Mechanical Systems and Signal Processing 100 (2018) 289–310

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Automated modal parameter estimation using correlation analysis and bootstrap sampling

Vahid Yaghoubi, Majid K. Vakilzadeh *, Thomas J.S. Abrahamsson

Department of Applied Mechanics, Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 3 August 2016
Received in revised form 3 July 2017
Accepted 6 July 2017

Keywords:
Subspace-based identification
Modal parameter estimation
Frequency responses
Modal Observability Correlation
QR- and singular value decomposition
Bootstrapping
Correlation-based clustering
Fuzzy c-means clustering

ABSTRACT

The estimation of modal parameters from a set of noisy measured data is a highly judgmental task, with user expertise playing a significant role in distinguishing between estimated physical and noise modes of a test-piece. Various methods have been developed to automate this procedure. The common approach is to identify models with different orders and cluster similar modes together. However, most proposed methods based on this approach suffer from high-dimensional optimization problems in either the estimation or clustering step. To overcome this problem, this study presents an algorithm for autonomous modal parameter estimation in which the only required optimization is performed in a three-dimensional space. To this end, a subspace-based identification method is employed for the estimation and a non-iterative correlation-based method is used for the clustering. This clustering is at the heart of the paper. The keys to success are correlation metrics that are able to treat the problems of spatial eigenvector aliasing and nonunique eigenvectors of coalescent modes simultaneously. The algorithm commences by the identification of an excessively high-order model from frequency response function test data. The high number of modes of this model provides bases for two subspaces: one for likely physical modes of the tested system and one for its complement dubbed the subspace of noise modes. By employing the bootstrap resampling technique, several subsets are generated from the same basic dataset and for each of them a model is identified to form a set of models. Then, by correlation analysis with the two aforementioned subspaces, highly correlated modes of these models which appear repeatedly are clustered together and the noise modes are collected in a so-called *Trashbox* cluster. Stray noise modes attracted to the mode clusters are trimmed away in a second step by correlation analysis. The final step of the algorithm is a fuzzy c-means clustering procedure applied to a three-dimensional feature space to assign a degree of physicalness to each cluster. The proposed algorithm is applied to two case studies: one with synthetic data and one with real test data obtained from a hammer impact test. The results indicate that the algorithm successfully clusters similar modes and gives a reasonable quantification of the extent to which each cluster is physical.

* Corresponding author.
E-mail address: khorsand@chalmers.se (M.K. Vakilzadeh).

http://dx.doi.org/10.1016/j.ymssp.2017.07.004

1. Introduction

1.1. Pertinent literature

Over the last decades, much effort has been put to develop efficient algorithms for identification of the modal parameters using time or frequency domain data [1,2]. A central problem in most of these algorithms is to determine the true model order to capture the physical modes of the test-piece. However, this model order determination often demands considerable interaction from an experienced user. This hinders the use of developed modal analysis techniques for the applications which require a periodic estimation of the modal parameters like continuous health monitoring of structures.

In the framework of system identification, there exists an extensive literature for order estimation of linear dynamical models. The Akaike Information Criterion (AIC) [3] for maximum likelihood estimator and the Singular Value Criterion (SVC) [4] for subspace-based methods are two such examples of model order selection criteria. They share the idea of comparing the significance of the inclusion of yet another mode for increasing the prediction capability of the model with a penalty cost of including it. Such cost is somewhat sensitive to the choice of specific user parameters. Although these criteria can perform well for model validation in general, they often provide a slight overestimation of the model order [5,6] and are also inadequate to detect and reject the physically irrelevant modes which often appear in the identified models [7]. Such irrelevant modes are here called noise modes without considering of their origin.

In the contrast, in the modal analysis community, the primary interest is often in the physical relevance of the individual modes of the identified model rather than a related model's prediction capacity. Therefore, the common practice is to identify a model with an order that is much higher than motivated by physics to ensure that all physical eigenmodes within the frequency band of interest are safely captured [8–10]. However, this inevitably results in the appearance of noise modes in the identified model, *i.e.,* modes which are present in the model due to measurement noise or computational imprecision but have no relevance to the physics of the tested system. Various tools have been developed to detect and eliminate such noise modes from a model. The most widespread tool is undoubtedly the so-called stabilization diagram [11,12]. This diagram is constructed using estimated eigenfrequencies of models with increasing order. Ideally, for a physical mode, the estimated eigenfrequencies show up with the same value for increasing model order while for a noise mode they do not [6]. However, the interpretation of the stabilization diagram is an art which often requires a lot of user interaction. Specifically, for highly noisy data its outcome highly depends on user decisions.

In recent years, many studies attempted to automate the interpretation of stabilization diagram or the modal parameter estimation algorithm in general [13–17]. Owing to the fact that analyzing the stabilization diagram reduces to finding modes with similar properties, the majority of automation strategies borrow methods from statistical machine learning with supervised and unsupervised learning algorithms. Goethal et al. [12] proposed to utilize a supervised learