>3.74</td>
<td>3.42</td>
<td>3.13</td>
</tr>
<tr>
<td>(3)</td>
<td>3.38</td>
<td>3.70</td>
<td>3.12</td>
<td>3.00</td>
<td>3.06</td>
<td>3.03</td>
<td>3.35</td>
<td>3.85</td>
<td>3.10</td>
<td>3.84</td>
<td>3.84</td>
<td>3.82</td>
<td>3.95</td>
</tr>
<tr>
<td>(4)</td>
<td>3.58</td>
<td>4.04</td>
<td>3.27</td>
<td>3.07</td>
<td>3.42</td>
<td>3.06</td>
<td>3.35</td>
<td>3.93</td>
<td>3.16</td>
<td>5.76</td>
<td>5.85</td>
<td>5.77</td>
<td>5.25</td>
</tr>
<tr>
<td>(5)</td>
<td>3.76</td>
<td>5.99</td>
<td>3.55</td>
<td>3.55</td>
<td>3.80</td>
<td>3.29</td>
<td>3.83</td>
<td>4.90</td>
<td>3.83</td>
<td>6.04</td>
<td>5.95</td>
<td>6.02</td>
<td>6.01</td>
</tr>
<tr>
<td>(6)</td>
<td>4.16</td>
<td>6.08</td>
<td>3.77</td>
<td>3.77</td>
<td>4.25</td>
<td>3.82</td>
<td>4.05</td>
<td>5.31</td>
<td>5.95</td>
<td>6.12</td>
<td>6.00</td>
<td>6.97</td>
<td>6.90</td>
</tr>
<tr>
<td>(7)</td>
<td>5.96</td>
<td>6.95</td>
<td>4.23</td>
<td>4.18</td>
<td>5.97</td>
<td>4.53</td>
<td>4.25</td>
<td>5.69</td>
<td>6.00</td>
<td>6.63</td>
<td>6.95</td>
<td>7.92</td>
<td>6.91</td>
</tr>
<tr>
<td>(8)</td>
<td>6.01</td>
<td>8.04</td>
<td>5.59</td>
<td>6.02</td>
<td>6.03</td>
<td>6.02</td>
<td>4.34</td>
<td>5.86</td>
<td>6.92</td>
<td>6.93</td>
<td>7.81</td>
<td>8.11</td>
<td>7.27</td>
</tr>
<tr>
<td>(9)</td>
<td>6.95</td>
<td>8.51</td>
<td>5.73</td>
<td>6.89</td>
<td>6.94</td>
<td>6.93</td>
<td>6.01</td>
<td>6.02</td>
<td>7.31</td>
<td>7.85</td>
<td>7.97</td>
<td>8.35</td>
<td>7.97</td>
</tr>
<tr>
<td>(10)</td>
<td>7.48</td>
<td>9.80</td>
<td>6.01</td>
<td>7.01</td>
<td>7.47</td>
<td>7.31</td>
<td>6.26</td>
<td>6.95</td>
<td>7.98</td>
<td>7.92</td>
<td>8.35</td>
<td>8.36</td>
<td>8.34</td>
</tr>
<tr>
<td>(11)</td>
<td>7.71</td>
<td>10.23</td>
<td>6.79</td>
<td>7.13</td>
<td>7.63</td>
<td>7.39</td>
<td>6.91</td>
<td>7.96</td>
<td>8.08</td>
<td>7.96</td>
<td>8.45</td>
<td>8.90</td>
<td>8.42</td>
</tr>
<tr>
<td>(12)</td>
<td>8.01</td>
<td>13.05</td>
<td>6.93</td>
<td>7.82</td>
<td>7.68</td>
<td>8.00</td>
<td>8.20</td>
<td>8.34</td>
<td>8.39</td>
<td>8.85</td>
<td>10.10</td>
<td>8.90</td>
<td>8.90</td>
</tr>
<tr>
<td>(15)</td>
<td>11.66</td>
<td>14.12</td>
<td>8.02</td>
<td>8.29</td>
<td>8.79</td>
<td>8.75</td>
<td>8.63</td>
<td>9.89</td>
<td>10.62</td>
<td>10.12</td>
<td>10.40</td>
<td>10.60</td>
<td>10.03</td>
</tr>
<tr>
<td>(19)</td>
<td>14.48</td>
<td>15.77</td>
<td>9.98</td>
<td>11.29</td>
<td>11.90</td>
<td>11.70</td>
<td>12.46</td>
<td>12.08</td>
<td>12.28</td>
<td>11.90</td>
<td>12.80</td>
<td>13.29</td>
<td>12.73</td>
</tr>
<tr>
<td>(27)</td>
<td>17.42</td>
<td>18.18</td>
<td>17.28</td>
<td>16.75</td>
<td>16.24</td>
<td>16.71</td>
<td>17.08</td>
<td>15.63</td>
<td>16.96</td>
<td>16.94</td>
<td>17.27</td>
<td>16.97</td>
<td>17.11</td>
</tr>
<tr>
<td>(28)</td>
<td>17.79</td>
<td>19.70</td>
<td>17.15</td>
<td>16.88</td>
<td>17.06</td>
<td>16.74</td>
<td>17.04</td>
<td>17.06</td>
<td>21.22</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(29)</td>
<td>–</td>
<td>–</td>
<td>31.24</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Note: Taking into account Eq. (4.25). \( k = 1, \ldots , 10 \) stands for glued or merged mode number. \( n = 58 \) by the SSI (real eigenvalues have not been removed). \( E_q (k) \) in Eq. (4.25) is the additional effect of the 4.7th mode due to the finite range (4.6th mode).
Analysis of multiple setups of sensors

<table>
<thead>
<tr>
<th>SSI</th>
<th>proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>$f$ (Hz)</td>
</tr>
<tr>
<td>(1)</td>
<td>2.97</td>
</tr>
<tr>
<td>(2)</td>
<td>3.08</td>
</tr>
<tr>
<td>(3)</td>
<td>3.82</td>
</tr>
<tr>
<td>(4)</td>
<td>6.02</td>
</tr>
<tr>
<td>(5)</td>
<td>6.94</td>
</tr>
<tr>
<td>(6)</td>
<td>7.99</td>
</tr>
</tbody>
</table>

Table 4.5: Mean natural frequencies and mean damping ratios of the merged modes estimated at $n_s = 58$ using SSI (Ninove footbridge). The corresponding modes estimated using the proposed method are also included. The last column shows the MAC values between them (MAC is computed taking into account the global modes).

<table>
<thead>
<tr>
<th>Setup 2</th>
<th>Setup 1</th>
<th>Setup 3</th>
<th>Setup 4</th>
<th>Setup 5</th>
<th>Setup 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.12 Hz</td>
<td>0.0107</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0008</td>
<td>0.0007</td>
</tr>
<tr>
<td>14.54 Hz</td>
<td>0.4566</td>
<td>0.4602</td>
<td>0.4450</td>
<td>0.4726</td>
<td>0.4461</td>
</tr>
<tr>
<td>14.77 Hz</td>
<td>0.0032</td>
<td>0.0008</td>
<td>0.0024</td>
<td>0.0003</td>
<td>0.0030</td>
</tr>
<tr>
<td>15.22 Hz</td>
<td>0.0278</td>
<td>0.0030</td>
<td>0.0028</td>
<td>0.0026</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

Table 4.6: MAC values computed between some modes estimated at Setup 2 and mode (10) estimated at setup 1, 3, 4, 5 and 6 (see Table 4.4).

- The setups showed here are not probably the optimum: some points are measured very few and others are measured in all the setups; or even other points should be measured in order to obtain more complete mode shapes. The optimal sensor configurations for a given structure is out of the scope of this work.
- The input signals of the shaker are available, but we did not use them in our analysis. We considered the shaker input as part of the ambient excitation. We think that the results obtained under this hypothesis are valid for the purposes of this work. Nevertheless, the proposed model can be extended to take into account input signals as well.
- Mixing data measured in 2010 with data measured in 2012 has to be done with care: changes in the structure, in the environment or in the climate conditions can affect to the modal parameters.

The proposed method was applied to the data and the resulting stabilization diagram is shown in Figure 4.8. We have chosen the modes estimated at state space order equal to 36. At this order, the stable modes are: 4, 6, 7, 8, 11, 14, and 18. These modes are stable too at most orders used in the stabilization diagram, so they seem to be physical ones.

The identified mode shapes are included in Figure 4.9. Modes 4 and 6 on the one hand, and modes 7 and 8 on the other hand have very similar mode shapes at the gym room. Sensors in the offices or in horizontal directions are needed to be able to differentiate
these mode shapes.

Because we used only vertical accelerometers and above all, the main excitation source was the shaker (what was vertical too), the identified modes are vertical bending ones: modes 4-6 represent the first vertical bending of the gym room; modes 7-8 the second vertical bending; mode 11 the third vertical bending; mode 14 the fourth vertical bending and finally, mode 18 the fifth vertical bending.

![Gym building pictures](image1)

![Gym building geometry and sensor location](image2)

Figure 4.7: (a) Gym building pictures; (b) Gym building geometry and sensor location (●); the dashed lines stand for the beams.

### 4.8 Conclusions

Estimating the modal parameters of a structure from ambient vibration measurements has become a standard methodology. In the time domain, the procedure involves the estimation of the well known state space model, for which many algorithms have been proposed. However, model and algorithms were developed to process one setup of sensors. When there are several setups, as in the case of estimating the mode shapes by parts, the existing methods have been adapted to try to solve the problem.

In this Chapter we have proposed a state space model that can be used for multiple setups of sensors, and we have also presented how this model can be estimated using the Expectation Maximization algorithm. This model has the same A and C matrices for all the setups, what results in a single value for the natural frequency, damping ratio and global mode shape for each mode.
Table 4.7: Points measured in each configuration of the gym building. • means a sensor and ‘s’ means the shaker.

The A and C matrices have an important property too: they have been estimated taking into account the information given for all the setups, or in other words, they contain the information common to all the setups. This is specially important when the excitation is non-white in given setups or quite different among setups, because the estimation process keeps the common parameters and isolate the parameters specific to one setup. On the other hand, the common parameters are better estimated, because all the information is used.

The stabilization diagram is made as in the case of single setups, but now the MAC values are computed using the global mode shapes. The stabilization diagram is useful to distinguish between structural modes and spurious modes.

So the problem of multiple setups of sensors can be addressed in the same way as with the well known single setup: the corresponding state space model, the algorithm to estimate the model, and the computation of the modal parameters from the model.

The proposed method has additional properties:

- The method does not need permanent sensors, but overlapped sensors can be used in place. In other words, the sensors shared by setup 1 and setup 2 can be different to the sensors shared by setup 1 and setup 3, and different to the sensors shared by setup 2 and setup 3, and so on. Besides, the number of overlapped can be different from one setup to the next.

- The method uses all the available data to estimate the modal parameters. Therefore, if one DOF has been measured in two different setups, the data from the
two setups will be used to estimate the mode shape at this DOF. This means the estimate of the mode shapes will be better at the DOFs measured in more setups.

These properties might be used to plan more efficient tests: using a lower number of setups, measuring some places more than one time in different setups to improve the estimates... These features are important in practice, and we expect them to be confirmed in further research.

4.9 Appendix

Proof of Property 4.2

The proof of Eqs. (4.16) - (4.19) are given in Prop. 3.2. For Eqs. (4.20) and (4.21):

- $\frac{\partial}{\partial C} \mathbb{E}[l_{XN,YN}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial C} \sum_{r=1}^{M} \mathbb{E}[l_3(C, R)|Y_N, \theta_j] = 0$;

$$
\mathbb{E} \left[ l_3(C, R(r))|Y_N^{(r)}, \theta_j^{(r)} \right] = N \log |R^{(r)}| +
+ \text{tr} \left( (R^{(r)})^{-1} \left[ S_{yy}^{(r)} - S_{g2}^{(r)} C^T \left( L^{(r)} \right)^T - L^{(r)} C S_{x2}^{(r)} C^T \left( L^{(r)} \right)^T \right] \right);
$$
Figure 4.9: Identified modes at the gym building using $n_s = 36$. 
\[ \frac{\partial}{\partial C} \sum_{r=1}^{M} N \log |R^{(r)}| = 0; \]
\[ \frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} S_{yx}^{(r)} \right) = 0; \]
\[ \frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( - (L^{(r)})^{-1} S_{yx}^{(r)} C^T \left( L^{(r)} \right)^T \right) = - \sum_{r=1}^{M} \left( L^{(r)} \right)^T (R^{(r)})^{-1} S_{yx}^{(r)}; \]
\[ \frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( - (L^{(r)})^{-1} L^{(r)} C S_{yx}^{(r)} \right) = - \sum_{r=1}^{M} \left( L^{(r)} \right)^T (R^{(r)})^{-1} \left( S_{yx}^{(r)} \right)^T; \]
\[ \frac{\partial}{\partial C} \sum_{r=1}^{M} \text{tr} \left( (R^{(r)})^{-1} L^{(r)} C S_{yx}^{(r)} \right) = \sum_{r=1}^{M} \left( L^{(r)} \right)^T (R^{(r)})^{-1} \left( R^{(r)} \right)^{-1} \left( L^{(r)} \right)^T C S_{yx}^{(r)}; \]

Equating to zero
\[ \sum_{r=1}^{M} \left( L^{(r)} \right)^T (R^{(r)})^{-1} \left( R^{(r)} \right)^{-1} \left( R^{(r)} \right)^{-1} \left( L^{(r)} \right)^T C S_{yx}^{(r)} = \sum_{r=1}^{M} \left( L^{(r)} \right)^T \left( R^{(r)} \right)^{-1} S_{yx}^{(r)}. \]

Taking the vec(□) operator
\[ \sum_{r=1}^{M} S_{yx}^{(r)} \otimes \left( \left( L^{(r)} \right)^T \left( R^{(r)} \right)^{-1} \left( L^{(r)} \right) \right) \text{vec}(C) = \sum_{r=1}^{M} \text{vec} \left( L^{(r)} \left( R^{(r)} \right)^{-1} S_{yx}^{(r)} \right); \]
\[ \text{vec}(C) = \left[ \sum_{r=1}^{M} S_{yx}^{(r)} \otimes \left( \left( L^{(r)} \right)^T \left( R^{(r)} \right)^{-1} \left( L^{(r)} \right) \right) \right]^{-1} \sum_{r=1}^{M} \text{vec} \left( L^{(r)} \left( R^{(r)} \right)^{-1} S_{yx}^{(r)} \right). \]

- \[ \frac{\partial}{\partial R^{(r)}} \text{E}[l_{X,y} Y_{N, \theta_j}|Y_{N}, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} \text{E}[l_3(C, R^{(r)})|Y_{N}^{(r)}, \theta_j^{(r)}] = 0, \]
\[ \text{E} \left[ l_3(C, R^{(r)})|Y_{N}^{(r)}, \theta_j^{(r)} \right] = N \log |R^{(r)}| + \text{tr} \left( (R^{(r)})^{-1} M_1^{(r)} \right), \]
where
\[ M_1^{(r)} = S_{yx}^{(r)} - S_{yx}^{(r)} C^T \left( L^{(r)} \right)^T - L^{(r)} C S_{yx}^{(r)} + L^{(r)} C S_{yx}^{(r)} C^T \left( L^{(r)} \right)^T. \]

Thus
\[ \frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} N \log |R^{(r)}| = N \left( R^{(r)} \right)^{-T}; \]
\[ \frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} \text{tr} \left( \left( R^{(r)} \right)^{-1} M_1^{(r)} \right) = - \left( \left( R^{(r)} \right)^{-1} M_1^{(r)} \left( R^{(r)} \right)^{-1} \right)^T. \]

Equating to zero
\[ N \left( R^{(r)} \right)^{-T} - \left( \left( R^{(r)} \right)^{-1} M_1^{(r)} \left( R^{(r)} \right)^{-1} \right)^T = 0 \Rightarrow R^{(r)} = \frac{1}{N} M_1^{(r)}. \]
Chapter 5

Operational modal analysis with unobserved autocorrelated inputs

5.1 Introduction and problem description

When input and output measurements are available in a structural system, the state space model we can use to estimate the modes of vibration is given by

\[ x_t = Ax_{t-1} + Bu_{t-1} + w_t \]  \hspace{1cm} (5.1a)
\[ y_t = Cx_t + Du_t + v_t, \]  \hspace{1cm} (5.1b)

where \( w_t \in \mathbb{R}^{n_s \times 1} \) and \( v_t \in \mathbb{R}^{n_i \times 1} \) are white noise processes (see Appendix A). In Operational Modal Analysis the system inputs \( u_t \in \mathbb{R}^{n_i \times 1} \) are not known. The option is to use the model

\[ x_t = Ax_{t-1} + w_t \]  \hspace{1cm} (5.2a)
\[ y_t = Cx_t + v_t. \]  \hspace{1cm} (5.2b)

Two different situations need to be considered:

1. The unobserved inputs \( u_t \in \mathbb{R}^{n_i \times 1} \) are white noise processes. This hypothesis is used in many real situations: for example, excitation due to wind in turbulent regime; roadway excitation, where many different cars act at or near the structure at different speeds (the road and pavement unevenness can also contribute to this hypothesis); and many others (see [87]).

If \( u_t \in \mathbb{R}^{n_i \times 1} \) are white noise processes, \((Bu_{t-1} + w_t)\) and \((Du_t + v_t)\) are also white noise processes. Then, model (5.2) can be used without any restrictions to compute the modes of vibration. This hypothesis is usually adopted in Operational Modal Analysis and gets very good results.

2. The unobserved inputs \( u_t \in \mathbb{R}^{n_i \times 1} \) are non-white noise or autocorrelated processes. This happens when the power spectral density function of the inputs is not constant. In these cases we can express \( u_t \in \mathbb{R}^{n_i \times 1} \) as a vector auto-regressive process of order \( p \), or \( \text{VAR}(p) \) (see [11], [93])

\[ u_t = H_1u_{t-1} + H_2u_{t-2} + \ldots + H_pu_{t-p} + e_t = \sum_{k=1}^{p} H_ku_{t-k} + e_t, \]  \hspace{1cm} (5.3)
where \( H_k \in \mathbb{R}^{n_k \times n_k} \) are coefficient matrices and \( e_t \in \mathbb{R}^{n_1 \times 1} \) is a zero mean white noise vector process with time invariant covariance matrix \( \Sigma \). From Eq. (5.3), we can write

\[
\begin{bmatrix}
  u_t \\
  u_{t-1} \\
  u_{t-2} \\
  \vdots \\
  u_{t-p+2} \\
  u_{t-p+1}
\end{bmatrix} =
\begin{bmatrix}
  H_1 & H_2 & H_3 & \ldots & H_{p-1} & H_p \\
  I_{n_1} & 0 & 0 & \ldots & 0 & 0 \\
  0 & I_{n_1} & 0 & \ldots & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & 0 & \ldots & I_{n_1} & 0
\end{bmatrix}
\begin{bmatrix}
  u_{t-1} \\
  u_{t-2} \\
  u_{t-3} \\
  \vdots \\
  u_{t-p+1} \\
  u_{t-p}
\end{bmatrix} +
\begin{bmatrix}
  \epsilon_t \\
  0 \\
  0 \\
  \vdots \\
  0 \\
  0
\end{bmatrix},
\]  

(5.4)

or in a more compact form

\[
\tilde{u}_t = H \tilde{u}_{t-1} + \tilde{\epsilon}_t,
\]  

(5.5)

where

\[
\tilde{u}_t =
\begin{bmatrix}
  u_t \\
  u_{t-1} \\
  u_{t-2} \\
  \vdots \\
  u_{t-p+2} \\
  u_{t-p+1}
\end{bmatrix} \in \mathbb{R}^{p_{n_1} \times 1},
\tilde{\epsilon}_t =
\begin{bmatrix}
  \epsilon_t \\
  0 \\
  0 \\
  \vdots \\
  0 \\
  0
\end{bmatrix} \in \mathbb{R}^{p_{n_1} \times 1},
\]  

(5.6)

\[
H =
\begin{bmatrix}
  H_1 & H_2 & H_3 & \ldots & H_{p-1} & H_p \\
  I_{n_1} & 0 & 0 & \ldots & 0 & 0 \\
  0 & I_{n_1} & 0 & \ldots & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & 0 & \ldots & I_{n_1} & 0
\end{bmatrix} \in \mathbb{R}^{p_{n_1} \times p_{n_1}}.
\]  

(5.7)

Matrix \( H \) is usually called the companion matrix. Taking into account simultaneously Eqs. (5.1) and (5.5), we can write

\[
\begin{bmatrix}
  x_t \\
  \tilde{u}_t
\end{bmatrix} =
\begin{bmatrix}
  A & \tilde{B} \\
  0 & H
\end{bmatrix}
\begin{bmatrix}
  x_{t-1} \\
  \tilde{u}_{t-1}
\end{bmatrix} +
\begin{bmatrix}
  w_t \\
  \tilde{\epsilon}_t
\end{bmatrix},
\]  

(5.8a)

\[
y_t =
\begin{bmatrix}
  C & \tilde{D}
\end{bmatrix}
\begin{bmatrix}
  x_t \\
  \tilde{u}_t
\end{bmatrix} + v_t,
\]  

(5.8b)

where

\[
\tilde{B} = [B \ 0_{n_x \times (p-1)n_1}], \quad \tilde{D} = [D \ 0_{n_x \times (p-1)n_1}].
\]  

(5.9)

In essence, Eq. (5.8) is equivalent to

\[
\tilde{x}_t = \tilde{A}\tilde{x}_{t-1} + \tilde{w}_t
\]  

(5.10a)

\[
y_t = \tilde{C}\tilde{x}_t + v_t,
\]  

(5.10b)

where

\[
\tilde{x}_t =
\begin{bmatrix}
  x_t \\
  \tilde{u}_t
\end{bmatrix}, \quad \tilde{A} =
\begin{bmatrix}
  A & \tilde{B} \\
  0 & H
\end{bmatrix}, \quad \tilde{w}_t =
\begin{bmatrix}
  w_t \\
  \tilde{\epsilon}_t
\end{bmatrix}, \quad \tilde{C} =
\begin{bmatrix}
  C & \tilde{D}
\end{bmatrix}.
\]  

(5.11)

So when we estimate model (5.2) from data recorded at a system with autocorrelated inputs, we actually obtain model (5.10), which includes an extra term, the matrix \( H \). It is easy to prove that the eigenvalues of matrix \( \tilde{A} \) are equal to the eigenvalues...
of matrix $A$ (modes of vibration) plus the eigenvalues of matrix $H$ (non-white components of the inputs), and it is very difficult to distinguish the ones from the others.

The problem of non-white noise or autocorrelated inputs has been analysed in some technical works, specially to handle harmonic loads. Extended OMA methods have been proposed in these cases, either assuming that the frequency of the harmonic disturbance is known \[ 72 \, 73 \], or via noise poles on the unit circle \[ 46 \, 74 \, 75 \]. A more general model was presented in Refs. \[ 83 \, 84 \]: it consists of the sum of a harmonic part (a sum of frequency modulated sinewaves) modelling the influence of the rotating components, and a noise part (filtered continuous-time band-limited white noise) modelling the response of the structure to the unobserved random perturbation.

In this Chapter we take advantage of the ideas presented in Chapters \[ 3 \, 4 \] to face with the problem. We propose to record the outputs of the structural system corresponding to different inputs and then to estimate a state space model based on Eq. (5.10): the eigenvalues that are common to all the records will correspond to the system and the eigenvalues that appear only in a given record will correspond to the input. As usual in this work, the model is estimated using the EM algorithm. A nice result of this model is that we obtain some information about the input acting in each record.

### 5.2 State space model for OMA with unobserved autocorrelated inputs

Consider $M$ different records for the output of a structural system that correspond to $M$ different (unknown) autocorrelated inputs,

$$Y^{(r)}_N = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_t^{(r)}, \ldots, y_N^{(r)}\}, \quad r = 1, 2, \ldots, M,$$

where $y_t^{(r)} \in \mathbb{R}^{n \times 1}$ is the observed vector for record $r$ and at time instant $t$. Taking into account Eqs. (5.10) and (5.11), we propose to use the following model to estimate the modes of vibration

$$\begin{bmatrix} x_{1,t}^{(r)} \\ x_{2,t}^{(r)} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12}^{(r)} \\ 0 & A_{22}^{(r)} \end{bmatrix} \begin{bmatrix} x_{1,t-1}^{(r)} \\ x_{2,t-1}^{(r)} \end{bmatrix} + \begin{bmatrix} w_{1,t}^{(r)} \\ w_{2,t}^{(r)} \end{bmatrix}, \quad \begin{bmatrix} w_{1,t}^{(r)} \\ w_{2,t}^{(r)} \end{bmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} Q_{11}^{(r)} & 0 \\ 0 & Q_{22}^{(r)} \end{bmatrix} \right),$$

$$y_t^{(r)} = \begin{bmatrix} C_{11} & C_{12}^{(r)} \end{bmatrix} \begin{bmatrix} x_{1,t}^{(r)} \\ x_{2,t}^{(r)} \end{bmatrix} + v_t^{(r)}, \quad v_t^{(r)} \sim \mathcal{N}(0, R^{(r)}),$$

$$r = 1, 2, \ldots, M.$$

The main property of this model is that $A_{11} \in \mathbb{R}^{n_1 \times n_1}$ and $C_{11} \in \mathbb{R}^{n_0 \times n_1}$ are constant for all the records, and $A_{12}^{(r)} \in \mathbb{R}^{n_1 \times n_2}$, $A_{22}^{(r)} \in \mathbb{R}^{n_2 \times n_2}$ and $C_{12}^{(r)} \in \mathbb{R}^{n_0 \times n_1}$ are different for each record. Assuming the system is time invariant, our hypothesis is that the eigenvalues of $A_{11}$ correspond to modes of vibration, and the eigenvalues of $A_{22}^{(r)}$ correspond to the autocorrelated inputs.

Another important issue of model (5.13) is that the covariances of $[w_{1,t}^{(r)} \ w_{2,t}^{(r)}]^T$, called $Q^{(r)} \in \mathbb{R}^{n_1 \times n_2}$, are block diagonal. This is important because forces the similarity
transformations (see Prop. A.2) to be block diagonal too:

\[
T^{(r)} = \begin{bmatrix} T_{11} & 0 \\ 0 & T_{22}^{(r)} \end{bmatrix}, \quad \left( T^{(r)} \right)^{-1} = \begin{bmatrix} T_{11}^{-1} & 0 \\ 0 & T_{22}^{(r)}^{-1} \end{bmatrix},
\]

(5.14)

where \( T_{11} \in \mathbb{R}^{n_x \times n_x} \) and \( T_{22}^{(r)} \in \mathbb{R}^{n_x \times n_x} \) are non-singular matrices. Therefore, when we estimate model \( 5.13 \) we actually obtain (see Section B.3.2):

\[
\begin{bmatrix} z_{1,t}^{(r)} \\ z_{2,t}^{(r)} \end{bmatrix} = \begin{bmatrix} T_{11}^{-1} A_{11} T_{11} \\ 0 \end{bmatrix} \begin{bmatrix} T_{11}^{-1} A_{12} T_{22}^{(r)} \\ T_{22}^{(r)} \end{bmatrix} \begin{bmatrix} z_{1,t-1}^{(r)} \\ z_{2,t-1}^{(r)} \end{bmatrix} + \begin{bmatrix} T_{11}^{-1} w_{1,t}^{(r)} \\ T_{22}^{(r)} \end{bmatrix},
\]

(5.15a)

\[
y_t^{(r)} = \begin{bmatrix} C_{11} T_{11} \\ C_{12} T_{22}^{(r)} \end{bmatrix} \begin{bmatrix} z_{1,t}^{(r)} \\ z_{2,t}^{(r)} \end{bmatrix} + v_t^{(r)},
\]

(5.15b)

\[ r = 1, 2, \ldots, M, \]

where \( T_{11} \in \mathbb{R}^{n_x \times n_x} \) and \( T_{22}^{(r)} \in \mathbb{R}^{n_x \times n_x} \) are not known, and

\[
\begin{bmatrix} T_{11}^{-1} w_{1,t}^{(r)} \\ T_{22}^{(r)} \end{bmatrix} \sim N \left( 0, \begin{bmatrix} T_{11}^{-1} Q_{11}^{(r)} \left( T_{11}^{-1} \right)^T \\ 0 \\ 0 \\ T_{22}^{(r)} \end{bmatrix} \begin{bmatrix} T_{22}^{-1} Q_{22}^{(r)} \left( T_{22}^{(r)} \right)^{-T} \end{bmatrix} \right). \]

So if matrices \( Q^{(r)} \) are block diagonal, then matrices \( T^{(r)} \) are block diagonal too, and vice-versa. Two additional important comments can be made:

1. The states \( z_t \) in Eq. 5.15 are related to the states \( x_t \) in Eq. 5.13 by mean of

\[
\begin{bmatrix} z_{1,t}^{(r)} \\ z_{2,t}^{(r)} \end{bmatrix} = \begin{bmatrix} T_{11}^{-1} \\ 0 \end{bmatrix} \begin{bmatrix} x_{1,t}^{(r)} \\ x_{2,t}^{(r)} \end{bmatrix} = \begin{bmatrix} T_{11}^{-1} x_{1,t}^{(r)} \\ T_{22}^{(r)} \end{bmatrix}.
\]

(5.17)

Therefore, the system states \( x_{1,t}^{(r)} \) and the states corresponding to the inputs \( x_{2,t}^{(r)} \) do not get mixed. However, if matrices \( Q^{(r)} \in \mathbb{R}^{n_x \times n_x} \) were full, the similarity transformations would be then

\[
\tilde{T} = \begin{bmatrix} \tilde{T}_{11} & \tilde{T}_{12}^{(r)} \\ 0 & \tilde{T}_{22}^{(r)} \end{bmatrix}, \quad \tilde{T}^{-1} = \begin{bmatrix} \tilde{T}_{11}^{-1} & -\tilde{T}_{12}^{(r)} \left( \tilde{T}_{22}^{(r)} \right)^{-1} \\ 0 & \left( \tilde{T}_{22}^{(r)} \right)^{-1} \end{bmatrix},
\]

(5.18)

and the states

\[
\begin{bmatrix} z_{1,t}^{(r)} \\ z_{2,t}^{(r)} \end{bmatrix} = \begin{bmatrix} \tilde{T}_{11}^{-1} \\ \tilde{T}_{11}^{-1} \tilde{T}_{12}^{(r)} \left( \tilde{T}_{22}^{(r)} \right)^{-1} \end{bmatrix} \begin{bmatrix} x_{1,t}^{(r)} \\ x_{2,t}^{(r)} \end{bmatrix} = \begin{bmatrix} \tilde{T}_{11}^{-1} x_{1,t}^{(r)} + \tilde{T}_{11}^{-1} \tilde{T}_{12}^{(r)} \left( \tilde{T}_{22}^{(r)} \right)^{-1} x_{2,t}^{(r)} \\ \left( \tilde{T}_{22}^{(r)} \right)^{-1} x_{2,t}^{(r)} \end{bmatrix}.
\]

(5.19)

Therefore, system states and input would be mixed. In conclusion, model 5.13 with block diagonal \( Q^{(r)} \) matrices will separate system states from inputs better than with full \( Q^{(r)} \) matrices.
2. We cannot estimate the input series \( u_t \) using model (5.13): indeed we obtain

\[
z_{2,\lambda}^{(r)} = T_{22}^{-1} x_{2,\lambda}^{(r)} = T_{22}^{-1} u_t^{(r)}
\]  

where \( T_{22}^{-1} \) is not known. So we estimate a linear transformation of the inputs.

Summarising, the state space model we propose to use with ambient data and non-white inputs is

\[
x_t^{(r)} = A^{(r)} x_{t-1}^{(r)} + w_t^{(r)}, \quad w_t^{(r)} \sim N(0, Q^{(r)}),
\]

\[
y_t^{(r)} = C^{(r)} x_t^{(r)} + v_t^{(r)}, \quad v_t^{(r)} \sim N(0, R^{(r)}),
\]

where

\[
A^{(r)} = \begin{bmatrix} A_{11}^{(r)} & A_{12}^{(r)} \\ 0 & A_{22}^{(r)} \end{bmatrix}, \quad C^{(r)} = \begin{bmatrix} C_{11}^{(r)} & C_{12}^{(r)} \end{bmatrix}, \quad Q^{(r)} = \begin{bmatrix} Q_{11}^{(r)} & 0 \\ 0 & Q_{22}^{(r)} \end{bmatrix}. \]  

The unknown parameters of this model are

\[
\theta = \{ A_{11}^{(r)}, A_{12}^{(r)}, A_{22}^{(r)}, C_{11}^{(r)}, C_{12}^{(r)}, Q_{11}^{(r)}, Q_{22}^{(r)}, R^{(r)}, x_0^{(r)}, P_0^{(r)} \}, \quad r = 1, 2, \ldots, M, \]

where \( x_0^{(r)} \) and \( P_0^{(r)} \) are the mean and variance of the initial states \( x_0^{(r)} \) respectively (which is assumed to be normal distributed).

Next we will show how to estimate this model using the EM algorithm.

### 5.3 Maximum likelihood estimation: EM algorithm

In essence, Eq. (5.21) is a model for multiple records. The likelihood for this kind of models was derived in Eq. (3.11)

\[
\log L_{Y_N}(\theta) = \log l_{Y_N}(\theta) = - \frac{n_N N M}{2} \log 2\pi - \frac{1}{2} \sum_{r=1}^{M} \sum_{t=1}^{N} \log \left| \Sigma_t^{(r)}(\theta^{(r)}) \right| - \frac{1}{2} \sum_{r=1}^{M} \sum_{t=1}^{N} \left( \epsilon_t^{(r)}(\theta^{(r)}) \right) ^T \left( \Sigma_t^{(r)}(\theta^{(r)}) \right)^{-1} \epsilon_t^{(r)}(\theta^{(r)}). \]

We describe now how to maximize Eq. (5.24) using the Expectation-Maximization algorithm. Consider we know both the observed outputs \( Y_N^{(r)} = \{ y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)} \} \) and the states \( X_N^{(r)} = \{ x_1^{(r)}, x_2^{(r)}, \ldots, x_N^{(r)} \} \). Then, the density function for one individual record is given by

\[
f_{\theta^{(r)}}(X_N^{(r)}, Y_N^{(r)}) = f_{\bar{x}_0^{(r)}, P_0^{(r)}}(x_0^{(r)}) \prod_{t=1}^{N} f_{A^{(r)}, Q^{(r)}}(X_t^{(r)} | X_{t-1}^{(r)}) \prod_{t=1}^{N} f_{C^{(r)}, R^{(r)}}(Y_t^{(r)} | X_t^{(r)}),
\]

where under Gaussian assumption

\[
f_{\bar{x}_0^{(r)}, P_0^{(r)}}(x_0^{(r)}) = \frac{1}{(2\pi)^{n_s/2} | P_0^{(r)} |^{1/2}} \exp \left( -\frac{1}{2} (x_0^{(r)} - \bar{x}_0^{(r)})^T (P_0^{(r)})^{-1} (x_0^{(r)} - \bar{x}_0^{(r)}) \right),
\]
where, ignoring constants, are functions

\[
f_{A(r), Q(r)}(X_t^{(r)} | X_{t-1}^{(r)}) = \frac{1}{(2\pi)^{n/2}|Q(r)|^{1/2}} \exp \left( -\frac{1}{2} (x_t^{(r)} - A(r)x_{t-1}^{(r)})^T Q(r)^{-1} (x_t^{(r)} - A(r)x_{t-1}^{(r)}) \right),
\]

\[
f_{C(r), R(r)}(X_t^{(r)} | x_t^{(r)}) = \frac{1}{(2\pi)^{n/2}|R(r)|^{1/2}} \exp \left( -\frac{1}{2} (y_t^{(r)} - C(r)x_t^{(r)})^T R(r)^{-1} (y_t^{(r)} - C(r)x_t^{(r)}) \right),
\]

Thus, if we consider \( M \) independent setups, the joint density function \( f_{\theta}(X_N, Y_N) \) will be the product of individual ones

\[
f_{\theta}(X_N, Y_N) = \prod_{r=1}^{M} f_{\theta(r)}(X_N^{(r)}, Y_N^{(r)}).
\]

The complete data likelihood is defined by \( L_{X_N, Y_N}(\theta) = f_{\theta}(X_N, Y_N) \). In practice the log-likelihood is used, so information is combined by addition and it can be written as a sum of the log-likelihood of each individual record:

\[
l_{X_N, Y_N}(\theta) = \log L_{X_N, Y_N}(\theta) = \sum_{r=1}^{M} l_{X_N^{(r)}, Y_N^{(r)}}(\theta^{(r)}).
\]

Finally, the log-likelihood of record \( r \) can be written as the sum of three uncoupled functions

\[
l_{X_N^{(r)}, Y_N^{(r)}}(\theta^{(r)}) = -\frac{1}{2}[l_1(\xi_0^{(r)}, P_0^{(r)}) + l_2(A^{(r)}, Q^{(r)}) + l_3(C^{(r)}, R^{(r)})],
\]

where, ignoring constants, are

\[
l_1(\xi_0^{(r)}, P_0^{(r)}) = \log |P_0^{(r)}| + (x_0^{(r)} - \bar{x}_0^{(r)})^T P_0^{(r)^{-1}} (x_0^{(r)} - \bar{x}_0^{(r)}),
\]

\[
l_2(A^{(r)}, Q^{(r)}) = N \log |Q^{(r)}| + \sum_{t=1}^{N} (x_t^{(r)} - A^{(r)}x_{t-1}^{(r)})^T Q^{(r)^{-1}} (x_t^{(r)} - A^{(r)}x_{t-1}^{(r)}),
\]

\[
l_3(C^{(r)}, R^{(r)}) = N \log |R^{(r)}| + \sum_{t=1}^{N} (y_t^{(r)} - C^{(r)}x_t^{(r)})^T R^{(r)^{-1}} (y_t^{(r)} - C^{(r)}x_t^{(r)}).
\]

The maximum likelihood estimation of the parameters \( \theta \) (see Eq. \ref{eq:5.23}) is obtained by maximizing the log-likelihood given by Eq. \ref{eq:5.26} using the EM algorithm. The details about the Expectation and Maximization steps are included in the following sections.

### 5.3.1 E-step: expectation step.

**Property 5.1.** Given the observed vectors \( Y_N^{(r)} = \{y_1^{(r)}, y_2^{(r)}, \ldots, y_N^{(r)}\} \) and a value for the parameters \( \theta_j \), then

\[
E[l_{X_N, Y_N}(\theta)|Y_N, \theta_j] = \sum_{r=1}^{M} E[l_{X_N^{(r)}, Y_N^{(r)}}(\theta^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] =
\]

\[
\sum_{r=1}^{M} \left( E[l_1(\xi_0^{(r)}, P_0^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] + E[l_2(A^{(r)}, Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] + E[l_3(C^{(r)}, R^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] \right)
\]
5.3 Maximum likelihood estimation: EM algorithm

with

\[
\mathbb{E} \left[ \log | P_0^{(r)} | + \operatorname{tr} \left( \left( P_0^{(r)} \right)^{-1} \left[ P_0^{N,(r)} + \left( x_0^{N,(r)} - \bar{x}_0^{(r)} \right) \left( x_0^{N,(r)} - \bar{x}_0^{(r)} \right)^T \right] \right) \right] =
\]

(5.32)

\[
\mathbb{E} \left[ l_2(A^{(r)}, Q^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] =
\]

(5.33)

\[
\mathbb{E} \left[ l_3(C^{(r)}, R^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] =
\]

(5.34)

\[
\text{The definition of } x_0^{N,(r)}, P_0^{N,(r)}, P_{1,d-1}^{N,(r)}, S_{xx}^{(r)}, S_{bb}^{(r)}, S_{xb}^{(r)}, S_{by}^{(r)} \text{ and } S_{yy}^{(r)} \text{ can be found at Def. 3.1 and Prop. 3.1.}
\]

Proof. We do not include the proof of this Property because is easily obtained from the proof of Prop. 3.1.

5.3.2 M-step: maximization step.

Property 5.2. The maximum of \( \mathbb{E}[l_{X,Y}^{(r)}(\theta) | Y_N, \theta_j] \) for model given by Eq. (5.21) is attained at

\[
\bar{x}_0^{(r)} = x_0^{N,(r)}, \quad r = 1, \ldots, M;
\]

(5.35)

\[
P_0^{(r)} = P_0^{N,(r)}, \quad r = 1, \ldots, M;
\]

(5.36)

\[
\text{vec}(A_{11}) = \sum_{r=1}^{M} \left( S_{bb1}^{(r)} - S_{bb2}^{(r)} \left( S_{bb2}^{(r)} \right)^{-1} S_{bb1}^{(r)} \right)^{-1} Q_{i11}^{(r)} + \sum_{r=1}^{M} \text{vec} \left[ Q_{i11}^{(r)} \left( S_{bb1}^{(r)} - S_{bb2}^{(r)} \left( S_{bb2}^{(r)} \right)^{-1} S_{bb1}^{(r)} \right) \right];
\]

(5.37)

\[
A_{22}^{(r)} = S_{bb2}^{(r)} \left( S_{bb2}^{(r)} \right)^{-1}, \quad r = 1, \ldots, M;
\]

(5.38)

\[
A_{12}^{(r)} = \left( S_{bb2}^{(r)} - A_{11}^{(r)} S_{bb2}^{(r)} \right) \left( S_{bb2}^{(r)} \right)^{-1}; \quad r = 1, \ldots, M;
\]

(5.39)

\[
Q_{11}^{(r)} = \frac{1}{N} \left[ S_{xx11}^{(r)} - S_{bb1}^{(r)} A_{11}^{(r)} - S_{bb2}^{(r)} \left( A_{12}^{(r)} \right)^T - A_{11}^{(r)} S_{bb1}^{(r)} - A_{21}^{(r)} S_{bb2}^{(r)} + A_{11}^{(r)} S_{bb1}^{(r)} A_{11}^{(r)} + A_{12}^{(r)} S_{bb2}^{(r)} A_{11}^{(r)} + A_{11}^{(r)} S_{bb2}^{(r)} \left( A_{12}^{(r)} \right)^T + A_{12}^{(r)} S_{bb2}^{(r)} \left( A_{12}^{(r)} \right)^T \right],
\]

(5.40)

\[
Q_{22}^{(r)} = \frac{1}{N} \left[ S_{xx22}^{(r)} - S_{bb2}^{(r)} \left( A_{22}^{(r)} \right)^T - A_{22}^{(r)} S_{bb2}^{(r)} + A_{22}^{(r)} S_{bb2}^{(r)} \left( A_{22}^{(r)} \right)^T \right], \quad r = 1, \ldots, M;
\]

(5.41)
5.3.3 Overall procedure

The two steps, expectation and maximization, have to be repeated iteratively until the likelihood is maximized. The overall method can be summarized as follows:

- Initialize the procedure \( (j = 0) \) selecting starting values for the parameters \( \theta_0 \) and a stop tolerance \( \delta_{adm} \).

- Repeat

1. Perform the E-Step. Apply the Kalman filter (Properties A.4, A.5 and A.6) to each record \( r \) to obtain the expected values \( x_t^{N,(r)}, P_t^{N,(r)}, \) and \( P_{t,t-1}^{N,(r)} \) with \( \theta_j \) as data. Use them to compute the matrices \( S_{xx}^{(r)}, S_{yy}^{(r)}, S_{xy}^{(r)} \) and \( S_{yx}^{(r)} \) given by (3.26)-(3.29).

2. Perform the M-Step. Compute the updated value of the parameters \( \theta_{j+1} \) using (3.39)-(3.44).

3. Compute the likelihood \( l_{Y_N}(\theta_{j+1}) \) with Equation (5.24).

4. Compute the actual tolerance

\[
\delta = \frac{|l_{Y_N}(\theta_{j+1}) - l_{Y_N}(\theta_j)|}{|l_{Y_N}(\theta_j)|} \tag{5.45}
\]

(a) If \( \delta > \delta_{adm} \), perform a new iteration with \( \theta_{j+1} \) as the value of the parameters.

(b) If \( \delta \leq \delta_{adm} \), stop the iterations. The estimate is \( \theta_{j+1} \).

5.4 Starting values for the EM algorithm

In this section we propose a procedure to build the starting point for the estimation of model (5.21) using the EM algorithm. Consider we know the model

\[
x_t^{(r)} = Ax_{t-1}^{(r)} + w_t^{(r)}, \quad w_t^{(r)} \sim N(0,Q^{(r)}),
\]

\[
y_t^{(r)} = Cx_t^{(r)} + v_t^{(r)}, \quad v_t^{(r)} \sim N(0,R^{(r)}),
\]

\( r = 1,2,\ldots,M, \)
that is, we know the common parameters of the records. If the records have specific components, then \( v_t^{(r)} \) will not be white noise because they will include the specific part of record \( r \). Therefore, we can estimate the model

\[
\begin{align*}
    x_t^{(r)} &= A_v^{(r)} x_{t-1}^{(r)} + b_t^{(r)}, \quad b_t^{(r)} \sim N(0, Q_v^{(r)}), \\
    v_t^{(r)} &= C_v^{(r)} x_t^{(r)} + a_t^{(r)}, \quad a_t^{(r)} \sim N(0, R_v^{(r)}),
\end{align*}
\]

for record \( r = 1, 2, \ldots, M \). Combining both models we have

\[
\begin{align*}
    \begin{bmatrix} x_t^{(r)} \\ z_t^{(r)} \end{bmatrix} &= \begin{bmatrix} A & 0 \\ 0 & A_v^{(r)} \end{bmatrix} \begin{bmatrix} x_{t-1}^{(r)} \\ z_{t-1}^{(r)} \end{bmatrix} + \begin{bmatrix} w_t^{(r)} \\ b_t^{(r)} \end{bmatrix}, \quad \begin{bmatrix} w_t^{(r)} \\ b_t^{(r)} \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q_v^{(r)} & 0 \\ 0 & Q_v^{(r)} \end{bmatrix} \right), \\
    y_t^{(r)} &= \begin{bmatrix} C & C_v^{(r)} \end{bmatrix} \begin{bmatrix} x_t^{(r)} \\ z_t^{(r)} \end{bmatrix} + a_t^{(r)}, \quad a_t^{(r)} \sim N(0, R_v^{(r)}),
\end{align*}
\]

We propose to use Eq. (5.46) and SSI to build the starting point (as usual, the objective is the starting point to be easy to build and take advantage of the available information). The procedure is:

- Select one record, \( y_t^{(k)} \), and apply SSI to it using \( n_s = n_{s1} \). The obtained parameters are \( A_k, C_k, Q_k \) and \( R_k \). Consider these parameters are common to all the records.
- Apply the Kalman filter (Property A.4) to each record using \( A_k, C_k, Q_k \) and \( R_k \). The result are the innovations \( e_t^{(r)} \), what can be considered as an estimate of \( v_t^{(r)} \).
- Apply SSI to the innovations using \( n_s = n_{s2} \). The obtained parameters are \( A_v^{(r)}, C_e^{(r)}, Q_e^{(r)} \) and \( R_e^{(r)} \) for each record \( r \).
- Finally, taking into account Eq. (5.46), the starting point we propose to use for the EM algorithm is

\[
\theta_0 = \left( A_0^{(r)}, C_0^{(r)}, Q_0^{(r)}, R_0^{(r)}, \bar{x}_0^{(r)}, \bar{P}_0^{(r)} \right)
\]

where

\[
\begin{align*}
    A_0^{(r)} &= \begin{bmatrix} A_k & 0 \\ 0 & A_v^{(r)} \end{bmatrix}, \\
    Q_0^{(r)} &= \begin{bmatrix} Q_k & 0 \\ 0 & Q_v^{(r)} \end{bmatrix}, \\
    C_0^{(r)} &= \begin{bmatrix} C_k & C_e^{(r)} \end{bmatrix}, \\
    R_0^{(r)} &= R_e^{(r)}.
\end{align*}
\]

This starting point is easy and fast to build, and gets very good results in practice.

### 5.5 Example 1: simulated data

#### 5.5.1 Description of the data

We use the simulated system presented in Sec. 3.5 but in this Chapter the inputs have the following characteristics:
We simulate three different records (\(M = 3\)):

- **Records 1 and 2**: Each input is an AR(2) process, that is, it responds to the equation

\[
  u_t = h_1 u_{t-1} + h_2 u_{t-2} + e_t, \quad e_t \sim N(0, \sigma_e^2), \quad (5.47)
\]

where \(h_1, h_2\) and \(\sigma_e^2\) are indicated in Table 5.1. According to Eq. (5.4), an AR(2) can be written as

\[
  \begin{bmatrix}
  u_t \\
  u_{t-1}
  \end{bmatrix} = \begin{bmatrix}
  h_1 & h_2 \\
  1 & 0
  \end{bmatrix} \begin{bmatrix}
  u_{t-1} \\
  u_{t-2}
  \end{bmatrix} + \begin{bmatrix}
  e_t \\
  0
  \end{bmatrix}. \quad (5.48)
\]

The eigenvalues of the companion matrix are the roots of the characteristic polynomial, that is, the roots of

\[
  \lambda^2 - h_1 \lambda - h_2 = 0. \quad (5.49)
\]

When \((h_1^2 + 4h_2 < 0)\), the roots \(\lambda_1\) and \(\lambda_2\) are complex (see [11]). If we express the complex eigenvalues as

\[
  \lambda_1, \lambda_2 = \exp \left[ -\zeta u \omega_u \pm i \omega u \sqrt{1 - \zeta^2} \right] \Delta t, \quad (5.50)
\]

we can compute an equivalent frequency and damping for the inputs by mean of

\[
  \omega_u = \frac{\ln (\lambda_1)}{\Delta t}, \quad (5.51)
\]

\[
  \zeta_u = -\text{Real} \left[ \frac{\ln (\lambda_1)}{\omega_u \Delta t} \right]. \quad (5.52)
\]

This information has also been included in Table 5.1. The resulting inputs are shown in Figure 5.1 both in the time and frequency domain. Input \(u_1\) is applied to mass number 2 and input \(u_1\) is applied to mass number 6.

- **Record 3**: The input is Gaussian white noise, \(u_t \sim N(0, \sigma_u^2 = 0.5)\), and has been applied to all masses.

- **Sampling frequency** \(f_s = 50\ \text{Hz}\). Total duration of signals, 100 seconds (5000 time steps).

- The observed value is the sum of the structure response and a sensor Gaussian noise with variance equal to the 20% of the largest acceleration response variance.
Figure 5.1: Inputs for simulated records 1 and 2 (PSD has been estimated using Yule-Walker AR method).

5.5.2 Results of the individual analysis

First, we are going to analyse each record individually. Figure 5.2 shows the stabilization diagram for the simulated record 1. We have estimated the state space model (2.2),

\[ x_t = Ax_{t-1} + w_t, \quad w_t^{(r)} \sim N(0, Q^{(r)}), \]
\[ y_t = Cx_t + v_t, \quad v_t^{(r)} \sim N(0, R^{(r)}), \]

for model orders 4 to 32 using the EM algorithm as explained in Chapter 2 (the starting point was the parameters estimated using SSI). We observe that seven stable poles have been identified in the range of the analysed orders:

- Six of them correspond to modes of vibration: modes 1, 2, 3, 4, 5, 6 (\( f_1 = 2.94 \text{ Hz} \), \( f_2 = 5.87 \text{ Hz} \), \( f_3 = 8.60 \text{ Hz} \), \( f_4 = 11.19 \text{ Hz} \), \( f_5 = 13.78 \text{ Hz} \), \( f_6 = 16.52 \text{ Hz} \));

- The other stable pole corresponds to the input (\( f_{u_1} = 7.24 \text{ Hz} \)). This example shows clearly that when we estimate the state space model

\[ x_t = Ax_{t-1} + w_t \]
\[ y_t = Cx_t + v_t \]

from output only data in the presence of autocorrelated inputs, we actually obtain the extended model

\[ \tilde{x}_t = \tilde{A}\tilde{x}_{t-1} + \tilde{w}_t \]
\[ y_t = \tilde{C}\tilde{x}_t + v_t, \]
where
\[ \tilde{x}_t = \begin{bmatrix} x_t \\ \tilde{u}_t \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} A & B \\ 0 & H \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C & D \end{bmatrix}. \]

Matrix \( H \) contains the input characteristics. In this example, the eigenvalues of matrix \( H \) are the roots of the characteristic polynomial for input \( u_t^{(1)} \) (see Table 5.1).

- Modes 7 (\( f_7 = 19.54 \) Hz) and 8 (\( f_8 = 23.12 \) Hz) have not been identified, even using a state space order equal to 32 (twice the theoretical, 16). This is because these modes have not been excited (see the frequency content of the input \( u_t^{(1)} \) in Figure 5.1).

- The results obtained with the theoretical order, \( n_s = 16 \), are given in Table 5.2.

In simulated record 2, the input has been built with a dominant frequency very close to the natural frequency of mode 5. The objective is to analyse what happens when we have input frequencies close to system natural frequencies. Figure 5.3 shows the stabilization diagram obtained for this record and Table 5.2 includes the results obtained with the theoretical order, \( n_s = 16 \). The main conclusions are:

- Eight stable poles have been identified.
- Seven correspond to modes 2, 3, 4, 5, 6, 7 and 8 (\( f_2 = 5.87 \) Hz, \( f_3 = 8.60 \) Hz, \( f_4 = 11.19 \) Hz, \( f_5 = 13.78 \) Hz, \( f_6 = 16.52 \) Hz, \( f_7 = 19.54 \) Hz, \( f_8 = 23.12 \) Hz);
- The other stable pole corresponds to the load \( u_2 \) (\( f_{u_2} = 14.33 \) Hz).
- Mode 1 (\( f_1 = 2.94 \) Hz) has not been identified in any of the analysed orders: this mode has not been excited (see the frequency content of the input \( u_t^{(2)} \) in Figure 5.1).

Last, in Record 3, whose input is white noise, eight stable poles are obtained: they correspond to the modes of vibration. The corresponding stabilization diagram is given in Figure 5.3 and the values obtained with the theoretical order, in Table 5.2.

5.5.3 Results of the joint analysis

After that we estimated the joint model (3.5)
\[
\begin{align*}
\begin{bmatrix} x_t^{(r)} \\ u_t^{(r)} \end{bmatrix} & = \begin{bmatrix} A & B \\ 0 & H \end{bmatrix} \begin{bmatrix} x_{t-1}^{(r)} \\ u_{t-1}^{(r)} \end{bmatrix} + \begin{bmatrix} w_t^{(r)} \\ v_t^{(r)} \end{bmatrix}, \\
\begin{bmatrix} y_t^{(r)} \\ v_t^{(r)} \end{bmatrix} & = \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} x_t^{(r)} \\ u_t^{(r)} \end{bmatrix} + \begin{bmatrix} v_t^{(r)} \end{bmatrix}, \\
& \quad w_t^{(r)} \sim N(0, Q^{(r)}), \\
& \quad v_t^{(r)} \sim N(0, R^{(r)}), \\
& \quad r = 1, 2, \ldots, M,
\end{align*}
\]

using the three simulated records, that is, we estimated the parameters common to all the records. We want to check the performance of this model when there are “specific eigenvalues” due to the inputs in addition to the eigenvalues of the system. Figure 5.3 shows the stabilization diagram obtained using this model for model orders 4 to 32; the starting point in this case was built applying SSI to record 3 (the input in this record is white noise, so the parameters estimated by SSI will be a good estimation for the common parameters). Because of a single matrix \( A \) and \( C \) are obtained, the stabilization diagram is built in the usual way. Some interesting facts are observed:
5.5 Example 1: simulated data

Figure 5.2: Stabilization diagram for simulated record 1 using model (2.2) and the EM algorithm.

<table>
<thead>
<tr>
<th>Mode</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>$u_{1}^{(1)}$</th>
<th>$u_{2}^{(2)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural frequencies (Hz)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td>2.94</td>
<td>5.87</td>
<td>8.60</td>
<td>11.19</td>
<td>13.78</td>
<td>16.52</td>
<td>19.54</td>
<td>23.12</td>
<td>7.24</td>
<td>14.33</td>
</tr>
<tr>
<td>Record 1</td>
<td>2.92</td>
<td>5.87</td>
<td>8.60</td>
<td>11.16</td>
<td>13.80</td>
<td>16.60</td>
<td>20.02</td>
<td>-</td>
<td>7.23</td>
<td>-</td>
</tr>
<tr>
<td>Record 2</td>
<td>-</td>
<td>5.89</td>
<td>8.59</td>
<td>11.19</td>
<td>13.88</td>
<td>14.33</td>
<td>19.51</td>
<td>22.88</td>
<td>-</td>
<td>16.51</td>
</tr>
<tr>
<td>Record 3</td>
<td>3.22</td>
<td>5.94</td>
<td>8.70</td>
<td>11.27</td>
<td>13.86</td>
<td>16.60</td>
<td>19.59</td>
<td>23.17</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Damping ratios (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td>2.00</td>
<td>1.24</td>
<td>1.10</td>
<td>1.10</td>
<td>1.15</td>
<td>1.23</td>
<td>1.35</td>
<td>1.50</td>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Record 1</td>
<td>2.02</td>
<td>1.13</td>
<td>1.64</td>
<td>2.06</td>
<td>1.28</td>
<td>2.36</td>
<td>15.16</td>
<td>-</td>
<td>2.08</td>
<td>-</td>
</tr>
<tr>
<td>Record 2</td>
<td>-</td>
<td>1.26</td>
<td>1.25</td>
<td>1.44</td>
<td>1.367</td>
<td>1.33</td>
<td>1.03</td>
<td>1.46</td>
<td>-</td>
<td>1.04</td>
</tr>
<tr>
<td>Record 3</td>
<td>21.92</td>
<td>3.26</td>
<td>1.77</td>
<td>1.25</td>
<td>0.85</td>
<td>0.93</td>
<td>0.96</td>
<td>0.87</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MAC values between theoretical and estimated modes shapes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 1</td>
<td>0.989</td>
<td>0.999</td>
<td>0.995</td>
<td>0.997</td>
<td>0.995</td>
<td>0.846</td>
<td>0.004</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Record 2</td>
<td>-</td>
<td>0.981</td>
<td>0.994</td>
<td>0.991</td>
<td>0.957</td>
<td>1.000</td>
<td>0.998</td>
<td>0.986</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Record 3</td>
<td>0.998</td>
<td>0.900</td>
<td>0.985</td>
<td>0.973</td>
<td>0.980</td>
<td>0.971</td>
<td>0.966</td>
<td>0.888</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.2: Natural frequencies, damping ratios and MAC values estimated from multiple records (simulated data) using model (2.2) and the EM algorithm with $n_s = 16$. The theoretical values are also included for comparison.

- For state space orders comprise between 4 and 16, model (5.5) has estimated the modal parameters (common parameters), and not the eigenvalues corresponding to the inputs (specific parameters). However, when we analysed each record individually, the input poles were estimated at very low orders.
At the theoretic state space order, $n_s = 16$, the modal parameters were estimated with precision (except mode 1). The numeric results obtained at this order are given in Table 5.3.

We observe the benefits of the joint estimation because for this order the parameters were not estimated properly in Record 1 and 2.

For state orders higher than 16, model (3.5) also estimated the input eigenvalues.

<table>
<thead>
<tr>
<th>Mode</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$ (Hz)</td>
<td>3.11</td>
<td>5.94</td>
<td>8.60</td>
<td>11.18</td>
<td>13.81</td>
<td>16.32</td>
<td>19.48</td>
<td>23.14</td>
</tr>
<tr>
<td>$\zeta$ (%)</td>
<td>14.68</td>
<td>1.73</td>
<td>1.55</td>
<td>1.65</td>
<td>1.12</td>
<td>1.94</td>
<td>2.26</td>
<td>2.46</td>
</tr>
<tr>
<td>MAC</td>
<td>0.998</td>
<td>0.995</td>
<td>0.998</td>
<td>0.996</td>
<td>0.986</td>
<td>0.977</td>
<td>0.987</td>
<td>0.992</td>
</tr>
</tbody>
</table>

Table 5.3: Natural frequencies, damping ratios and MAC values estimated from simulated data using model (3.5) and $n_s = 16$.

### 5.5.4 Results using the proposed model

Our objective now is to estimate the modes of vibration inherent to the system and present in greater or lesser extent in all the records, and the input poles, present only in specific records. Figure 5.6 shows the stabilization diagram computed using the proposed state space model, that is, Eq. (5.21). This plot has been built varying $n_{s1}$ from 4 to 32;
5.5 Example 1: simulated data

Figure 5.4: Stabilization diagram for simulated record 3 using model (2.2) and the EM algorithm.

Figure 5.5: Stabilization diagram built using the joint state space model (3.5).
the order of the specific part, \( n_{s2} \), was equal to 2 for all the values of \( n_{s1} \). The starting point for the EM has been built according to Section 5.4 with \( y_k^{(1)} = y_k^{(3)} \).

We see that the proposed model works very well in practice: the eigenvalues of the system and the eigenvalues of the input are properly estimated. And what is more important, we know what eigenvalues correspond to the system and what to the input. In Table 5.4 we present the results obtained with the proposed model for estate space orders \( n_{s1} = 16 \) and \( n_{s2} = 2 \) (the theoretical orders). The results are quite accurate, both in the modal parameters and in the inputs.

![Figure 5.6: Stabilization diagram for the simulated example using state space model (5.21). \( \oplus \) stands for stable common parameters, \( \Box \) for stable specific parameters, and \( + \) for unstable parameters.

<table>
<thead>
<tr>
<th>Mode</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>( w_1^{(1)} )</th>
<th>( w_1^{(2)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural frequencies (Hz)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td>2.94</td>
<td>5.87</td>
<td>8.60</td>
<td>11.19</td>
<td>13.78</td>
<td>16.52</td>
<td>19.54</td>
<td>23.12</td>
<td>7.24</td>
<td>14.33</td>
</tr>
<tr>
<td>Model (5.21)</td>
<td>2.95</td>
<td>5.87</td>
<td>8.59</td>
<td>11.18</td>
<td>13.78</td>
<td>16.51</td>
<td>19.53</td>
<td>23.11</td>
<td>7.24</td>
<td>14.36</td>
</tr>
<tr>
<td>Damping ratios (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td>2.00</td>
<td>1.24</td>
<td>1.10</td>
<td>1.10</td>
<td>1.15</td>
<td>1.23</td>
<td>1.35</td>
<td>1.50</td>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Model (5.21)</td>
<td>4.50</td>
<td>1.27</td>
<td>1.27</td>
<td>1.13</td>
<td>1.15</td>
<td>1.32</td>
<td>1.39</td>
<td>1.48</td>
<td>2.20</td>
<td>0.98</td>
</tr>
<tr>
<td>MAC values between theoretical and estimated modes shapes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model (5.21)</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>0.997</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.4: Natural frequencies, damping ratios and MAC values estimated from multiple records (simulated data) using model (5.21) with \( n_{s1} = 16 \) and \( n_{s2} = 2 \).
5.5.5 Input estimation

An important question is that using model \( (5.21) \) we can estimate the input at each record. According to Eq. \( (5.20) \), the states \( x^{(r)}_{2,t} \) are a linear combination of the inputs, but not the inputs. So we cannot take the estimate of \( x^{(r)}_{2,t} \) as the estimate of the inputs. Another possibility is to consider the data given by the term \( C^{(r)} x^{(r)}_{2,t} \), that is, the projection of the inputs onto the measured accelerations. The advantage of this approach is that we can estimate the input acting at a given point; the disadvantage is that we obtain a scaled input, that is, the product of a constant by the input (and this constant is not known).

This approach is based in Property B.7: when the measured values in a structural system are accelerations, the observation equation is

\[
y_t = C x_t + D u_t + v_t = \begin{bmatrix} -C_a M^{-1} K & -C_a M^{-1} C \end{bmatrix} \begin{bmatrix} q_t \\ \dot{q}_t \end{bmatrix} + C_a M^{-1} B u_t + v_t.\]

(see Appendix B for the meaning of the symbols). Taking into account this equation, we can estimate a “scaled input” by mean of:

- Estimate model \( (5.21) \), that is, \( \hat{A}^{(r)} \), \( \hat{C}^{(r)} \), \( \hat{Q}^{(r)} \) and \( \hat{R}^{(r)} \) using the EM algorithm.
- Apply the Kalman filter (Property A.4) using the above parameters to obtain an estimate of the states, \( \hat{x}_{t-1,1}^{(r)} \), and the innovations \( \hat{e}_t^{(r)} \). Therefore, we can express the observation equation of model \( (5.21) \) as

\[
y_t^{(r)} = \hat{C}^{(r)} \hat{x}_{t-1,1}^{(r)} + \hat{e}_t^{(r)} = \hat{C}_{11} \hat{x}_{1,t}^{t-1,1} + \hat{C}_{12} \hat{x}_{2,t}^{t-1,1} + \hat{e}_t^{(r)}. \tag{5.53}
\]

- We can take \( \hat{C}_{11} \hat{x}_{1,t}^{t-1,1} \) as an estimate of \( C x_t \) and \( \hat{C}_{12} \hat{x}_{2,t}^{t-1,1} \) as an estimate of \( C_a M^{-1} B u_t \). Due to the matrices \( C_a \) and \( B_u \), the projection on a given channel is a good approximation to the loads acting near that sensor.

We have applied this procedure to the simulated example using \( n_{s1} = 16 \) and \( n_{s2} = 2 \). According to the above indications, the best values for \( u_t^{(1)} \) will be obtained using the acceleration of mass 2 in record 1, and the best values for \( u_t^{(2)} \) will be obtained using the acceleration of mass 6 in record 2. The results are shown in Figure 5.7, both in the time and in the frequency domain. We observe that the projection of the estimated states \( \hat{x}_{2,t}^{t-1,1} \) onto the selected accelerations are a good estimation of the inputs except for the presence of a constant.

5.6 Example 2: gym building

In addition to the tests presented in Section 4.7, the gym building was tested using people jumping at certain frequencies. The tests were classified in function of the number of people jumping (groups of 1, 6, 12, 18, 24, and 30 people) and the frequency of the jump (1.5 Hz, 2.0 Hz and 3.0 Hz). The frequencies of the jumps were indicated using a metronome, and all the groups jumped at all the frequencies.

Thirteen sensors were placed in the gym floor to measure the vibrations due to the jump. The position of these sensors in shown in Figure 5.8; this figure also includes the region devoted to people jumps. When only one person jumped, the jumping load was
Figure 5.7: Estimation of $u_t^{(1)}$ and $u_t^{(2)}$. $u_t^{(1)}$: projection of the estimated states $\hat{x}_{2,t}^{t-1,(r)}$ onto the recorded accelerations at channel 2, record 1; $u_t^{(2)}$: projection of the estimated states $\hat{x}_{2,t}^{t-1,(r)}$ onto the recorded accelerations at channel 6, record 2; $k_1 = \frac{\text{std}[u_t^{(1)}]}{\text{std}[u_{t,p}^{(1)}]} = 7.60$, $k_2 = \frac{\text{std}[u_t^{(2)}]}{\text{std}[u_{t,p}^{(2)}]} = 4.60$, where std means standard deviation.
5.6 Example 2: gym building

Figure 5.8: Sensor position (●) and jumping place in the gym building.

measured by mean of a force plate. Finally, the vibrations due to ambient loads were also recorded.

In this section we are going to analyse two records to check the proposed method:

• The first record corresponds to one person (87 kg) jumping at a 2.5 Hz. The data measured by the force plate is shown in Figure 5.9 both in the time domain and in the frequency domain. We see that the force plate gives \( F = 0 \) N when the person is not jumping \((t = 0)\). Therefore, when the person is in the air, the obtained value is minus the person weight \( F_z = -mg \approx -850 \) N.

In the frequency domain, we see that the jump is not perfectly harmonic with frequency 2.5 Hz, because the energy is distributed among 2.5 Hz and integer harmonics of it, that is, 5.0 Hz, 7.5 Hz, 10.0 Hz, 12.5 Hz and 15.0 Hz.

The output measured by the sensor number 8 due to jump is shown in Figure 5.10.

• The second record corresponds to ambient vibration. We use this record because we want the two record are very different so the model is able to distinguish common parameters (the modes of vibration) from specific parameters (the jump).

The output measured by the sensor number 8 due to ambient load is shown in Figure 5.10 as well. We observe that: the vibration due to jumping load is higher than the ambient vibration; the Power Spectral Density (PSD) function of the jumping load has some peaks that are not present in the ambient load PSD. Specifically, these peaks correspond to the jumping frequency and harmonics.

This example has only two records but in one of then the input is clearly non-white noise (or autocorrelated), so we can check the performance of the proposed model. First, we have built the stabilization diagram estimating the state space model (5.21) from order \( n_{s1} = 4 \) to \( n_{s1} = 70 \). An important question is the order for \( A_{22}(r) \): since we have six harmonics in the range of frequencies \( 0 - 16 \) Hz, we are going to take \( n_{s2} = 12 \).

We have built two different stabilization diagrams: one for the common part of model (5.21), that is, taking into account matrices \( A_{11} \) and \( C_{11} \) (Figure 5.11(a)), and another for the specific part, taking into account matrices \( A_{22}(r) \) and \( C_{12}(r) \) (Figure 5.11(b)). This figures show that the method is able to estimate the modes of vibration and the parameters of the input, even in real an complicated cases like the analysed one.

Finally, Figure 5.12 shows the estimated jumping force using the same procedure that in the simulated case. The values have been obtained using \( n_{s1} = 40 \) and \( n_{s2} = 12 \).
Figure 5.9: Jumping force (2.5 Hz) recorded by the force plate in the gym building.

Figure 5.10: Vibration recorded by sensor number 8 due to ambient loading and jumping at 2.5 Hz.
5.6 Example 2: gym building

(a) Stabilization diagram for $A_{11}$ and $C_{11}$ in model (5.21).

(b) Stabilization diagram for $A_{22}^{(r)}$ and $C_{22}^{(r)}$ in model (5.21).

Figure 5.11: Stabilization diagrams for the simulated example using state space model (5.21).
5.7 Conclusions

Estimating the state space model from output-only data when the (unmeasured) system inputs are autocorrelated has the problem that additional eigenvalues of matrix A are obtained, and they correspond to the input and not to the modes of vibration. To separate the ones from the others, some extra information is needed.

In this Chapter we have proposed a state space model that can be used for estimating the modes of vibration in these cases. The idea is to measure the response of the structure to different inputs; then, we assume that the parameters common to all the records will correspond to the structure, and the rest to the inputs.

The A and C matrices of the proposed model have two parts: one common to all the records and other specific to each record. The eigenvalues of the common A matrix give us the modal parameters, meanwhile the eigenvalues of the specific A matrix give us the characteristics of the inputs. This model can be estimated successfully using the EM algorithm. We would like to remark that no prior assumptions about the inputs are needed in this model.

The proposed method has been tested with simulated data and with measured data. In both cases, the results obtained with the proposed method were very good.

Finally, some information about the inputs can be obtained. However, this aspect needs to be analysed in more detail in future research.

Figure 5.12: Jumping load estimated using the vibration recorded by sensor number 8 (the shown values have been multiplied by $5\times10^4$ to get adequate scale).
5.8 Appendix

Proof of Property 5.2

The proof of Eqs. (5.35) - (5.36) is similar to the one given in Prop. 3.2. For the proof of Eqs. (5.37), (5.38), (5.39), (5.40) and (5.41), the starting equation is Eq. (5.33)

\[ E \left[ l_2(A^{(r)}, Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)} \right] = \]

\[ = N \log |Q^{(r)}| + tr \left( Q^{(r)}_i \left[ S_{xx}^{(r)} - S_{xb}^{(r)} (A^{(r)})^T - A^{(r)} S_{bx}^{(r)} + A^{(r)} S_{bb}^{(r)} (A^{(r)})^T \right] \right), \]

where \( Q_i^{(r)} \) stands for \( (Q^{(r)})^{-1} \). Taking into account Eq. (5.22), we can write

\[ E \left[ l_2(A^{(r)}, Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)} \right] = N \log |Q_{11}^{(r)}| + N \log |Q_{22}^{(r)}| + \]

\[ + tr \left\{ \begin{bmatrix} Q_{11}^{(r)} & 0 \\ 0 & Q_{22}^{(r)} \end{bmatrix} \begin{bmatrix} S_{xx11}^{(r)} & S_{xx12}^{(r)} \\ S_{xx12}^{(r)} & S_{xx22}^{(r)} \end{bmatrix} \right\} \]

\[ - tr \left\{ \begin{bmatrix} Q_{11}^{(r)} & 0 \\ 0 & Q_{22}^{(r)} \end{bmatrix} \begin{bmatrix} S_{bb11}^{(r)} & S_{bb12}^{(r)} \\ S_{bb12}^{(r)} & S_{bb22}^{(r)} \end{bmatrix} \right\} \]

\[ - tr \left\{ \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} S_{bb11}^{(r)} & S_{bb12}^{(r)} \\ S_{bb12}^{(r)} & S_{bb22}^{(r)} \end{bmatrix} \right\} \]

\[ + tr \left\{ \begin{bmatrix} Q_{11}^{(r)} & 0 \\ 0 & Q_{22}^{(r)} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} S_{bb11}^{(r)} & S_{bb12}^{(r)} \\ S_{bb12}^{(r)} & S_{bb22}^{(r)} \end{bmatrix} \right\} \]

Multiplying matrices and grouping terms we obtain

\[ E \left[ l_2(A^{(r)}, Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)} \right] = N \log |Q_{11}^{(r)}| + N \log |Q_{22}^{(r)}| + \]

\[ + tr \left\{ Q_{11}^{(r)} \left( S_{xx11}^{(r)} - S_{xx12}^{(r)} A_{11}^{(r)} + S_{xx12}^{(r)} A_{12}^{(r)} - A_{11} S_{bx11}^{(r)} + A_{12} S_{bx12}^{(r)} + A_{11} S_{bb11}^{(r)} + A_{12} S_{bb12}^{(r)} + A_{11} S_{bb21}^{(r)} + A_{12} S_{bb22}^{(r)} \right) \right\} \]

\[ + tr \left\{ Q_{22}^{(r)} \left( S_{xx22}^{(r)} - S_{xx22}^{(r)} A_{22}^{(r)} + S_{xx22}^{(r)} A_{22}^{(r)} + A_{22} S_{bb22}^{(r)} \right) \right\}. \]

Now we take the derivative of this expression to find the maximum:

- \( \frac{\partial}{\partial A_{11}} E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial A_{11}} \sum_{r=1}^{M} E[l_2(A^{(r)}, Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] = 0; \)

\[ \frac{\partial}{\partial A_{11}} \sum_{r=1}^{M} tr \left( -Q_{11}^{(r)} S_{xb11}^{(r)} A_{11}^{(r)} \right) = - \sum_{r=1}^{M} Q_{11}^{(r)} S_{xb11}^{(r)}; \]
\[ \frac{\partial}{\partial A_{11}} \sum_{r=1}^{M} \text{tr} \left( -Q_{i11}^{(r)} A_{11} S_{b11}^{(r)} \right) = - \sum_{r=1}^{M} Q_{i11}^{(r)} S_{xb11}^{(r)}; \]

\[ \frac{\partial}{\partial A_{11}} \sum_{r=1}^{M} \text{tr} \left( Q_{i11}^{(r)} A_{21}^{(r)} S_{b21}^{(r)} A_{11}^T \right) = \sum_{r=1}^{M} Q_{i11}^{(r)} A_{21}^{(r)} S_{b21}^{(r)}; \]

\[ \frac{\partial}{\partial A_{11}} \sum_{r=1}^{M} \text{tr} \left( Q_{i11}^{(r)} A_{11} S_{b12}^{(r)} \left( A_{21}^{(r)} \right)^T \right) = \sum_{r=1}^{M} Q_{i11}^{(r)} A_{21}^{(r)} S_{b21}^{(r)}; \]

\[ \frac{\partial}{\partial A_{11}} \sum_{r=1}^{M} \text{tr} \left( Q_{i11}^{(r)} A_{11} S_{b12}^{(r)} A_{11}^T \right) = 2 \sum_{r=1}^{M} Q_{i11}^{(r)} A_{11} S_{b12}^{(r)}; \]

Equating to zero

\[ \sum_{r=1}^{M} (Q_{i11}^{(r)} A_{11} S_{b12}^{(r)} + Q_{i11}^{(r)} A_{21}^{(r)} S_{b21}^{(r)} = Q_{i11}^{(r)} S_{xb11}^{(r)} ) \]. \quad (5.54) \]

\[ \frac{\partial}{\partial A_{12}} \text{E}[l_{X_N,Y_N}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial A_{12}} \sum_{r=1}^{M} \text{E}[l_2(A^{(r)}, Q^{(r)})|Y_N^{(r)}, \theta_j^{(r)}] = 0; \]

\[ \frac{\partial}{\partial A_{12}} \sum_{r=1}^{M} \text{tr} \left( -Q_{i11}^{(r)} S_{xb12}^{(r)} \left( A_{12}^{(r)} \right)^T \right) = -Q_{i11}^{(r)} S_{xb12}^{(r)}; \]

\[ \frac{\partial}{\partial A_{12}} \sum_{r=1}^{M} \text{tr} \left( -Q_{i11}^{(r)} A_{12}^{(r)} S_{b21}^{(r)} \right) = -Q_{i11}^{(r)} S_{xb12}^{(r)}; \]

\[ \frac{\partial}{\partial A_{12}} \sum_{r=1}^{M} \text{tr} \left( Q_{i11}^{(r)} A_{12}^{(r)} S_{b21}^{(r)} A_{11}^T \right) = Q_{i11}^{(r)} A_{11} S_{b12}^{(r)}; \]

\[ \frac{\partial}{\partial A_{12}} \sum_{r=1}^{M} \text{tr} \left( Q_{i11}^{(r)} A_{11} S_{b12}^{(r)} \left( A_{12}^{(r)} \right)^T \right) = Q_{i11}^{(r)} A_{11} S_{b12}^{(r)}; \]

\[ \frac{\partial}{\partial A_{12}} \sum_{r=1}^{M} \text{tr} \left( Q_{i11}^{(r)} A_{12}^{(r)} S_{b22}^{(r)} \left( A_{12}^{(r)} \right)^T \right) = 2Q_{i11}^{(r)} A_{12}^{(r)} S_{b22}^{(r)}; \]

Equating to zero

\[ Q_{i11}^{(r)} A_{11} S_{b12}^{(r)} + Q_{i11}^{(r)} A_{12}^{(r)} S_{b22}^{(r)} = Q_{i11}^{(r)} S_{xb12}^{(r)} , \quad r = 1, \ldots, M; \] \quad (5.55)

From Eq. (5.55) it is straightforward

\[ A_{12}^{(r)} = \left( S_{xb12}^{(r)} - A_{11} S_{b12}^{(r)} \right) \left( S_{b22}^{(r)} \right)^{-1}, \quad r = 1, \ldots, M, \] \quad (5.56)

what corresponds to Eq. (5.59). Substituting Eq. (5.56) into Eq. (5.54) and operating we have

\[ \sum_{r=1}^{M} \left[ Q_{i11}^{(r)} A_{11} \left( S_{b11}^{(r)} - S_{b12}^{(r)} \left( S_{b22}^{(r)} \right)^{-1} S_{b21}^{(r)} \right) \right] = \]
\[ = \sum_{r=1}^{M} \left[ Q_{1i1}^{(r)} \left( S_{x_{b11}}^{(r)} - S_{x_{b12}}^{(r)} \left( S_{b_{b22}}^{(r)} \right)^{-1} S_{b_{b21}}^{(r)} \right) \right]; \]

Taking the \( \text{vec}(\Box) \) operator we finally obtain Eq. (5.37).

- \( \frac{\partial}{\partial A_{22}^{(r)}} \text{E}[l_{X_N, Y_N}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial A_{22}^{(r)}} \sum_{r=1}^{M} \text{E}[l_2(A^{(r)}, Q^{(r)})|Y_N, \theta_j^{(r)}] = 0; \)

\[
\frac{\partial}{\partial A_{22}^{(r)}} \sum_{r=1}^{M} \text{tr} \left( -Q_{122}^{(r)} S_{x_{b22}}^{(r)} \left( A_{22}^{(r)} \right)^T \right) = -Q_{122}^{(r)} S_{x_{b22}}^{(r)};
\]

\[
\frac{\partial}{\partial A_{22}^{(r)}} \sum_{r=1}^{M} \text{tr} \left( -Q_{122}^{(r)} A_{22}^{(r)} S_{b_{b22}}^{(r)} \right) = -Q_{122}^{(r)} S_{b_{b22}}^{(r)};
\]

\[
\frac{\partial}{\partial A_{22}^{(r)}} \sum_{r=1}^{M} \text{tr} \left( Q_{122}^{(r)} A_{22}^{(r)} S_{b_{b22}}^{(r)} \left( A_{22}^{(r)} \right)^T \right) = 2Q_{122}^{(r)} A_{22}^{(r)} S_{b_{b22}}^{(r)};
\]

Equating to zero

\[
Q_{122}^{(r)} A_{22}^{(r)} S_{b_{b22}}^{(r)} = Q_{122}^{(r)} S_{x_{b22}}^{(r)}; \quad r = 1, \ldots, M, \quad (5.57)
\]

what easily gives Eq. (5.38).

- \( \frac{\partial}{\partial Q_{11}^{(r)}} \text{E}[l_{X_N, Y_N}(\theta)|Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial Q_{11}^{(r)}} \sum_{r=1}^{M} \text{E}[l_2(A^{(r)}, Q^{(r)})|Y_N, \theta_j^{(r)}] = 0; \)

\[
\frac{\partial}{\partial Q_{11}^{(r)}} \left( N \log |Q_{11}^{(r)}| \right) = N \left( (Q_{11}^{(r)})^{-1} \right)^T = N(Q_{11}^{(r)})^T;
\]

\[
\frac{\partial}{\partial Q_{11}^{(r)}} \text{tr} \left( Q_{1i1}^{(r)} \left[ S_{x_{x11}}^{(r)} - S_{x_{x12}}^{(r)} \left( A_{12}^{(r)} \right)^T - A_{12}^{(r)} S_{x_{x2}}^{(r)} + A_{12}^{(r)} S_{x_{b2}}^{(r)} \left( A_{12}^{(r)} \right)^T \right] \right) =
\]

\[
- \left( Q_{1i1}^{(r)} \left[ S_{x_{x11}}^{(r)} - S_{x_{x12}}^{(r)} \left( A_{12}^{(r)} \right)^T - A_{12}^{(r)} S_{x_{x2}}^{(r)} + A_{12}^{(r)} S_{x_{b2}}^{(r)} \left( A_{12}^{(r)} \right)^T \right] \right) Q_{1i1}^{(r)}^T;
\]

\[
\frac{\partial}{\partial Q_{11}^{(r)}} \text{tr} \left\{ Q_{1i1}^{(r)} \left[ S_{x_{x11}}^{(r)} - S_{x_{x12}}^{(r)} A_{12}^{(r)} T - A_{12}^{(r)} S_{x_{x2}}^{(r)} + A_{12}^{(r)} S_{x_{b2}}^{(r)} \left( A_{12}^{(r)} \right)^T \right] \right\} =
\]

\[
- \left[ Q_{1i1}^{(r)} \left[ S_{x_{x11}}^{(r)} - S_{x_{x12}}^{(r)} A_{12}^{(r)} T - A_{12}^{(r)} S_{x_{x2}}^{(r)} + A_{12}^{(r)} S_{x_{b2}}^{(r)} \left( A_{12}^{(r)} \right)^T \right] \right] Q_{1i1}^{(r)}^T; \]

Equating to zero we obtain Eq. (5.40):

\[
N(Q_{1i1}^{(r)})^T = \left[ Q_{1i1}^{(r)} \left[ S_{x_{x11}}^{(r)} - S_{x_{x12}}^{(r)} A_{12}^{(r)} T - A_{12}^{(r)} S_{x_{x2}}^{(r)} + A_{12}^{(r)} S_{x_{b2}}^{(r)} \left( A_{12}^{(r)} \right)^T \right] \right] Q_{1i1}^{(r)}^T.
\]
\[ Q_{i11}^{(r)} = \frac{1}{N} \left[ S_{x_{i11}}^{(r)} - S_{x_{2i1}}^{(r)} A_{i11}^{(r)} - S_{x_{b_{1i2}}}^{(r)} (A_{11}^{(r)})^T - A_{11} S_{x_{b_{1i1}}}^{(r)} - 2 A_{21} \phi^{(r)}_{b_{1i2}} + A_{11} S_{b_{1i1}}^{(r)} A_{i11}^{(r)} + A_{11} S_{x_{b_{1i1}}}^{(r)} A_{i11}^{(r)} + A_{11} S_{b_{2i1}}^{(r)} (A_{11}^{(r)})^T + A_{11} S_{b_{1i2}}^{(r)} (A_{11}^{(r)})^T \right]. \]

- \[ \frac{\partial}{\partial Q_{22}^{(r)}} \log |Q_{22}^{(r)}| \right) = N \left( (Q_{22}^{(r)})^{-1} \right)^T = N Q_{22}^{(r)T}; \]

\[ \frac{\partial}{\partial Q_{22}^{(r)}} \left( (Q_{22}^{(r)})^{-1} \right)^T = \frac{N}{Q_{22}^{(r)T}} \]

Equating to zero we obtain Eq. (5.11):

\[ N Q_{22}^{(r)T} = \left( Q_{i22}^{(r)} \left[ S_{x_{2i2}}^{(r)} - S_{x_{2i2}}^{(r)} A_{i22}^{(r)} + A_{22} S_{x_{2i2}}^{(r)} (A_{22}^{(r)})^T \right] \right)^T \]

\[ \Rightarrow Q_{22}^{(r)} = \frac{1}{N} \left[ S_{x_{2i2}}^{(r)} - S_{x_{2i2}}^{(r)} A_{i22}^{(r)} + A_{22} S_{x_{2i2}}^{(r)} (A_{22}^{(r)})^T \right]. \]

For the proof of Eqs. (5.42), (5.43) and (5.44), the starting equation is Eq. (5.32)

\[ E \left[ \log (C(r), R^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] = N \log |R^{(r)}| + \]

\[ + \text{tr} \left( R_i^{(r)} \left[ S_{y_{ix}}^{(r)} - S_{y_{ix}}^{(r)} C_{C_{11}}^{(r)} C_{C_{12}}^{(r)} - C_{C_{12}}^{(r)} S_{y_{ix}}^{(r)} + C_{C_{12}}^{(r)} S_{y_{ix}}^{(r)} \right] \right), \]

where we have used \( R_i^{(r)} \) for representing \( (R^{(r)})^{-1} \). Taking into account Eq. (5.22), we can write

\[ E \left[ \log (C(r), R^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] = N \log |R^{(r)}| + \text{tr} \left\{ R_i^{(r)} S_{y_{ix}}^{(r)} \right\} + \text{tr} \left\{ R_i^{(r)} \right\} \left[ \begin{array}{cc} S_{y_{ix11}}^{(r)} & S_{y_{ix12}}^{(r)} \\ C_{C_{11}}^{T} & C_{C_{12}}^{T} \end{array} \right] + \text{tr} \left\{ R_i^{(r)} \right\} \left[ \begin{array}{cc} S_{x_{ix11}}^{(r)} & S_{x_{ix12}}^{(r)} \\ C_{C_{11}}^{T} & C_{C_{12}}^{T} \end{array} \right] \]

Multiplying matrices we have

\[ E \left[ \log (C(r), R^{(r)}) | Y_N^{(r)}, \theta_j^{(r)} \right] = N \log |R^{(r)}| + \text{tr} \left\{ R_i^{(r)} \right\} \left[ \begin{array}{cc} S_{y_{ix11}}^{(r)} & S_{y_{ix12}}^{(r)} \\ C_{C_{11}}^{T} & C_{C_{12}}^{T} \end{array} \right] + \text{tr} \left\{ R_i^{(r)} \right\} \left[ \begin{array}{cc} S_{x_{ix11}}^{(r)} & S_{x_{ix12}}^{(r)} \\ C_{C_{11}}^{T} & C_{C_{12}}^{T} \end{array} \right] \]

Taking derivatives and equating to zero:
Thus, we obtain Eq. (5.42)

$$\partial = \frac{\partial C}{\partial C_{11}} \sum_{r=1}^{M} \text{tr} \left( -R_i^{(r)} S_{y_{11}y}^{(r)} C_{11}^{T} \right) = - \sum_{r=1}^{M} R_i^{(r)} S_{y_{11}y}^{(r)};$$

$$\partial = \frac{\partial C}{\partial C_{11}} \sum_{r=1}^{M} \text{tr} \left( -R_i^{(r)} C_{11} S_{y_{11}y}^{(r)} \right) = - \sum_{r=1}^{M} R_i^{(r)} S_{y_{11}y}^{(r)};$$

$$\partial = \frac{\partial C}{\partial C_{11}} \sum_{r=1}^{M} \text{tr} \left( R_i^{(r)} C_{11} S_{y_{11}y}^{(r)} C_{11}^{T} \right) = 2 \sum_{r=1}^{M} R_i^{(r)} C_{11} S_{y_{11}y}^{(r)};$$

$$\partial = \frac{\partial C}{\partial C_{11}} \sum_{r=1}^{M} \text{tr} \left( R_i^{(r)} C_{11} S_{y_{11}y}^{(r)} C_{11}^{T} \right) = 2 \sum_{r=1}^{M} R_i^{(r)} C_{11} S_{y_{11}y}^{(r)};$$

Equating to zero:

$$\sum_{r=1}^{M} \left( R_i^{(r)} C_{11} S_{y_{11}y}^{(r)} + R_i^{(r)} C_{12} S_{y_{21}y}^{(r)} = R_i^{(r)} S_{y_{11}y}^{(r)} \right). \quad (5.58)$$

Thus, we obtain Eq. (5.52)

$$C_{12}^{(r)} = \left( S_{y_{12}y}^{(r)} - C_{11} S_{y_{11}y}^{(r)} \right) \left( S_{y_{22}y}^{(r)} \right)^{-1}, \quad r = 1, \ldots, M; \quad (5.59)$$
On the other hand, substituting Eq. (5.59) into (5.58) we have:

\[
\sum_{r=1}^{M} R_i^{(r)} C_{11} \left( S_{xx11}^{(r)} - S_{xx12}^{(r)} \left( S_{xx22}^{(r)} \right)^{-1} S_{xx21}^{(r)} \right) =
\]

\[
= \sum_{r=1}^{M} \left[ \left( R^{(r)} \right)^{-1} \left( S_{yy11}^{(r)} - S_{yy12}^{(r)} \left( S_{xx22}^{(r)} \right)^{-1} S_{xx21}^{(r)} \right) \right].
\]

Taking the vec(□) operator we finally obtain Eq. (5.37):

\[
\operatorname{vec}(C_{11}) = \sum_{r=1}^{M} \left( S_{xx11}^{(r)} - S_{xx12}^{(r)} \left( S_{xx22}^{(r)} \right)^{-1} S_{xx21}^{(r)} \right) \otimes \left( R^{(r)} \right)^{-1}.
\]

\[
\cdot \sum_{r=1}^{M} \operatorname{vec} \left[ \left( R^{(r)} \right)^{-1} \left( S_{yy11}^{(r)} - S_{yy12}^{(r)} \left( S_{xx22}^{(r)} \right)^{-1} S_{xx21}^{(r)} \right) \right];
\]

\[
\bullet \frac{\partial}{\partial R^{(r)}} \mathbb{E}[\lambda_{X_N,Y_N}(\theta) | Y_N, \theta_j] = 0 \Rightarrow \frac{\partial}{\partial R^{(r)}} \sum_{r=1}^{M} \mathbb{E}[\lambda_3(C^{(r)}, R^{(r)}) | Y_N^{(r)}, \theta_j^{(r)}] = 0;
\]

\[
\frac{\partial}{\partial R^{(r)}} (N \log |R^{(r)}|) = N(R_i^{(r)})^T;
\]

\[
\frac{\partial}{\partial R^{(r)}} \operatorname{tr} \left( R_i^{(r)} \left[ S_{yy}^{(r)} - S_{yx}^{(r)} \left( C^{(r)} \right)^T - C^{(r)} S_{xy}^{(r)} + C^{(r)} S_{xx}^{(r)} \left( C^{(r)} \right)^T \right] \right) =
\]

\[
- \left( R_i^{(r)} \left[ S_{yy}^{(r)} - S_{yx}^{(r)} \left( C^{(r)} \right)^T - C^{(r)} S_{xy}^{(r)} + C^{(r)} S_{xx}^{(r)} \left( C^{(r)} \right)^T \right] R_i^{(r)} \right)^T;
\]

Equating to zero we obtain Eq. (5.44):

\[
N(R_i^{(r)})^T - \left( R_i^{(r)} \left[ S_{yy}^{(r)} - S_{yx}^{(r)} \left( C^{(r)} \right)^T - C^{(r)} S_{xy}^{(r)} + C^{(r)} S_{xx}^{(r)} \left( C^{(r)} \right)^T \right] R_i^{(r)} \right)^T = 0
\]

\[
\Rightarrow R^{(r)} = \frac{1}{N} \left( S_{yy}^{(r)} - S_{yx}^{(r)} \left( C^{(r)} \right)^T - C^{(r)} S_{xy}^{(r)} + C^{(r)} S_{xx}^{(r)} \left( C^{(r)} \right)^T \right).\]
Chapter 6

Conclusions and future research

6.1 Conclusions

This thesis discussed the applications of the EM algorithm in Operational Modal Analysis of structural systems. The work contains both theoretical developments and application to real cases. Although detailed conclusions are provided at the end of each chapter, in our opinion two main conclusions have been derived:

1. Several state space models can be used in order to solve different problems arising in an OMA context. Using these specialized state space models improves the estimation of the modal parameters.

2. The EM algorithm can be applied to obtain the maximum likelihood estimation of these models and gets very good results.

More specifically, the main contributions of this work can be summarized as follows:

- Chapter 2
  - The EM algorithm for state space model is presented, including the derivation of the equations.
  - Two methods for constructing the starting point are proposed and compared.
  - The EM algorithm is applied to simulated data and to data measured in a real structure. The results are compared to SSI.
  - The main results of this chapter have been presented at [16, 17, 18].

- Chapter 3
  - The problem of multiple records is presented: sometimes we have several records corresponding to the same structure. The objective is to estimate the modal parameters using all the records.
  - The state space model is extended to estimate the modal parameter from multiple records.
  - The model is estimated using the EM algorithm. The equations are derived in detail.
  - The method is applied to an steel arch bridge.
Conclusions and future research

The main results of this chapter were presented at [19, 20, 15].

• Chapter 4

– A new state space model is proposed to solve the problem of multiple setups of sensors.
– The model is estimated using the EM algorithm. The equations are derived in detail.
– The method is applied to a footbridge with 13 setups of sensors. Each setup is composed by 21 accelerometers (longitudinal, transversal and vertical directions are measured at seven different points per setup).
– The proposed method is compared with the traditional approach of estimating the modal parameters from each setup and them combining these partial estimates.
– The main results of this chapter were presented at [21, 22].

• Chapter 5

– In Operational Modal Analysis, the non-white components of the inputs become part of the system (that is, become part of $A$ matrix in the state space model). This is shown mathematically.
– A new method is proposed to separate the components of the inputs from the system.
– A new state space model is defined and estimated using the EM algorithm. The equations are derived in detail.
– Some characteristics of the unmeasured inputs are estimated.

• Appendix B

– The equations for computing the modal parameters from the state space model are derived in detail.

6.2 Future research

During the development of this work, some interesting future lines of research were identified. Some of them are here provided and briefly explained.

1. To extend the EM algorithm to estimate the model

\[ x_t = Ax_{t-1} + w_t \]
\[ y_t = Cx_t + v_t, \]

considering the white noise processes $\{w_t\}$ and $\{v_t\}$ are correlated, that is

\[ \text{Cov}[w_t, v_t] = S. \]

According to the theory, the general model is

\[ x_t = Ax_{t-1} + Bu_{t-1} + w_t \]
\[ y_t = Cx_t + Du_t + v_t; \]

(6.1) (6.2)
6.2 Future research

In OMA, \((Bu_{t-1} + w_t)\) and \((Du_t + v_t)\) are modelled by white noise processes, but it is clear they are correlated, so \(S \neq 0\). It is important to analyse the influence of the hypothesis adopted in this work, that is, \(S = 0\).

2. To apply the EM algorithm to estimate the estate space model

\[
x_t = Ax_{t-1} + Bu_{t-1} + w_t \tag{6.3}
\]

\[
y_t = Cx_t + Du_t + v_t. \tag{6.4}
\]

when some inputs are known (OMAX).

3. The state space model for multiple setups of sensors (Eq. (4.11)) can be extended to OMAX too. Then, the data corresponding to the shaker can be used in the estimation of the modes of the gym (see Section 4.7).

4. It is important to note that model (5.13) is equivalent to model

\[
\begin{bmatrix}
\tilde{x}_{1,t}^{(r)} \\
\tilde{x}_{2,t}^{(r)}
\end{bmatrix}
= \begin{bmatrix}
\tilde{A}_{11} & 0 \\
0 & \tilde{A}_{22}
\end{bmatrix}
\begin{bmatrix}
\tilde{x}_{1,t-1}^{(r)} \\
\tilde{x}_{2,t-1}^{(r)}
\end{bmatrix}
+ \begin{bmatrix}
\tilde{w}_{1,t}^{(r)} \\
\tilde{w}_{2,t}^{(r)}
\end{bmatrix}, \quad \begin{bmatrix}
\tilde{x}_{1,t}^{(r)} \\
\tilde{x}_{2,t}^{(r)}
\end{bmatrix}
\sim N(0, \tilde{Q}^{(r)}) \tag{6.5a}
\]

\[
y_t^{(r)} = \begin{bmatrix}
\tilde{C}_{11} & \tilde{C}_{12}
\end{bmatrix}
\begin{bmatrix}
\tilde{x}_{1,t}^{(r)} \\
\tilde{x}_{2,t}^{(r)}
\end{bmatrix}
+ \tilde{v}_t^{(r)}, \quad \tilde{v}_t^{(r)} \sim N(0, \tilde{R}^{(r)}), \tag{6.5b}
\]

because the state representation is not unique (Property A.2): we can define a matrix \(T\)

\[
T = \begin{bmatrix}
T_{11} & T_{12} \\
0 & T_{22}
\end{bmatrix}, \quad T^{-1} = \begin{bmatrix}
T_{11}^{-1} & T_{12}^{-1}T_{11}^{-1}T_{22}^{-1} \\
0 & T_{22}^{-1}
\end{bmatrix}, \tag{6.6}
\]

and a state transformation

\[
\begin{bmatrix}
\tilde{x}_{1,t}^{(r)} \\
\tilde{x}_{2,t}^{(r)}
\end{bmatrix}
= \begin{bmatrix}
T_{11} & T_{12} \\
0 & T_{22}
\end{bmatrix}
\begin{bmatrix}
\tilde{x}_{1,t}^{(r)} \\
\tilde{x}_{2,t}^{(r)}
\end{bmatrix},
\]

so the following relationships hold

\[
A_{11} = T_{11}^{-1}\tilde{A}_{11}T_{11}, \quad A_{22} = T_{22}^{-1}\tilde{A}_{22}T_{22}, \tag{6.7}
\]

\[
A_{12}^{(r)} = T_{11}^{-1}\tilde{A}_{11}T_{12} - T_{11}^{-1}T_{12}T_{22}^{-1}\tilde{A}_{22}T_{22},
\]

\[
C_{11} = \tilde{C}_{11}T_{11}, \quad C_{12}^{(r)} = \tilde{C}_{11}T_{12} + \tilde{C}_{12}T_{22}.
\]

Taking into account Eq. (6.7) it is easy to prove that the eigenvalues of matrix \(\tilde{A}_{11}\) are the same than the eigenvalues of \(A_{11}\), and the eigenvalues of matrix \(\tilde{A}_{22}\) are the same than the eigenvalues of \(A_{22}\). Therefore, similar to model (5.13), the eigenvalues of \(\tilde{A}_{11}\) correspond to the structural system, and the eigenvalues of \(\tilde{A}_{12}\) correspond to the non-white inputs.

Both, model (6.5) and model (5.5), can be used in the case of autocorrelated inputs, and Model (6.5) is simpler than (5.13).
5. One important problem in all the analysed models is the determination of the model order. In this work most of order have been chosen by inspection in a stabilization diagram. We want to explore other approaches, for example the concept of modal contribution presented in [23]. It would be interesting to check this method with the EM and to extend it to the case of multiple records.
Appendix A

The state space model

A.1 Introduction

The state space model is a well known model in the field of control theory. In fact, this is where it was born \[58\] and where it has reached maturity as a scientific discipline.

The first major development of the theory of control took place during the Second World War in the Radiation Laboratory of the Massachusetts Institute of Technology. The Radiation Laboratory team consisted of engineers, physicists and mathematicians who were engaged in solving various engineering problems associated with war, like the development of radar and advanced fire control. The laboratory, which dealt with control problems, worked with frequency response methods that were based on the works done by engineers such as Nyquist and Bode. These methods are still used today.

In the late 1950’s, at the University of Columbia, a researcher named Rudolf E. Kalman defended the use of state space instead of the classical methods in frequency. In a first moment his ideas were received with scepticism that prompted him to publish his results in a mechanical journal \[58\] though he was an electrical engineer. The definite support to the new method took place when Kalman visited the NASA Ames Research Center in 1960, who successfully used the new method in the Apollo Program\(^1\).

In this work we describe the state space model, and more specifically, the discrete, linear and time invariant state space model. If one looks at the number of adjectives used, it may seem that we are staying with a restricted class of model (especially by the fact that it is linear), but surprisingly, many processes can be described by this type of models.

The idea of state is a basic concept in the representation of systems in state space. If we try to describe intuitively what means a state in a system, we must remember that we are dealing with dynamic systems. In a generic dynamic system, the output at a given instant of time depends on the inputs to the system at that moment and the situation where the system is due to the previous inputs. In R. E. Kalman words \[58\], “How is a dynamic system (linear or non-linear) described? The fundamental concept is the notion of the state. By this is meant, intuitively, some quantitative information (a set of numbers, a function, etc.) which is the least amount of data one has to know about the past behaviour of the system in order to predict its future behaviour. The dynamics is then described in terms of state transitions, i.e., one must specify how one state is

\(^1\)On July 20, 1969 Neil Armstrong and Edwin Buzz Aldrin aboard Apollo 11, landed on the Moon in the Sea of Tranquillity (Mare Tranquillitatis).
transformed into another as time passes”.

Methods of analysis and modelling of systems based on state space are best suited for problems with multiple inputs and outputs, as well as for problems with time-varying parameters and even non-linear systems.

### A.2 State space model in continuous time

#### A.2.1 Linear Time-Invariant systems

A system converts a certain set of signals, called input signals $u(t)$, into another set of signals, called output signals $y(t)$. Systems analysed in this work are dynamical systems. These systems have, in addition to the input and output signals, a third signal that is called the state. This is an internal signal of the system, which can be thought of as the memory of the system.

**Definition A.1.** The state of a system at time $t_0$ is defined as the minimal information that is sufficient to determine the state and the output of the system for $t \geq t_0$ when the input to the system is known for $t \geq t_0$.

A continuous-time state space system can be defined as:

**Definition A.2.** A continuous-time state space system can be represented as

\[
\dot{x}(t) = f(t, x(t), u(t)) \tag{A.1a}
\]
\[
y(t) = g(t, x(t), u(t)), \tag{A.1b}
\]

where

- $u(t) \in \mathbb{R}^{n_i}$ is the vector of $n_i$ input signals;
- $y(t) \in \mathbb{R}^{n_o}$ is the vector of $n_o$ output signals;
- $x(t) \in \mathbb{R}^{n_s}$ is the state vector of the system at time $t$.

The first equation means that the evolution of the state depends on the state at time $t$ and on the input. The second equation means that the system output can be obtained from the state of the system and from the inputs.

An important class of systems are the linear systems.

**Definition A.3** (Linear system). The state space system (A.1a)-(A.1b) is a linear system if the functions $f$ and $g$ are linear functions with respect to $x(t)$ and $u(t)$.

For example, the following equations represent a linear state space system:

\[
\dot{x}(t) = A_c(t)x(t) + B_c(t)u(t) \tag{A.2a}
\]
\[
y(t) = C_c(t)x(t) + D_c(t)u(t), \tag{A.2b}
\]

with $A_c(t) \in \mathbb{R}^{n_s \times n_s}$, $B_c(t) \in \mathbb{R}^{n_s \times n_i}$, $C_c(t) \in \mathbb{R}^{n_o \times n_s}$ and $D_c(t) \in \mathbb{R}^{n_o \times n_i}$.

Another important class of systems is the class of time-invariant systems.
A.2 State space model in continuous time

Definition A.4 (Time-invariant system). The state space system (A.1a) of (A.1b) is a time-invariant system if it is described by functions $f$ and $g$ that do not change over time:

\[
\begin{align*}
\dot{x}(t) &= f(x(t), u(t)) \quad \text{(A.3a)} \\
y(t) &= g(x(t), u(t)) \quad \text{(A.3b)}
\end{align*}
\]

The system given by (A.2) is a linear time-varying (LTV) state space system. If the system matrices $A_c(t)$, $B_c(t)$, $C_c(t)$ and $D_c(t)$ do not depend on the time $t$, the system is time-invariant. In this work, the attention will be focused on linear time-invariant (LTI) systems.

A.2.2 Equations of state space model in continuous time

According to the previous section, the LTI state space model is defined by the following set of differential equations (subscript $c$ indicates continuous time):

Definition A.5. The state space model in continuous time for a LTI system is defined by

\[
\begin{align*}
\dot{x}(t) &= A_c x(t) + B_c u(t) \quad \text{(A.4a)} \\
y(t) &= C_c x(t) + D_c u(t) \quad \text{(A.4b)}
\end{align*}
\]

where

- $y(t) \in \mathbb{R}^{n_o}$ is the output vector;
- $u(t) \in \mathbb{R}^{n_i}$ is the input vector;
- $x(t) \in \mathbb{R}^{n_s}$ is the state vector;
- $A_c \in \mathbb{R}^{n_s \times n_s}$ is the transition state matrix describing the dynamics of the system;
- $B_c \in \mathbb{R}^{n_s \times n_i}$ is the input matrix;
- $C_c \in \mathbb{R}^{n_o \times n_s}$ is the output matrix, which is describing how the internal state is transferred to the output measurements $y(t)$;
- $D_c \in \mathbb{R}^{n_o \times n_i}$ is the direct transmission matrix.

Equation (A.4a) is known as the State Equation and equation (A.4b) is known as the Observation Equation. Thus, the system dynamics is described by mean of first-order differential equations.

A.2.3 Solution of state equation in continuous time

In this section we are going to seek a solution to the initial value problem given by

\[
\begin{align*}
\dot{x}(t) &= A_c x(t) + B_c u(t) \quad \text{(A.5a)} \\
t = t_0 \Rightarrow x(t) &= x(t_0) \quad \text{(A.5b)}
\end{align*}
\]

using the integration factor method ([39], [94]). Rewriting Equation (A.5a) as

\[
\left( \frac{dx(t)}{dt} - A_c x(t) \right) = B_c u(t),
\]

(A.6)
and premultiplying each side by $e^{-A_c t}$:

$$e^{-A_c t} \left( \frac{d}{dt} x(t) - A_c x(t) \right) = e^{-A_c t} B_c u(t). \quad (A.7)$$

We note that:

$$\frac{d}{dt} (e^{-A_c t} x(t)) = e^{-A_c t} \frac{d}{dt} x(t) - A_c e^{-A_c t} x(t) = e^{-A_c t} \left( \frac{d}{dt} x(t) - A_c x(t) \right).$$

Therefore, Equation \((A.7)\) becomes

$$\frac{d}{dt} (e^{-A_c t} x(t)) = e^{-A_c t} B_c u(t), \quad (A.8)$$

and, upon integrating between $t_0$ and $t$ yields:

$$e^{-A_c t} x(t) = k + \int_{t_0}^{t} e^{-A_c s} B_c u(s) ds,$$

$$x(t) = e^{A_c t} k + e^{A_c t} \int_{t_0}^{t} e^{-A_c s} B_c u(s) ds.$$

Taking into account the initial values

$$x(t_0) = e^{A_c t_0} k \Rightarrow k = e^{-A_c t_0} x(t_0), \quad (A.9)$$

$$x(t) = e^{A_c (t-t_0)} x(t_0) + e^{A_c t} \int_{t_0}^{t} e^{-A_c s} B_c u(s) ds. \quad (A.10)$$

The first term of the solution is the system response due to the initial conditions $x(t_0)$, and the second one is the system response due to the inputs $u(t)$.

### A.3 State space model in discrete time

Measurements are taken in discrete time rather than continuous time, so equations must be expressed in discrete time too. Typical for the sampling of a continuous-time signal is a Zero-Order Hold (ZOH) assumption, which means that the signal is piecewise constant over the sampling period (see Figure \ref{fig:discrete_time}). For the continuous state space model \((A.4)\), this is expressed as

$$\forall t \in [t_k, t_{k+1}) = [k\Delta t, (k+1)\Delta t) \Rightarrow x(t) = x(t_k) = x_k, \ u(t) = u(t_k) = u_k, \ y(t) = y(t_k) = y_k. \quad (A.11)$$

**Property A.1.** Assuming zero order hold discretization of input and output signals with sampling period equal to $\Delta t$, the continuous state space model given by Eq. \((A.4)\) is converted into

$$x_{k+1} = A_d x_k + B_d u_k \quad (A.12a)$$

$$y_k = C_d x_k + D_d u_k, \quad (A.12b)$$

where
A.3 State space model in discrete time

Figure A.1: Zero-Order Hold approximation of \( \sin x \).

\[ y_k \in \mathbb{R}^{n_o} \text{ is the discrete output vector}; \]
\[ u_k \in \mathbb{R}^{n_i} \text{ is the discrete input vector}; \]
\[ x_k \in \mathbb{R}^{n_s} \text{ is the discrete state vector}; \]
\[ A_d \in \mathbb{R}^{n_s \times n_s} \text{ is the discrete state matrix}; \]
\[ B_d \in \mathbb{R}^{n_s \times n_i} \text{ is the discrete input matrix}; \]
\[ C_d \in \mathbb{R}^{n_o \times n_s} \text{ is the discrete output matrix}; \]
\[ D_d \in \mathbb{R}^{n_o \times n_i} \text{ is the discrete direct transmission matrix}. \]

Eqs. (A.12) is known as the state space model in discrete time. The discrete state space matrices are related to their continuous-time counterparts by

\[ A_d = e^{A_c \Delta t}, \quad (A.13) \]
\[ B_d = (A_d - I) A_c^{-1} B_c, \quad (A.14) \]
\[ C_d = C_c, \quad (A.15) \]
\[ D_d = D_c. \quad (A.16) \]

**Proof.** First, it is convenient to obtain the discrete form of the solution of state equation (A.10). Hence, in Equation (A.10) the values at time \( x(t + \Delta t) \) can be related to those at time \( t \) by

\[ x(t + \Delta t) = e^{A_c \Delta t} x(t_0) + e^{A_c (t+\Delta t)} \int_{t_0}^{t+\Delta t} e^{-A_c s} B_c u(s) ds. \quad (A.17) \]

Multiplying (A.10) by \( e^{A_c \Delta t} \):

\[ e^{A_c \Delta t} x(t) = e^{A_c (t+\Delta t)} x_0 + e^{A_c (t+\Delta t)} \int_{0}^{t} e^{-A_c s} B_c u(s) ds \Rightarrow \]
\begin{equation}
e^{A_c(t+\Delta t)}x_0 = e^{A_c(t+\Delta t)}x(t) - \int_0^t e^{-A_c s} B_c u(s) ds.
\tag{A.18}
\end{equation}

Substituting Equation (A.18) into Equation (A.17) yields

\begin{equation}
x(t + \Delta t) = e^{A_c \Delta t} x(t) - e^{A_c (t+\Delta t)} \int_0^t e^{-A_c s} B_c u(s) ds + e^{A_c (t+\Delta t)} \int_0^{(t+\Delta t)} e^{-A_c s} B_c u(s) ds
\Rightarrow x(t + \Delta t) = e^{A_c \Delta t} x(t) + e^{A_c (t+\Delta t)} \int_t^{(t+\Delta t)} e^{-A_c s} B_c u(s) ds.
\tag{A.19}
\end{equation}

No approximations have been introduced in Equation (A.19). A frequently used approximation is to consider \( u(s) \) as being constant in the interval \((t, t+\Delta t)\). Considering that \( u(s) \) is approximated by the value of \( u \) at the beginning of the interval (zero order hold):

\begin{equation}
u(s) = u(t_k) \quad t_k \leq s \leq t_k + \Delta t,
\tag{A.20}
\end{equation}

Solving the integral

\begin{equation}
x(t_k + \Delta t) = e^{A_c \Delta t} x(t_k) + e^{A_c (t_k+\Delta t)} \left(-e^{-A_c (t_k+\Delta t)} A_c^{-1} + e^{-A_c t_k} A_c^{-1}\right) B_c u(t_k),
\end{equation}

and grouping terms

\begin{equation}
x(t_k + \Delta t) = e^{A_c \Delta t} x(t_k) + (-I + e^{A_c \Delta t}) A_c^{-1} B_c u(t_k).
\end{equation}

Thus, using (A.11)

\begin{equation}
x_{k+1} = A_d x_k + B_d u_k,
\tag{A.21}
\end{equation}

where

\begin{equation}
A_d = e^{A_c \Delta t},
\tag{A.22}
\end{equation}

\begin{equation}
B_d = (A - I) A_c^{-1} B_c.
\tag{A.23}
\end{equation}

On the other hand, the discretization of observation equation is immediate

\begin{equation}
y(t_k) = C_c x(t_k) + D_c u(t_k) \Rightarrow
\end{equation}

\begin{equation}
y_k = C_d x_k + D_d u_k,
\tag{A.24}
\end{equation}

\begin{equation}
C_d = C_c, \quad D_d = D_c.
\tag{A.25}
\end{equation}

We include now two important properties of the LTI state space model (A.12) that will be used several times in this work. The derivation of these properties for the continuous-time case is similar to the discrete case.

**Property A.2.** Consider the LTI system given by (A.12). Matrices \( A_d, B_d, C_d \) and \( D_d \) and the state vector \( x_k \) are not unique.
Proof. We can transform the state $x_k$ into another state vector $z_k$ as follows:

$$x_k = T z_k,$$

where $T$ is an arbitrary non-singular matrix that is called a state transformation or a similarity transformation. Substituting in (A.12a) and (A.12b) gives:

$$T z_{k+1} = A_d T z_k + B_d u_k$$
$$y_k = C_d T z_k + D_d u_k.$$

Pre-multiplying by $T^{-1}$:

$$z_{k+1} = T^{-1} A_d T z_k + T^{-1} B_d u_k$$
$$y_k = C_d T z_k + D_d u_k.$$

Finally:

$$z_{k+1} = A_t z_k + B_t u_k$$
$$y_k = C_t z_k + D_t u_k,$$

where

$$A_t = T^{-1} A_d T, \quad B_t = T^{-1} B_d$$
$$C_t = C_d T, \quad D_t = D_d.$$

So the state of a system is not unique: there are different state representations that yield the same dynamic relation between $u_k$ and $y_k$, that is, the same input-output behaviour.

Property A.3. Consider a system given by (A.12a) and (A.12b), and the eigenproblem of matrix $A_d$:

$$A_d = V^{-1} D V,$$

where $D$ is the eigenvalues matrix and $V$ is the eigenvectors matrix (see Section C.1). Then, $D$, $C_d V$ and $V^{-1} B_d$ are system invariants, that is, they are not affected by similarity transformations.

Proof. Let be two different state representations of the system, that is

$$r_{k+1} = A_r r_k + B_r r_k$$
$$y_k = C_r r_k + D_r u_k,$$
$$s_{k+1} = A_s s_k + B_s s_k$$
$$y_k = C_s s_k + D_s u_k,$$

with transformation matrix

$$r_k = T s_k.$$

Therefore, taking into account Property A.2

$$A_s = T^{-1} A_r T,$$
$$B_s = T^{-1} B_r,$$
$$C_s = C_r T.$$
\[ D_s = D_r. \]  

The eigenproblem of matrices \( A_r \) and \( A_s \) is given by (A.33):

\[ A_r = V_r D_r V_r^{-1}, \]  
\[ A_s = V_s D_s V_s^{-1}. \]  

Combining equations (A.31), (A.35), and (A.36)

\[ A_s = T^{-1} A_r T = T^{-1} V_r D_r V_r^{-1} T = V_s D_s V_s^{-1}. \]

\( D_r \) and \( D_s \) are two diagonal matrices, and since the diagonalization of a matrix is unique:

\[ D_r = D_s, \]  
\[ V_s = T^{-1} V_r. \]  

Equation (A.37) shows that the eigenvalues of the state matrix are independent of the base chosen to represent the states. On the other hand, substituting (A.38) in (A.33)

\[ C_s = C_r T = C_r V_r V_r^{-1} \Rightarrow C_s V_s = C_r V_s. \]

So the product of the output matrix and the eigenvectors of the state matrix is invariant. Finally

\[ B_s = T^{-1} B_r \Rightarrow V_s^{-1} B_s = V_s^{-1} T^{-1} B_r \Rightarrow V_s^{-1} B_s = (T^{-1} V_r)^{-1} T^{-1} B_r = V_r^{-1} B_r. \]

\[ \square \]

## A.4 Stochastic state space models

### A.4.1 Model for input-output data

Up to now it was assumed that the input signals \( u_k \) and the output signals \( y_k \) are known with precision. However, in practice we can have

- Noise generated by the sensors, \( y_k^* \). Then
  \[ x_{k+1} = A_d x_k + B_d u_k \]  
  \[ y_k = C_d x_k + D_d u_k + y_k^*; \]

- The presence of unmeasured inputs. The system inputs can be expressed as
  \[ u_k = \begin{bmatrix} \bar{u}_k \\ u_k^* \end{bmatrix}, \]

where \( \bar{u}_k \) are the measured inputs and \( u_k^* \) are the unmeasured inputs. Then
  \[ x_{k+1} = A_d x_k + \bar{B}_d \bar{u}_k + B_d^* u_k^* \]  
  \[ y_k = C_d x_k + \bar{D}_d \bar{u}_k + D_d^* u_k^*; \]
The model does not include all the system states. Consider the state space model (A.12) written as

\[
\begin{bmatrix}
\bar{x}_{k+1} \\
\bar{x}_k^\ast + 1
\end{bmatrix} = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix} \begin{bmatrix}
\bar{x}_k \\
\bar{x}_k^\ast
\end{bmatrix} + \begin{bmatrix}
B_1 \\
B_2
\end{bmatrix} u_k
\]

\begin{align*}
y_k &= [C_1 & C_2] \begin{bmatrix}
\bar{x}_k \\
\bar{x}_k^\ast
\end{bmatrix} + D_d u_k.
\end{align*}

If the states \(x_k^\ast\) are not considered, we have

\[
\begin{bmatrix}
\bar{x}_{k+1} \\
\bar{x}_k + 1
\end{bmatrix} = \begin{bmatrix}
A_{11} \\
A_{21}
\end{bmatrix} \begin{bmatrix}
\bar{x}_k \\
\bar{x}_k^\ast
\end{bmatrix} + \begin{bmatrix}
B_1 \\
B_2
\end{bmatrix} u_k 
\]

\begin{align*}
y_k &= C_1 \bar{x}_k + D_d u_k + C_2 x_k^\ast;
\end{align*}

In practice, we have a combination of these situations.

The unknown terms in the above equations (\(y_k^\ast, u_k^\ast\) and \(x_k^\ast\)) are usually characterized as stochastic processes. The state space model used in these cases is the so-called stochastic state space model. We can find two options in the literature (see \([5, 55, 93]\)):

- the state space model

\[
\begin{align*}
x_{k+1} &= A_d x_k + B_d u_k + w_k \\
y_k &= C_d x_k + D_d u_k + v_k,
\end{align*}
\]

- and the model

\[
\begin{align*}
x_{k+1} &= A x_k + B u_k + w_{k+1} \\
y_k &= C x_k + D u_k + v_k,
\end{align*}
\]

where \(w_k \in \mathbb{R}^{n_s}\) is known as the input noise process and \(v_k \in \mathbb{R}^{n_o}\) as the output noise process.

Model (A.39) is generally used when \(\{w_k\} \) and \(\{v_k\}\) are not independent processes, that is, \(\text{Cov}(w_k, v_j) \neq 0\); on the contrary, model (A.40) emerges because it is usual to consider that the error for the states at \(k + 1\) is given by \(w_{k+1}\) and not by \(w_k\). In any case, the differences between both models when \(\text{Cov}(w_k, v_j) = 0\) are very little.

We shall make the following assumptions for \(\{w_k\}\) and \(\{v_k\}\) in this work:

1. White noise processes with

\[
\begin{align*}
\text{E}[w_k] &= 0, & \text{E}[v_k] &= 0; \quad (A.41)
\end{align*}
\]

\[
\text{Cov}(w_k, w_j) = \begin{cases} 0 & \text{if } k \neq j \\ Q & \text{if } k = j \end{cases}, \quad \text{Cov}(v_k, v_j) = \begin{cases} 0 & \text{if } k \neq j \\ R & \text{if } k = j \end{cases}; \quad (A.42)
\]

2. Independent processes

\[
\text{Cov}(w_k, v_j) = 0 \quad \forall \ j, k; \quad (A.43)
\]

(so we use model (A.40) in this work).

3. Gaussian processes

\[
w_k \sim N(0, Q), \quad v_k \sim N(0, R). \quad (A.44)
\]
Finally, we need to specify an initial condition for the difference equation (A.40a). Under normal circumstances, one might expect some prescribed vector at $k = 0$, $x_0$. However, we are not observing the states $x_k$, so it is unlikely that $x_0$ will be available. Then it is usual to adopt a random initial condition for the system [5]. In particular, we shall assume that $x_0$ is a Gaussian random variable of known mean $\bar{x}_0$ and known covariance $P_0$, i.e.,

$$E[x_0] = \bar{x}_0, \quad E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0.$$  

(A.45)

$$x_0 \sim N(\bar{x}_0, P_0).$$  

(A.46)

### A.4.2 Model for output-only data

When the input is totally unknown, like in OMA, the option is to use the model

$$x_{k+1} = Ax_k + w_{k+1}$$  

(A.47a)

$$y_k = Cx_k + v_k.$$  

(A.47b)

In this model, the white noise assumptions for $w_k$ and $v_k$, Eqs. (A.41) - (A.44), and the conditions for the initial state $x_0$, Eqs. (A.45) and (A.46), are also valid:

$$w_t \sim N(0, Q), \quad v_t \sim N(0, R), \quad x_0 \sim N(\bar{x}_0, P_0).$$

(A.48)

The main effect of using this approximation is that matrix $A$ includes now eigenvalues and eigenvectors corresponding to the system and eigenvalues and eigenvectors corresponding to the input (see Section 5.1).

### A.5 Filtering, smoothing and forecasting in the state space model

Three signals define a state space system (see Definition A.2): two signals are measured, $y_k$ and $u_k$. However, the state signal $x_k$ is not observed. The question is whether $x_k$ can be estimated given the data $Y_s = \{y_1, y_2, \ldots, y_s\}$, to time $s$. When $s < k$, the problem is called forecasting or prediction. When $s = k$, the problem is called filtering, and when $s > k$, the problem is called smoothing. In addition to these estimates, we would also want to measure their precision. The solution to these problems is accomplished via the Kalman filter and smoother.

The following notation is used in this document

**Definition A.6.** Given the output data for $s$ time steps $Y_s = \{y_1, y_2, \ldots, y_s\}$, it is defined:

$$x^s_k = E[x_k|Y_s],$$

$$P^s_{k_1, k_2} = E[(x_{k_1} - x^s_{k_1})(x_{k_2} - x^s_{k_2})^T|Y_s],$$

where $E[\square|\square]$ is the conditional expectation operator.

When $k_1 = k_2 = k$ it will be written $P^s_k$:

$$P^s_k = E[(x_k - x^s_k)(x_k - x^s_k)^T|Y_s] = \text{Var}[x_k|Y_s].$$
A.5 Filtering, smoothing and forecasting in the state space model

A.5.1 The Kalman filter

First, we present the Kalman filter, which gives the filtering equations, \( x_k^k = \text{E}[x_k|Y_k] \), and the forecasting equations, \( x_k^{-1} = \text{E}[x_k|Y_{k-1}] \). The advantage of the Kalman filter is that it specifies how to update the filter from \( x_k^{-1} \) to \( x_k^k \) once a new observation \( y_k \) is obtained, without having to reprocess the entire data set \( y_1, \ldots, y_k \).

**Property A.4** (The Kalman Filter). For the state space model specified in (A.40) with initial conditions \( \bar{x}_0 = x_0^0 \) and \( P_0 = P_0^0 \), for \( k = 1, 2, \ldots, N \),

\[
x_k^{k-1} = A_d x_{k-1}^{k-1} + B_d u_{k-1},
\]

\[
P_k^{k-1} = A_d P_{k-1}^{k-1} A_d^T + Q,
\]

with

\[
x_k^k = x_k^{k-1} + K_k \epsilon_k,
\]

\[
P_k^k = [I - K_k C] P_k^{k-1},
\]

where

\[
K_k = P_k^{k-1} C^T \Sigma_k^{-1},
\]

\[
\epsilon_k = y_k - \text{E}[y_k|Y_{k-1}] = y_k - C_d x_k^{k-1} - D_d u_k,
\]

\[
\Sigma_k = \text{Var}[\epsilon_k] = \text{Var}[C_d (x_k - x_k^{k-1}) + v_k] = C_d P_k^{k-1} C_d^T + R.
\]

\( K_k \) is called the Kalman gain and \( \epsilon_k \) are the innovations (prediction errors). Prediction for \( k > N \) is accomplished via (A.49) and (A.50) with initial conditions \( x_N^N \) and \( P_N^N \).

A.5.2 The Kalman smoother

Now we consider the problem of obtaining estimators for \( x_k \) based on the entire data sample \( y_1, \ldots, y_N \), where \( k \leq N \), namely \( x_N^N \). These estimators are called smoothers because a time plot of the sequence \( \{x_k^N; k = 1, \ldots, N\} \) is typically smoother than the forecast \( \{x_k^{-1}; k = 1, \ldots, N\} \) or the filters \( \{x_k^k; k = 1, \ldots, N\} \). As is obvious from the above remarks, smoothing implies that each estimated value is a function of the present, future and past, whereas the filtered estimator depends on the present and past. The forecast depends only on the past.

**Property A.5** (The Kalman Smoother). For the state space model specified in (A.40) with initial conditions \( x_N^N \) and \( P_N^N \) obtained via Property A.4 for \( k = N, N-1, \ldots, 1 \),

\[
x_{k-1}^N = x_{k-1}^{k-1} + J_{k-1} \left( x_k^N - x_k^{k-1} \right),
\]

\[
P_{k-1}^N = P_{k-1}^{k-1} + J_{k-1} \left( P_k^N - P_k^{k-1} \right) J_{k-1}^T,
\]

where

\[
J_{k-1} = P_{k-1}^{k-1} A_d^T \left[ P_k^{k-1} \right]^{-1}.
\]

In this work we will need a set of recursions for obtaining the covariance smoother \( P_{k,k-1}^N \) as defined in (A.6). We give the necessary recursions in the following property.
Property A.6 (The Lag-One Covariance Smoother). For the state space model specified in (A.40), with $K_k$, $J_k$ ($k = 1, 2, \ldots, N$), and $P_N^N$ obtained from Properties A.4 and A.5, and with initial condition

$$P_{N,N-1}^N = (I - K_N^C) A P_{N-1}^{N-1}$$

(A.59)

for $k = N, N - 1, \ldots, 2$

$$P_{k-1,k-2}^N = P_{k-1}^{k-1} J_{k-2}^T + J_{k-1}^T \left( P_{k,k-1}^N - A P_{k-1}^{k-1} \right) J_{k-2}^T.$$  (A.60)

The proof of Properties A.4, A.5 and A.6 can be found in many texts, for example, [5, 55, 93].
Appendix B

Modal analysis

B.1 Introduction

The essential idea of modal analysis is to describe the vibrations of a structural system with simple components, the so-called vibration modes or modal parameters. In this sense, we could trace an analogy with the Fourier analysis, where complex functions are described by the sum of harmonic functions.

The vibration modes naturally emerge when solving the equation of vibration of a structural system and, mathematically, it is reduced to an eigenvalue problem. Using the equations, we can prove that the dynamic response of an structural system is completely characterized by its vibration modes.

This chapter will focus on deriving the vibration modes, their properties and their relation with the state space model. These relationships are the base of the connection between modal analysis and system identification in structural mechanics.

B.2 Equation of vibration of a structural system

Several mathematical models can be constructed to analyse the vibrations of a structural system. Sometimes, the system can be directly modelled as a set of localized masses, springs and elements dissipating energy or dampers; more general, the Finite Element Method (FEM) consists on approximate the behaviour of continuous systems by a finite number of displacements and rotations, the Degrees of Freedom (DOF). In any case, a discrete model is generally built to analyse the vibrations of a continuous system.

Let us consider a discrete model with $n_d$ DOF for a structural system. The mass matrix $\mathcal{M} \in \mathbb{R}^{n_d \times n_d}$ and the stiffness matrix $\mathcal{K} \in \mathbb{R}^{n_d \times n_d}$ can be computed using the FEM [53, 85], what results in symmetric and positive-definite matrices. On the contrary, it is difficult to formulate explicit mathematical expressions for the damping forces in a structure (see [1] for details and some examples). The main model used to study damping in structural systems is the viscous damping model, where the damping is proportional to the velocity. Viscous damping is modelled by mean of the viscous damping matrix $\mathcal{C} \in \mathbb{R}^{n_d \times n_d}$. In principle, the only property required for this matrix is to be symmetric.

Applying the Second Newton Law to the discrete model, the equation of motion is obtained [28]:

$$\dot{\mathcal{M}}q(t) + \mathcal{C}q(t) + \mathcal{K}q(t) = F(t), \quad (B.1)$$
where \( q(t) \in \mathbb{R}^{n_d \times 1} \) is the vector containing the solution for the considered DOFs (the dots mean derivative with respect to time) and \( F(t) \in \mathbb{R}^{n_d \times 1} \) is the vector of the external forces applied to the system.

**B.3 Modes of vibration of a structural system**

**B.3.1 Definition**

Eq. (B.1) is a linear differential equation, so its solution can be computed as

\[
q(t) = q_h(t) + q_p(t),
\]

(B.2)

where \( q_p(t) \) is a particular or complementary solution and \( q_h(t) \) is the solution of the homogeneous version of Eq. (B.1), that is,

\[
[M \ddot{q}_h(t) + C \dot{q}_h(t) + K q_h(t) = 0],
\]

(B.3)

\[
q(t_0), \dot{q}(t_0),
\]

(B.4)

where \( q(t_0), \dot{q}(t_0) \) are the initial conditions. According to the theory of differential equations, the solutions of Eq. (B.3) have the form

\[
q_h(t) = \psi e^{\lambda t}.
\]

(B.5)

In this equation, \( \lambda \) is a scalar (the eigenvalue), and \( \psi \) is a vector with dimension \( n_d \) (the eigenvector). Deriving Eq. (B.5) two times and substituting in Eq. (B.3) it is obtained

\[
(M \lambda^2 + C \lambda + K) \psi = 0.
\]

(B.6)

This equation is known as the quadratic eigenvalue problem [95]. Since matrix \( M \) is nonsingular, the solution is composed by \( 2n_d \) eigenvalues \( \lambda_k \) and \( 2n_d \) eigenvectors \( \psi_k \) (see [95]). For simplicity, we are going to assume:

- The multiplicity of the eigenvalues is one (all the eigenvalues are distinct);
- The eigenvalues and eigenvectors are complex valued. Indeed, since \( M, K \) and \( C \) are real and symmetric, the eigenvalues and eigenvectors are real or come in pairs \( (\lambda, \lambda^*) \), where \( \square^* \) means complex conjugate ([95]). Real eigenvalues stand for rigid-body modes, and complex pairs of eigenvalues represent vibrations modes (see [41] and section B.3.2).

If we define the eigenvalues matrix \( \Lambda \in \mathbb{C}^{n_d} \) as a diagonal matrix with eigenvalues \( \lambda_j, j = 1, \ldots, n_d \) in the diagonal and zeros elsewhere, and the eigenvectors matrix \( \Psi \in \mathbb{C}^{n_d \times n_d} \) formed by the corresponding eigenvectors \( \psi_j, j = 1, \ldots, n_d \) as columns, Eq. (B.6) is then equivalent to

\[
(M \Lambda^2 + C \Lambda + K) \Psi = 0,
\]

(B.7)

\[
(M(\Lambda^*)^2 + C \Lambda^* + K) \Psi^* = 0.
\]

(B.8)

The modes of vibration are defined in terms of this eigen-problem in the following definition.
Definition B.1 (Modes of vibration). Consider a structural system with mass matrix \( M \in \mathbb{R}^{n_d \times n_d} \), stiffness matrix \( K \in \mathbb{R}^{n_d \times n_d} \) and viscous damping matrix \( C \in \mathbb{R}^{n_d \times n_d} \). The modes of vibration of the system are the non-trivial solutions of Eq. (B.6)

\[
(M \lambda^2 + C \lambda + K)\psi = 0.
\]

That is, the modes of vibration are given by the matrices \( \{ \Lambda, \Lambda^*, \Psi, \Psi^* \} \).

An efficient way to solve Eq. (B.6) and therefore, to find the modes of vibration of a structural system, is to transform it to the generalized eigenvalue problem of matrix

\[
A = \begin{bmatrix}
0_{n_d} & I_{n_d} \\
-\mathcal{M}^{-1}K & -\mathcal{M}^{-1}C
\end{bmatrix},
\]

The following property gives the details.

Property B.1. Consider a structural system with mass matrix \( M \in \mathbb{R}^{n_d \times n_d} \), stiffness matrix \( K \in \mathbb{R}^{n_d \times n_d} \) and viscous damping matrix \( C \in \mathbb{R}^{n_d \times n_d} \). The eigen-decomposition of matrix

\[
A = \begin{bmatrix}
0_{n_d} & I_{n_d} \\
-\mathcal{M}^{-1}K & -\mathcal{M}^{-1}C
\end{bmatrix},
\]

is given by \(^1\)

\[
A = \mathcal{V} \mathcal{D} \mathcal{V}^{-1}, \quad \mathcal{D} = \begin{bmatrix}
\Lambda & 0 \\
0 & \Lambda^*
\end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix}
\Psi & \Psi^* \\
\Psi \Lambda & \Psi^* \Lambda^*
\end{bmatrix},
\]

where \( \Lambda \) and \( \Psi \) are the the eigenvalues and eigenvectors matrices of the eigenproblem given by Eq. (B.6).

Proof. From Eq. (B.7)

\[-\mathcal{M}^{-1}K \Psi - \mathcal{M}^{-1}C \Psi \Lambda = \Psi \Lambda^2,\]

and taking into account the trivial equation \( \Psi \Lambda = \Psi \Lambda \) and the matrix conjugate properties, it is obtained

\[
\begin{bmatrix}
0_{n_d} & I_{n_d} \\
-\mathcal{M}^{-1}K & -\mathcal{M}^{-1}C
\end{bmatrix}\begin{bmatrix}
\Psi & \Psi^* \\
\Psi \Lambda & \Psi^* \Lambda^*
\end{bmatrix} = \begin{bmatrix}
\Psi & \Psi^* \\
\Psi \Lambda & \Psi^* \Lambda^*
\end{bmatrix}\begin{bmatrix}
\Lambda & 0 \\
0 & \Lambda^*
\end{bmatrix}
\]

Note that because \( \Lambda \) is a diagonal matrix, this equation constitutes the diagonalization problem of matrix

\[
\begin{bmatrix}
0_{n_d} & I_{n_d} \\
-\mathcal{M}^{-1}K & -\mathcal{M}^{-1}C
\end{bmatrix}.
\]

We have defined the modes of vibration (Definition [B.1]), and we have shown how to compute them (Property [B.1]). The following sections state the importance of the modes of vibration.

\(^1\)Because of \( \mathcal{M}, \mathcal{C} \) and \( \mathcal{K} \) are real and symmetric, matrix \( A \) given by Eq. (B.9) can be always factored like \( A = \mathcal{V} \mathcal{D} \mathcal{V}^{-1} \), where \( \mathcal{D} \) is a diagonal matrix and \( \mathcal{V} \) is a non-singular matrix (see [95]).
B.3.2 Physical interpretation

The solution of the homogeneous differential equation (B.3)

\[ M \ddot{q}_h(t) + C \dot{q}_h(t) + K q_h(t) = 0 \]

is the sum of the \(2n_d\) linear independent solutions given by Eq. (B.5), that is

\[ q_h(t) = \sum_{j=1}^{2n_d} c_j \psi_j e^{\lambda_j t}, \]  

(B.10)

where \(c_j\) are general constants that can be determined with the initial conditions (B.4). Taking into account complex eigenvalues, \(\lambda_j = \alpha_j \pm i\beta_j\), and eigenvectors, then

\[ q_h(t) = \sum_{j=1}^{n_d} e^{\alpha_j(t-t_0)} \left( c_{2j-1} \psi_j e^{i\beta_j(t-t_0)} + c_{2j} \psi_j^* e^{-i\beta_j(t-t_0)} \right) \Rightarrow \]

\[ q_h(t) = \sum_{j=1}^{n_d} e^{\alpha_j(t-t_0)} \left( (c_{2j-1} \psi_j + i c_{2j} \psi_j^*) \cos(\beta_j(t-t_0)) + (c_{2j-1} \psi_j - i c_{2j} \psi_j^*) \sin(\beta_j(t-t_0)) \right). \]  

(B.11)

Since the system vibration is real,

\[ c_{2j-1} \psi_j + i c_{2j} \psi_j^* = \delta_j \in \mathbb{R}^{n_d \times 1}, \quad c_{2j-1} \psi_j - i c_{2j} \psi_j^* = \gamma_j \in \mathbb{R}^{n_d \times 1}. \]

Therefore

\[ q_h(t) = \sum_{j=1}^{n_d} e^{\alpha_j(t-t_0)} \left[ \delta_j \cos(\beta_j(t-t_0)) + \gamma_j \sin(\beta_j(t-t_0)) \right]. \]  

(B.11)

This is a damped harmonic signal: the real part of \(\lambda_j\) (\(\alpha_j\)) is a measure for the damping (\(\alpha_j < 0\) because the amplitude of vibration decays to zero with time), and the imaginary part of \(\lambda_j\) (\(\beta_j\)) is the frequency of the oscillations. The eigenvectors and the constants \(c_j\) give the amplitude of the vibrations.

B.3.3 Importance of the modes

We have seen that the vibration \(q(t)\) of a structural system is given by

\[ M \ddot{q}(t) + C \dot{q}(t) + K q(t) = F(t), \]

\[ q(t_0), \ \dot{q}(t_0), \]

\[ q(t) = q_h(t) + q_p(t). \]

The homogeneous part \(q_h(t)\) can be computed using Eq. (B.10) and Prop. (B.1). However, the particular solution \(q_p(t)\) depends on the force \(F(t)\), and it is not possible to find
analytically a particular solution \( q_p(t) \) for each \( F(t) \). Only a few specific cases such as constant forces and harmonic forces can be solved analytically. We include here another approach to the problem; first, we re-write (B.1) as

\[
\ddot{q}(t) = -M^{-1}Kq(t) - M^{-1}C\dot{q}(t) + M^{-1}F(t),
\]

and taking into account the trivial relationship \( \dot{q}(t) = \dot{q}(t) \), we have

\[
\begin{bmatrix}
\dot{q}(t) \\
\ddot{q}(t)
\end{bmatrix} =
\begin{bmatrix}
0_{n_d} & I_{n_d} \\
-M^{-1}K & -M^{-1}C
\end{bmatrix}
\begin{bmatrix}
q(t) \\
\dot{q}(t)
\end{bmatrix} +
\begin{bmatrix}
0 \\
M^{-1}F(t)
\end{bmatrix},
\]

or, in a more compact form

\[
x(t) = Ax(t) + B(t),
\]

where

\[
x(t) =
\begin{bmatrix}
q(t) \\
\dot{q}(t)
\end{bmatrix} \in \mathbb{R}^{2n_d \times 1}, \quad B =
\begin{bmatrix}
0 \\
M^{-1}F(t)
\end{bmatrix} \in \mathbb{R}^{2n_d \times n_d};
\]

\[
A =
\begin{bmatrix}
0_{n_d} & I_{n_d} \\
-M^{-1}K & -M^{-1}C
\end{bmatrix} \in \mathbb{R}^{2n_d \times 2n_d}.
\]

Equation (B.12) is a system of first order linear differential equations. The coefficient matrix of the system, \( A \), is constant. The general solution is

\[
x(t) = x_h(t) + x_p(t)
\]

where \( x_p(t) \) denotes any particular solution, and \( x_h(t) \) (called the homogeneous or complementary solution) represents the general solution of the associated homogeneous equation, that is

\[
\dot{x}_h(t) = Ax_h(t).
\]

Integrating both sides of Equation (B.14) it is obtained

\[
x_h(t) = e^{At}c,
\]

where \( c \in \mathbb{C}^{2n_d \times 1} \) is an arbitrary vector of constants.

The particular solution can be computed using several methods, for example, the method of variation of constants: we define a particular solution equal to the homogeneous solution, but the vector of constants \( c \) now is a time dependent vector

\[
x_p(t) = e^{At}c(t).
\]

Deriving with respect to time, and substituting into Eq. (B.12)

\[
\dot{c}(t) = e^{-At}B(t) \Rightarrow c(t) = \int_{t_0}^{t} e^{-As}B(s)ds \Rightarrow
\]

\[
x_p(t) = e^{At} \int_{t_0}^{t} e^{-As}B(s)ds.
\]

So the general solution of Eq. (B.12) is

\[
x(t) = e^{At}c + \int_{t_0}^{t} e^{A(t-s)}B(s)ds.
\]
Finally, we impose the initial conditions \( x(t_0) = [q(t_0) \quad \dot{q}(t_0)]^T \), and obtain

\[
x(t) = e^{At-t_0}x(t_0) + \int_{t_0}^{t} e^{A(t-s)}B(s)ds.
\]  

(B.19)

We see that the vibration of a structural system depends on the term \( e^{At} \). But according to (C.16)

\[
e^{At} = D e^{Vt} V^{-1},
\]  

(B.20)

where \( D \) and \( V \) are the eigenvalues and eigenvectors matrices of \( A \) respectively. Taking into account this property

\[
x(t) = [q(t) \quad \dot{q}(t)] = V e^{D(t-t_0)}V^{-1}x(t_0) + \int_{t_0}^{t} V e^{D(t-s)}V^{-1}B(s)ds.
\]  

(B.21)

Once \( x(t) \) is known, that is, the displacement and velocity are known, the acceleration is easily computed. From Eq. (B.1)

\[
\ddot{q}(t) = -M^{-1}Kq(t) - M^{-1}C\dot{q}(t) + M^{-1}F(t) \Rightarrow
\]  

(B.22)

\[
\ddot{q}(t) = [-M^{-1}K \quad -M^{-1}C] \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix} + M^{-1}F(t),
\]  

(B.23)

\[
\ddot{q}(t) = \begin{bmatrix} 0_{nd} & I_{nd} \end{bmatrix} A x(t) + M^{-1}F(t),
\]  

(B.24)

\[
\ddot{q}(t) = \begin{bmatrix} 0_{nd} & I_{nd} \end{bmatrix} (VDV^{-1}) x(t) + M^{-1}F(t).
\]  

(B.25)

In conclusion, the vibrations of a structural system due to a given force \( F(t) \), that is, \( \{q(t), \dot{q}(t), \ddot{q}(t)\} \), are determined by the eigenvalues and eigenvectors of matrix \( A \) (the modes of vibrations) and by matrix \( M \).

### B.4 Special cases of damping

#### B.4.1 Modes of vibration of non-damped systems

All real structures dissipate energy when they vibrate, so taking damping equal to zero is not a realistic case. However, some important results are derived taking into account this hypothesis.

When the damping is zero, Eq. (B.3) turns into

\[
\mathcal{M}\ddot{q}(t) + \mathcal{K}q(t) = 0,
\]  

(B.26)

and the modes of vibration are defined by the non-trivial solutions of

\[
(\mathcal{M}\lambda^2 + \mathcal{K})\psi = 0.
\]  

(B.27)

Considering the following variables

\[
\lambda = -i\omega, \quad \psi = \phi,
\]  

(B.28)

and substituting, one obtains

\[
(\mathcal{M}^{-1}\mathcal{K} - \omega^2 I)\phi = 0.
\]  

(B.29)
B.4 Special cases of damping

This is the eigen-problem of matrix $\mathbf{M}^{-1}\mathbf{K}$: $\omega_j^2$ are the eigenvalues and $\phi_j$ the eigenvectors ($j = 1, 2, \ldots, n_d$). All eigenvalue problem can be written in a matrix form, therefore

$$\mathbf{M}^{-1}\mathbf{K} = \Phi \Omega^2 \Phi^{-1}, \quad \text{(B.30)}$$

where $\Phi \in \mathbb{R}^{n_d \times n_d}$ contains the eigenvectors as columns, and $\Omega^2 \in \mathbb{R}^{n_d \times n_d}$ is a diagonal matrix formed with the eigenvalues, $[\Omega^2]_{ij} = \omega_j^2 \delta_{ij}$ ($\delta_{ij}$ is the Kronecker delta).

**Property B.2.** Consider a structural system with mass matrix $\mathbf{M} \in \mathbb{R}^{n_d \times n_d}$ and stiffness matrix $\mathbf{K} \in \mathbb{R}^{n_d \times n_d}$, and let $\Omega^2$ be the eigenvalue matrix and $\Phi$ the eigenvectors matrix of $\mathbf{M}^{-1}\mathbf{K}$. Then

- $\Omega$ and $\Phi$ are real matrices,
  $$\Omega, \Phi \in \mathbb{R}^{n_d \times n_d}; \quad \text{(B.31)}$$
- $\Phi^T \mathbf{M} \Phi$ is a diagonal matrix
  $$\Phi^T \mathbf{M} \Phi = \mathbf{M}_m, \quad [\mathbf{M}_m]_{ij} = m_j \delta_{ij}; \quad \text{(B.32)}$$
  $\mathbf{M}_m$ is called the modal mass matrix.
- $\Phi^T \mathbf{K} \Phi$ is a diagonal matrix
  $$\Phi^T \mathbf{K} \Phi = \mathbf{K}_m, \quad [\mathbf{K}_m]_{ij} = k_j \delta_{ij}; \quad \text{(B.33)}$$
  $\mathbf{K}_m$ is called the modal stiffness matrix.
- The eigenvalue matrix can be computed as
  $$\Omega^2 = \mathbf{M}_m^{-1} \mathbf{K}_m, \quad [\Omega^2]_{ij} = \omega_j^2 \delta_{ij}; \quad \text{(B.34)}$$
  Taking into account (B.32) and (B.33)
  $$\omega_j^2 = \frac{k_j}{m_j} \text{[rad}^2/\text{s}^2]\text{]; } \quad \text{(B.35)}$$
  $\omega_j$ is called the natural or undamped frequency of vibration.

**Proof.** The proof of this important property can be found in many structural dynamics tests, for example in [28].

It is important to remark that $\Omega^2$ and $\Phi$ are real matrices: the square root of the eigenvalues $\omega_j$ are called the natural frequencies of vibration, and the undamped eigenvectors $\phi_j$ are called mode shapes (their graphical representation have a nice visual interpretation as the deformation shape of the structure).

The modes of vibration of an undamped system are usually defined in terms of the above parameters. The following properties summarize the main results.

**Property B.3.** Consider a structural system with mass matrix $\mathbf{M} \in \mathbb{R}^{n_d \times n_d}$ and stiffness matrix $\mathbf{K} \in \mathbb{R}^{n_d \times n_d}$. If the viscous damping matrix $\mathbf{C} \in \mathbb{R}^{n_d \times n_d}$ is equal to zero, the modes of the system, that is, the non-trivial solutions of the equation

$$(\mathbf{M}\lambda^2 + \mathbf{K})\psi = 0, \quad \text{(B.36)}$$

are given by

$$\Lambda = i\Omega, \quad \text{(B.37)}$$
$$\Psi = \Phi. \quad \text{(B.38)}$$
Proof. The proof is straightforward taking into account Eqs. (B.28) and (B.30).

Property B.4. Consider a structural system with mass matrix $M \in \mathbb{R}^{n_d \times n_d}$ and stiffness matrix $K \in \mathbb{R}^{n_d \times n_d}$. The eigen-decomposition of matrix

$$A = \begin{bmatrix} 0 & \mathbf{I} \\ -M^{-1}K & 0 \end{bmatrix},$$

is given by

$$A = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}, \quad \mathbf{D} = \begin{bmatrix} i\Omega & 0 \\ 0 & -i\Omega \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \Phi & \Phi \Omega \\ i\Phi \Omega & -i\Phi \Omega \end{bmatrix},$$

where $\Omega$ and $\Phi$ are the modes of vibration of the undamped system (Prop. B.3).

B.4.2 Modes of vibration of proportionally damped systems

It is difficult to formulate explicit mathematical expressions for the damping forces acting in a structure. Simplified models have been developed which, in many cases, are based more on mathematical convenience than on physical representation. One example is the proportional viscous damping model. Proportional damping has found significant applications in finite element analysis where damping needs to be incorporated in order to carry out meaningful response analysis and predictions.

Definition B.2. Consider a structural system with mass matrix $M \in \mathbb{R}^{n_d \times n_d}$, stiffness matrix $K \in \mathbb{R}^{n_d \times n_d}$ and viscous damping matrix $C \in \mathbb{R}^{n_d \times n_d}$. And let $\Phi$ be the eigenvectors matrix of $M^{-1}K$. Matrix $C$ is called proportional viscous damping if

$$\Phi^T C \Phi = C_m,$$

where $C_m$ is a diagonal matrix

$$(C_m)_{ij} = c_j \delta_{ij}.$$

Although it may seem a very restrictive condition, this hypothesis is valid in most of structures, and even small non-proportional damping can be replaced with proportional damping without causing a significant error ([11], chapter 3).

A general class of matrices giving proportional damping, for nonsingular matrix $M$, is ([28])

$$C = \sum_j \alpha_j M(M^{-1}K)^j = \alpha_0 M + \alpha_1 K + \alpha_2 K M^{-1} K + \cdots,$$

in which as many terms can be included as desired. The well known Rayleigh damping, where $C = \alpha_0 M + \alpha_1 K$, is a particular case of proportional damping commonly used in practice. Proportional damping assumes that the energy loss mechanism is distributed over the structure in the same way as the mass and the stiffness.

Property B.5. Consider a structural system with mass matrix $M \in \mathbb{R}^{n_d \times n_d}$, stiffness matrix $K \in \mathbb{R}^{n_d \times n_d}$ and proportional viscous damping matrix $C \in \mathbb{R}^{n_d \times n_d}$. The modes of the system, that is, the non-trivial solutions of the equation

$$(M \Lambda^2 + C \Lambda + K) \Psi = 0$$

are

$$\Lambda = Z \Omega + i \Omega (I - Z^2)^{1/2} \in \mathbb{C}^{n_d \times n_d},$$
where $\Omega$ and $\Phi$ verify
\[ M^{-1}K = \Phi \Omega^2 \Phi^{-1}, \]
and $Z$ is a diagonal matrix
\[ [Z]_{ij} = \delta_{ij} \zeta_j, \quad \zeta_j = \frac{c_j}{2\sqrt{m_j k_j}}, \]
$\zeta_j$ is called the modal damping ratio.

**Proof.** Since $\phi_j$ is the eigenvector corresponding to the eigenvalue $\lambda_j$
\[ (\lambda_j^2 M + \lambda_j C + K) \phi_j = 0. \]
Pre-multiplying by $\phi_i^T$
\[ \phi_i^T (\lambda_j^2 M + \lambda_j C + K) \phi_j = 0, \]
\[ (\lambda_j^2 \phi_i^T M \phi_j + \lambda_j \phi_i^T C \phi_j + \phi_i^T K \phi_j) = 0. \]
Taking into account the orthogonality properties of $M$ and $K$ (Equations (B.32) and (B.33)), and the definition of proportional damping (B.40)
\[ \lambda_j^2 m_j + \lambda_j c_j + k_j = 0, \quad j = 1, \ldots, n_d, \]
with solution
\[ \lambda_j = -c_j \pm \sqrt{c_j^2 - 4m_j k_j} = -\zeta_j \omega_j \pm i\omega_j \sqrt{1 - \zeta_j^2}. \]

Finally, a similar property to Prop. B.4 is given.

**Property B.6.** Consider a structural system with mass matrix $M \in \mathbb{R}^{n_d \times n_d}$, stiffness matrix $K \in \mathbb{R}^{n_d \times n_d}$ and proportional viscous damping matrix $C \in \mathbb{R}^{n_d \times n_d}$. The eigen-decomposition of matrix
\[ A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix} \]
is given by
\[ A = \mathcal{V} \mathcal{D} \mathcal{V}^{-1}, \quad \mathcal{D} = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda^* \end{bmatrix}, \quad \mathcal{V} = [\Phi \Phi \Phi \Phi^*], \]
where $\Lambda$ and $\Phi$ are the modes of the system with proportional damping (Eqs. (B.44) and (B.45)).

### B.5 Computing the modes of vibration from vibration data

Given a structural system, the modes of vibration can be computed using Prop. B.1 if we know the mass, stiffness and viscous damping matrices. This section deals with the problem of computing the modes of vibration not from these matrices, but from vibration data measured in the system.

In this work we adopt the parametric approach, that is, first we fit a mathematical (or parametric) model to the measured data and then we compute the modes from the parameters of the model. We use the state space model (see Appendix A for details), so we need the relationship between the modes of vibration and the state space model.
B.5.1 Modes of vibration and the state space model

The equation of vibration of a structural system, Eq. (B.1), can be written as

\[ \ddot{q}(t) + Kq(t) = B_u u(t), \]  

where \( u(t) \) is the \( n_i \)-dimensional input vector, and \( B_u \in \mathbb{R}^{n_d \times n_i} \) is a matrix that verifies

\[ F(t) = B_u u(t), \]  

that is, \( B_u \) indicates where \( u(t) \) is applied. Then

\[ \ddot{q}(t) = -M^{-1}Kq(t) - M^{-1}C\dot{q}(t) + M^{-1}B_u u(t), \]  

and taking into account the trivial equation \( \dot{q}(t) = \dot{q}(t) \), it is obtained

\[
\begin{bmatrix}
\dot{q}(t) \\
\dot{\dot{q}}(t)
\end{bmatrix} = 
\begin{bmatrix}
0 & I \\
-M^{-1}K & -M^{-1}C
\end{bmatrix}
\begin{bmatrix}
q(t) \\
\dot{q}(t)
\end{bmatrix} + 
\begin{bmatrix}
0 \\
M^{-1}B_u
\end{bmatrix} u(t).
\]  

Note that this equation corresponds to the state equation of the state space model (A.4):

\[ \dot{x}(t) = A_c x(t) + B_c u(t), \]  

where

\[ x(t) = \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix} \in \mathbb{R}^{n_s \times 1}, \]  

\[ A_c = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix} \in \mathbb{R}^{n_s \times n_s}, \]  

\[ B_c = \begin{bmatrix} 0 \\ M^{-1}B_u \end{bmatrix} \in \mathbb{R}^{n_s \times n_i}. \]  

On the other hand, we can measure or observe system displacements, velocities or accelerations. However, due to sensor constraints, it is not usual to measure them at all DOFs given by \( q(t) \). Because this, we define the observation or measurement vector \( y(t) \) as

\[ y(t) = \begin{bmatrix} C_d q(t) \\ C_v \dot{q}(t) \\ C_a \ddot{q}(t) \end{bmatrix} \in \mathbb{R}^{n_o \times 1}, \]  

where \( C_d \in \mathbb{R}^{n_o \times n_d}, \) \( C_v \in \mathbb{R}^{n_o \times n_d} \) and \( C_a \in \mathbb{R}^{n_o \times n_d} \) are matrices composed by zeros and ones indicating what DOFs are measured. Therefore, \( y(t) \) can, in general, include position, velocity, and/or acceleration data obtained at selected coordinates, depending on the kind of sensor we use and the position of the sensors.

Taking into account again Equation (B.1) we obtain

\[ y(t) = \begin{bmatrix} C_d & 0 \\ 0 & C_v \\ -C_a M^{-1}K & -C_a M^{-1}C \end{bmatrix}
\begin{bmatrix}
q(t) \\
\dot{q}(t)
\end{bmatrix} + 
\begin{bmatrix}
0 \\
0 \\
C_a M^{-1}B_u
\end{bmatrix} u(t), \]  

that is, the observation equation of the state space model (A.4):

\[ y(t) = C_c x(t) + D_c u(t), \]  

(B.57)
B.5 Computing the modes of vibration from vibration data

where

\[ C_c = \begin{bmatrix} C_d & 0 \\ 0 & C_v \end{bmatrix} \in \mathbb{R}^{n_o \times n_s}, \quad (B.58) \]

\[ D_c = \begin{bmatrix} 0 \\ 0 \\ C_a M^{-1} B_u \end{bmatrix} \in \mathbb{R}^{n_o \times n_i}. \quad (B.59) \]

The following property summarizes the above equations:

**Property B.7** (State space model for structural systems). Given a structural system with mass matrix \( M \in \mathbb{R}^{n_d \times n_d} \), stiffness matrix \( K \in \mathbb{R}^{n_d \times n_d} \) and viscous damping matrix \( C \in \mathbb{R}^{n_d \times n_d} \), the dynamic behaviour of the system in continuous-time can be represented with the state space equations

\[ \dot{x}(t) = A_c x(t) + B_c u(t) \quad (B.60a) \]
\[ y(t) = C_c x(t) + D_c u(t), \quad (B.60b) \]

where

- \( y(t) \in \mathbb{R}^{n_o} \) is the output vector
  \[ y(t) = \begin{bmatrix} C_d q(t) \\ C_v \dot{q}(t) \\ C_a \ddot{q}(t) \end{bmatrix}; \quad (B.61) \]
  \( C_d \in \mathbb{R}^{n_o_d \times n_d} \), \( C_v \in \mathbb{R}^{n_o_v \times n_d} \) and \( C_a \in \mathbb{R}^{n_o_a \times n_d} \) are the output matrices;
- \( u(t) \in \mathbb{R}^{n_i} \) is the input vector, which verifies
  \[ F(t) = B_u u(t); \quad (B.62) \]
  \( B_u \in \mathbb{R}^{n_d \times n_i} \) is the excitation influence matrix;
- \( x(t) \in \mathbb{R}^{n_s} \) is the state vector
  \[ x(t) = \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix}; \quad (B.63) \]
- \( A_c \in \mathbb{R}^{n_s \times n_s} \) is the transition state matrix
  \[ A_c = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix}; \quad (B.64) \]
- \( B_c \in \mathbb{R}^{n_s \times n_i} \) is the input matrix;
  \[ B_c = \begin{bmatrix} 0 \\ M^{-1}B_u \end{bmatrix}; \quad (B.65) \]
- \( C_c \in \mathbb{R}^{n_o \times n_s} \) is the output matrix
  \[ C_c = \begin{bmatrix} C_d & 0 \\ 0 & C_v \\ -C_a M^{-1} K & -C_a M^{-1} C \end{bmatrix}; \quad (B.66) \]
• \( D_c \in \mathbb{R}^{n_o \times n_i} \) is the direct transmission matrix

\[
D_c = \begin{bmatrix}
0 \\
0 \\
C_a \mathcal{M}^{-1} B_u 
\end{bmatrix}.
\]  
(B.67)

Now we show the connection between the state space model and the modes of vibration by computing the system invariants given by Prop. A.3. In next Section we will see that these formulas are very useful for computing the modes of vibration from measured data.

Property B.8. Consider the continuous-time state space model \((B.60)\) to \((B.67)\), and the eigenvalue problem of matrix \(A_c\):

\[
A_c V_c = V_c D_c.
\]  
(B.68)

Then it holds

\[
D_c = \begin{bmatrix}
\Lambda & 0 \\
0 & \Lambda^\ast
\end{bmatrix},
\]  
(B.69)

\[
V_c = \begin{bmatrix}
\Psi \\
\Psi \Lambda \\
\Psi^\ast \Lambda \\
\Psi^\ast (\Lambda^\ast)^2
\end{bmatrix},
\]  
(B.70)

\[
C_c V_c = \begin{bmatrix}
C_d & 0 & 0 \\
0 & C_v & 0 \\
0 & 0 & C_a
\end{bmatrix} \begin{bmatrix}
\Psi & \Psi^\ast \\
\Psi \Lambda & \Psi^\ast \Lambda \\
\Psi \Lambda^2 & \Psi^\ast (\Lambda^\ast)^2
\end{bmatrix},
\]  
(B.71)

\[
V_c^{-1} B_c = \left[\Psi^{-1} \Psi^\ast \right] \left( \Psi^\ast \Lambda^\ast - \Psi \Lambda \Psi^{-1} \Psi^\ast \right)^{-1} \mathcal{M}^{-1} B_u,
\]  
(B.72)

where \(\mathcal{M}\) is the mass matrix, and \(\Lambda\) and \(\Psi\) are the modes of vibration.

Proof. The proof of Equations (B.69) and (B.70) are derived from Property B.1. The proof of Equation (B.71) is

\[
C_c V_c = \begin{bmatrix}
C_d & 0 & 0 \\
0 & C_v & 0 \\
0 & 0 & C_a
\end{bmatrix} \begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix} \begin{bmatrix}
\Psi & \Psi^\ast \\
\Psi \Lambda & \Psi^\ast \Lambda \\
\Psi \Lambda^2 & \Psi^\ast (\Lambda^\ast)^2
\end{bmatrix}.
\]

Taking into account Property B.1

\[
-\mathcal{M}^{-1} K \Psi - \mathcal{M}^{-1} C \Psi \Lambda = \Psi \Lambda^2, \\
-\mathcal{M}^{-1} K \Psi^\ast - \mathcal{M}^{-1} C \Psi^\ast \Lambda = \Psi^\ast (\Lambda^\ast)^2
\]

\[
\Rightarrow C_c V_c = \begin{bmatrix}
C_d & 0 & 0 \\
0 & C_v & 0 \\
0 & 0 & C_a
\end{bmatrix} \begin{bmatrix}
\Psi & \Psi^\ast \\
\Psi \Lambda & \Psi^\ast \Lambda \\
\Psi \Lambda^2 & \Psi^\ast (\Lambda^\ast)^2
\end{bmatrix}.
\]

For Equation (B.72) we need the inverse of \(V_c\)

\[
V_c^{-1} = \begin{bmatrix}
\Psi^{-1} + \Psi^{-1} \Psi \Lambda \Psi^{-1} & -\Psi^{-1} \Psi \Lambda \\
-\Psi^{-1} \Psi \Lambda \Psi^{-1} & \Psi^{-1} \Psi \Lambda \\
\end{bmatrix},
\]
where
\[ O^{-1} = \Psi^* \Lambda^* - \Psi \Lambda \Psi^{-1} \Psi^*. \]

Thus
\[
\begin{align*}
V_c^{-1} B_c &= \begin{bmatrix}
\Psi^{-1} + \Psi^{-1} \psi O^{-1} \Psi \Lambda^{-1} & -\Psi^{-1} \psi O^{-1} \\
O^{-1} \Psi \Lambda^{-1} & O^{-1}
\end{bmatrix}
\begin{bmatrix}
0 \\
-M^{-1} B_u
\end{bmatrix} \\
&= \begin{bmatrix}
\Psi^{-1} \psi \\
-I_{n_d}
\end{bmatrix} O^{-1} M^{-1} B_u = \begin{bmatrix}
\Psi^{-1} \psi \\
-I_{n_d}
\end{bmatrix} (\Psi^* \Lambda^* - \Psi \Lambda \Psi^{-1} \Psi^*)^{-1} M^{-1} B_u.
\end{align*}
\]

Now we give the counterpart of Prop. B.8 for the discrete-time state space model.

**Property B.9.** Consider the discrete-time version of the model (B.60) to (B.67):
\[
x_{t+1} = A_d x_t + B_d u_t \\
y_t = C_d x_t + D_d u_t,
\]
where \( \Delta t \) is the sampling interval. The eigenproblem of matrix \( A_d \) is given by
\[ A_d \mathcal{V}_d = \mathcal{V}_d D_d; \quad \text{(B.74)} \]
Then
\[ D_d = e^{D_d \Delta t}, \quad \text{(B.75)} \]
\[ \mathcal{V}_d = \mathcal{V}_c, \quad \text{(B.76)} \]
\[ C_d \mathcal{V}_d = \begin{bmatrix} C_d & 0 & 0 \\ 0 & C_d & 0 \\ 0 & 0 & C_d \end{bmatrix} \begin{bmatrix} \Psi & \Psi^* \\ \Psi \Lambda & \Psi^* \Lambda^* \\ \Psi \Lambda^2 & \Psi^* (\Lambda^*)^2 \end{bmatrix}, \quad \text{(B.77)} \]
\[ \mathcal{V}_d^{-1} B_d = \begin{bmatrix} (e^{\Lambda \Delta t} - I_{n_d}) \Lambda^{-1} \Psi^{-1} \Psi^* \\ -(e^{\Lambda \Delta t} - I_{n_d}) (\Lambda^{-1})^* \end{bmatrix} (\Psi^* \Lambda^* - \Psi \Lambda \Psi^{-1} \Psi^*)^{-1} M^{-1} B_u. \quad \text{(B.78)} \]

**Proof.** First we are going to prove Eq. (B.75) and Eq. (B.76).
\[ A_d = e^{A_d \Delta t} = e^{\mathcal{V}_d D_d \mathcal{V}_c^{-1} \Delta t} = \mathcal{V}_c e^{D_c \Delta t} \mathcal{V}_c^{-1}, \]
where \( e^{D_c \Delta t} \) is a diagonal matrix. Therefore, the above equation is the eigenvalue decomposition of \( A_d = \mathcal{V}_d D_d \mathcal{V}_c^{-1} \) with \( D_c = e^{D_c \Delta t} \) and \( \mathcal{V}_d = \mathcal{V}_c \).

Eq. (B.77) is due to \( C_d \mathcal{V}_d = C_c \mathcal{V}_c \). For Eq. (B.78)
\[ \mathcal{V}_d^{-1} B_d = \mathcal{V}_d^{-1} (A_d - I) A_c^{-1} B_c = \mathcal{V}_d^{-1} (\mathcal{V}_c D_c \Delta t \mathcal{V}_c^{-1} - I) \mathcal{V}_c D_c \mathcal{V}_c^{-1} B_c; \]

Taking into account Eq. (B.76)
\[ \mathcal{V}_d^{-1} B_d = (e^{D_c \Delta t} - I) D_c \mathcal{V}_c^{-1} B_c = \]
\[ = \begin{bmatrix} e^{\Lambda \Delta t} - I_{n_d} & 0 \\ 0 & e^{\Lambda \Delta t} - I_{n_d} \end{bmatrix} \begin{bmatrix} \Lambda^{-1} & 0 \\ 0 & (\Lambda^{-1})^* \end{bmatrix} \begin{bmatrix} \Psi^{-1} \Psi^* \\ -I_{n_d} \end{bmatrix} (\Psi^* \Lambda^* - \Psi \Lambda \Psi^{-1} \Psi^*)^{-1} M^{-1} B_u = \]
\[ = \begin{bmatrix} (e^{\Lambda \Delta t} - I_{n_d}) \Lambda^{-1} \Psi^{-1} \Psi^* \\ -(e^{\Lambda \Delta t} - I_{n_d}) (\Lambda^{-1})^* \end{bmatrix} (\Psi^* \Lambda^* - \Psi \Lambda \Psi^{-1} \Psi^*)^{-1} M^{-1} B_u. \]

\[ \square \]
B.5.2 Modes of vibration and system identification

Given $N$ measurements for the outputs $y_1, \ldots, y_N$ and for the inputs $u_1, \ldots, u_N$ of a structural system, and the state space model \[(A.40)\]

$$
x_{k+1} = A_d x_k + B_d u_k + w_k \\
y_k = C_d x_k + D_d u_k + v_k,
$$

where $w_k \sim N(0, Q)$, $v_k \sim N(0, R)$ and $x_0 \sim N(\bar{x}_0, P_0)$, the system identification problem consists on estimate matrices $A_d$, $B_d$, $C_d$, $D_d$, $Q$, $R$, $\bar{x}_0$ and $P_0$. The aims of doing system identification can be:

- Control of structural systems: control engineering needs accurate mathematical models, and the state space model is one of the most used. In this case, the estimation procedure finishes with the estimated state space model.
- Modal analysis: system identification and modal analysis are interrelated because the modal parameters $\Lambda$, $\Psi$ and the mass matrix $M$ can be computed from state space matrices $A_d, B_d, C_d$ according to Section B.5.1. Therefore, system identification is an intermediate step towards modal parameters.

Note that the solution of the identification problem is not unique, because we can obtain the model

$$
x_{k+1} = A_d x_k + B_d u_k + w_k \\
y_k = C_d x_k + D_d u_k + v_k,
$$
or the model

$$
z_{k+1} = (T^{-1} A_d T) z_k + (T^{-1} B_d) u_k + T^{-1} w_k \\
y_k = (C_d T) z_k + D_d u_k + v_k,
$$

where $T \in \mathbb{R}^{n_s \times n_s}$ is a non-singular matrix (see Prop. [A.3]). The input and output are the same in both models, so we can obtain any of them as the result of an identification process. The problem is that we can not find the matrix $T$. This is not important in control engineering, because the input-output relation is the same for any $T$. On the contrary, the solution of the modal problem is unique. To avoid the problem, we can use system invariants (Prop. [B.9] for the case of discrete systems).

Let us assume that $\{\hat{A}, \hat{B}, \hat{C}\}$ are the identified or estimated state space matrices from the system inputs and outputs. We are interested in to compute the estimate of $\hat{\Lambda}$, $\hat{\Psi}$ and $\hat{M}$. First, we perform the eigen-decomposition of matrix $\hat{A}$:

$$
\hat{A} = \hat{\Phi}^{-1} \hat{D} \hat{\Phi};
$$

Then, taking into account Prop. [B.9]

- The modal matrix $\hat{A}$ is computed from the eigenvalues of $\hat{A}$ since

$$
\hat{D} = \begin{bmatrix} \hat{\Lambda} & 0 \\ 0 & \hat{\Lambda}^* \end{bmatrix}.
$$
Note that the eigenvalues of matrix $\hat{A}$ are system invariant, and they are not affected by similarity transformations. It is usual to express the $j$th eigenvalue as

$$\hat{\lambda}_j = \exp\left[-\hat{\zeta}_j\hat{\omega}_j \pm i\hat{\omega}_j \sqrt{1 - \hat{\zeta}_j^2}\right] \Delta t;$$

(B.79)

Therefore

$$\hat{\omega}_j = \left|\frac{\ln(\hat{\lambda}_j)}{\Delta t}\right|, \quad \text{ (B.80)}$$

$$\hat{\zeta}_j = \frac{-\text{Real}\left[\ln(\hat{\lambda}_j)\right]}{\hat{\omega}_j \Delta t}. \quad \text{ (B.81)}$$

If proportional damping is admitted, $\hat{\omega}_j$ is the undamped natural frequency (see Eq. (B.35)), and $\hat{\zeta}_j$ is the damping ratio (Eq. (B.47)).

- The modal matrix $\Psi$ cannot be computed from the eigenvectors of $A$ because they depend on the similarity transformation. It can be computed from the system invariant $\hat{C}\hat{V}$:

$$\hat{C}\hat{V} = \begin{bmatrix} C_d\hat{\Psi} & C_d\hat{\Psi}^* \\ C_v\hat{\Psi}\hat{\Lambda} & C_v\hat{\Psi}\hat{\Lambda}^* \\ C_a\hat{\Psi}\hat{\Lambda}^2 & C_a\hat{\Psi}\hat{\Lambda}^2 \end{bmatrix} \quad \text{(B.82)}$$

The following observations can be made:

- Due to $C_d$, $C_v$, and $C_a$, the vectors $\hat{\Psi}$ can only be computed at those points where a sensor has been placed.
- If we do not mix displacement, velocity and acceleration measurements, we do not need to divide by $\Lambda$ or $\Lambda^2$ in the case of velocities and accelerations. This is due to the fact that if $\psi_j$ is an eigenvector of matrix $A$, then $c\psi_j$, $c \in \mathbb{C}$ is also an eigenvector. In particular, $\hat{\lambda}_j\psi_j$ and $\hat{\lambda}_j^2\psi_j$ are eigenvectors.
- If proportional damping is admitted, $\hat{\Psi} = \Phi$, that is, they are the undamped eigenvectors, so they are real vectors. We can take

$$\hat{\Phi} = \text{Real}[\hat{C}\hat{V}], \quad \hat{\Phi} = \text{Imag}[\hat{C}\hat{V}], \quad \text{or} \quad \hat{\Phi} = \text{Modulus}[\hat{C}\hat{V}]. \quad \text{(B.83)}$$

Any of these equations can be used because

$$\hat{C}\hat{V} = \begin{bmatrix} C_d & 0 & 0 \\ 0 & C_v & 0 \\ 0 & 0 & C_a \end{bmatrix} \begin{bmatrix} \hat{\Phi} & \hat{\Phi} \\ \hat{\Phi}\hat{\Lambda} & \hat{\Phi}\hat{\Lambda}^* \\ \hat{\Phi}\hat{\Lambda}^2 & \hat{\Phi}\hat{\Lambda}^2 \end{bmatrix}.$$  

- Finally, the mass matrix $\hat{M}$ can be computed from the invariant $\hat{V}^{-1}\hat{B}$

$$\hat{V}^{-1}\hat{B} = \begin{bmatrix} (e^{\hat{\Lambda}\Delta t} - I_{nd}) \hat{\Lambda}^{-1}\hat{\Psi}^{-1}\hat{\Psi}^* \\ -(e^{\hat{\Lambda}'\Delta t} - I_{nd}) (\hat{\Lambda}'^{-1})^* \end{bmatrix} \left(\hat{\Psi}^*\hat{\Lambda}^* - \hat{\Psi}\hat{\Lambda}\hat{\Psi}^*\hat{\Lambda}^*\right)^{-1}\hat{M}^{-1}B_u. \quad \text{(B.84)}$$

Again, some comments can be made:

- In operational modal analysis matrix $\hat{B}$ can not be estimated. Therefore, matrix $\hat{M}$ can not be estimated either.
– We can only estimate some the elements of matrix $\hat{M}$ at those points where simultaneously a sensor is placed and a force is applied: due to matrix $B_u$, the columns of $\hat{M}$ can be estimated at the DOFs where a force is applied; due to matrices $C_d$, $C_v$ and $C_a$, we only know the rows of $\hat{Ψ}$ at the DOFs where a sensor is placed.

– If proportional damping is admitted, the modal mass matrix can be estimated

$$\hat{\psi}^{-1}\hat{B} = \begin{bmatrix} \left( e^{\hat{Λ}\Delta t} - I_{n_d} \right) \hat{Λ}^{-1} - \left( e^{\hat{Λ}^*\Delta t} - I_{n_d} \right) \left( \hat{Λ}^{-1} \right)^* \end{bmatrix} \left( \hat{Λ}^* - \hat{Λ} \right)^{-1} \hat{Φ}^{-1} \hat{M}^{-1} B_u,$$

$$\hat{\psi}^{-1}\hat{B}_v = \begin{bmatrix} \left( e^{\hat{Λ}\Delta t} - I_{n_d} \right) \left( \left( \hat{Λ}^* - \hat{Λ} \right) \hat{Λ} \right)^{-1} - \left( e^{\hat{Λ}^*\Delta t} - I_{n_d} \right) \left( \left( \hat{Λ}^* - \hat{Λ} \right) \hat{Λ}^* \right)^{-1} \end{bmatrix} \hat{M}_m^{-1} \hat{Φ}^T B_u,$$

$$\hat{M}_m = \hat{Φ}^T B_u \hat{B}_v^\dagger \hat{ψ} \begin{bmatrix} \left( e^{\hat{Λ}\Delta t} - I_{n_d} \right) \left( \left( \hat{Λ}^* - \hat{Λ} \right) \hat{Λ} \right)^{-1} - \left( e^{\hat{Λ}^*\Delta t} - I_{n_d} \right) \left( \left( \hat{Λ}^* - \hat{Λ} \right) \hat{Λ}^* \right)^{-1} \end{bmatrix}, \quad (B.85)$$

where $\hat{Φ}^\dagger$ is the pseudo-inverse.
Appendix C

Mathematical appendix

C.1 Eigenvalues and eigenvectors

The eigenvalue problem consists on to compute the nontrivial solutions of the equation:

$$Ax = \lambda x, \quad A \in \mathbb{R}^{n \times n}, \ x \in \mathbb{R}^{n \times 1}, \ \lambda \in \mathbb{R}. \quad (C.1)$$

The $n$ values for $\lambda$ satisfying Eq. (C.1) are called the eigenvalues, and the corresponding vectors $x$ are called the eigenvectors (right eigenvectors). The problem is solved taking:

$$(A - \lambda I) x = 0. \quad (C.2)$$

This system of homogeneous equations has non-zero solution if

$$\det (A - \lambda I) = 0. \quad (C.3)$$

From this equation we obtain the eigenvalues and eigenvectors of matrix $A$:

$$[\lambda_1, x_1], [\lambda_2, x_2], \ldots, [\lambda_n, x_n]. \quad (C.4)$$

Consider matrix $D$ (eigenvalue matrix)

$$D \overset{\text{def}}{=} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}, \quad (C.5)$$

and matrix $V$ (eigenvectors matrix)

$$V \overset{\text{def}}{=} \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{nn} \end{bmatrix}, \quad (C.6)$$

whose column $i$ is the eigenvector corresponding to the eigenvalue $\lambda_i$

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{bmatrix}; \quad (C.7)$$
Then it holds

\[ AV = VD, \]  \hspace{1cm} (C.8)
\[ A = VDV^{-1}, \]  \hspace{1cm} (C.9)
\[ D = V^{-1}AV. \]  \hspace{1cm} (C.10)

On the other hand, the generalized eigenvalue problem consists on to compute the non-trivial solutions of:

\[ Ax = \lambda Bx, \quad A, B \in \mathbb{R}^{n \times n}, \quad x \in \mathbb{R}^{n \times 1}, \quad \lambda \in \mathbb{R}. \]  \hspace{1cm} (C.11)

If \( B \) is nonsingular, the problem is reduced to the classic eigenvalue problem by taking:

\[ B^{-1}Ax = \lambda x. \]  \hspace{1cm} (C.12)

Therefore we have:

\[ B^{-1}AV = VD, \]  \hspace{1cm} (C.13)
\[ B^{-1}A = VDV^{-1}, \]  \hspace{1cm} (C.14)
\[ D = V^{-1}B^{-1}AV. \]  \hspace{1cm} (C.15)

C.2 Matrix identities

A detailed set of matrix identities can be found at [42, 49].

C.2.1 Matrix exponential

Property C.1. Consider a square matrix \( A \in \mathbb{R}^{n \times n} \) and the scalar \( t \in \mathbb{R} \). Then

\[ e^{At} = Ve^{Dt}V^{-1} = V \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} V^{-1}, \]  \hspace{1cm} (C.16)

where \( V \) is the eigenvectors matrix of \( A \), and \( D \) is the eigenvalue matrix of \( A \) (the \( \lambda_j \) are the eigenvalues of matrix \( A \)).

Proof. The exponential of a matrix if defined as

\[ e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \]

The eigenvalue problem of matrix \( A \) gives

\[ AV = VD \Rightarrow A = VDV^{-1}; \]
\[ A^2 = VDV^{-1}VDV^{-1} = VD^2V^{-1}; \]
\[ A^3 = A^2A = VD^2V^{-1}VDV^{-1} = VD^3V^{-1}; \]

\[ \vdots \]
\[ A^n = VD^nV^{-1}; \]

Thus

\[ e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = V \left[ \sum_{n=0}^{\infty} \frac{D^n}{n!} \right] V^{-1} = Ve^DV^{-1}. \]
C.2.2 Matrix Trace

\[ tr(A) = tr(A^T); \quad (C.17) \]
\[ tr(A + B) = tr(A) + tr(B); \quad (C.18) \]
\[ tr(AB) = tr(BA); \quad (C.19) \]
\[ a^T A a = tr(Aaa^T); \quad (C.20) \]

C.2.3 Matrix Expectation

Assume \( x \) to be a stochastic vector with \( \mu = \mathbb{E}[x] \), \( \Sigma = \text{Var}[x] \). Assume \( A, B \) are symmetric matrices:

\[ \mathbb{E}[Ax + B] = A\mu + B; \quad (C.21) \]
\[ \mathbb{E}[x^T A x] = tr(A\Sigma) + \mu^T A \mu; \quad (C.22) \]

C.2.4 Matrix Derivatives

Consider \( A, B, C, X \) matrices:

\[ \frac{\partial \log (\det (X))}{\partial X} = (X^{-1})^T; \quad (C.23) \]
\[ \frac{\partial tr(AXB)}{\partial X} = A^T B^T; \quad (C.24) \]
\[ \frac{\partial tr(AX^T B)}{\partial X} = BA; \quad (C.25) \]
\[ \frac{\partial tr(AX^{-1} B)}{\partial X} = (X^{-1} BAX^{-1})^T; \quad (C.26) \]
\[ \frac{\partial tr(AXBX^T C)}{\partial X} = A^T C^T XB^T + CAXB; \quad (C.27) \]

C.3 Jensen’s inequality

**Theorem C.1.** Let \((\Omega, \mathcal{A}, \mu)\) be a measure space, such that \( \mu(\Omega) = 1 \). If \( f : I \rightarrow \mathbb{R} \) is \( \mu \)-integrable, and \( \varphi \) is a convex function on the real axis, then:

\[ \int_{\Omega} \varphi \circ f \, d\mu \geq \varphi \left( \int_{\Omega} f \, d\mu \right). \quad (C.28) \]

**Proof.** We begin by letting \( c = \int_{\Omega} f \, d\mu \). If \( c \) were an endpoint of \( I \), it would imply that Jensen’s inequality is satisfied by being an equality. The main case is that in which \( c \) is an interior point of \( I \), and then we let \( m \) denote the slope of the tangent line at \( t = c \) for the convex function \( \varphi \), with equation \( y = m(t - c) + \varphi(c) \). Thus, because \( \varphi \) is a convex function

\[ \varphi(t) \geq m(t - c) + \varphi(c), \quad \forall t \in I. \]
It follows from monotonicity of the integral that
\[
\int_{\Omega} \phi \circ f \, d\mu \geq \int_{\Omega} (m(f(x) - c) + \phi(c)) \, d\mu(x) = m \int_{\Omega} (f(x) - c) \, d\mu + \int_{\Omega} \phi(c) \, d\mu = \phi(c) = \phi \left( \int_{\Omega} f \, d\mu \right).
\]

**Corollary C.1.** The same result can be equivalently stated in probability theory. Let \((\Omega, \mathcal{A}, P)\) be a probability space, \(X\) an integrable real-valued random variable and \(\phi\) a convex function. Then:

\[
\mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[X]).
\]

**(C.29)**

**Proof.** The proof follows from (C.28) and \(\mathbb{E}[X] = \int_{\Omega} X \, dP\).

**Theorem C.2.** Kullback-Leibler divergence.

\[
\int_{-\infty}^{+\infty} \left[ \ln \left( \frac{f(x)}{g(x)} \right) \right] f(x) \, dx \geq 0.
\]

**(C.30)**

**Proof.** Consider:

\[
y \overset{\text{def}}{=} \frac{g(x)}{f(x)},
\]

\(\varphi(y) = -\ln(y)\) is a convex function.

Then, using (C.28)

\[
\int_{-\infty}^{+\infty} \left[ \ln \left( \frac{f(x)}{g(x)} \right) \right] f(x) \, dx = \int_{-\infty}^{+\infty} \left[ -\ln \left( \frac{g(x)}{f(x)} \right) \right] f(x) \, dx \overset{(C.28)}{=} -\ln \left( \int_{-\infty}^{+\infty} \frac{g(x)}{f(x)} f(x) \, dx \right) = -\ln \left( \int_{-\infty}^{+\infty} g(x) \, dx \right) = -\ln 1 = 0.
\]

**Corollary C.2.** Let \(f(x|\theta_1), f(x|\theta_2)\) be. Then

\[
\mathbb{E}[\ln (f(x|\theta_1))|\theta_1] \geq \mathbb{E}[\ln (f(x|\theta_2))|\theta_1].
\]

**(C.31)**

**Proof.** From (C.30)

\[
\int_{-\infty}^{+\infty} \left[ \ln \left( \frac{f(x|\theta_1)}{f(x|\theta_2)} \right) \right] f(x|\theta_1) \, dx \geq 0 \Rightarrow
\int_{-\infty}^{+\infty} \ln (f(x|\theta_1)) f(x|\theta_1) \, dx \geq \int_{-\infty}^{+\infty} \ln (f(x|\theta_2)) f(x|\theta_1) \, dx \Rightarrow
\mathbb{E}[\ln (f(x|\theta_1))|\theta_1] \geq \mathbb{E}[\ln (f(x|\theta_2))|\theta_1].
\]

Bibliography


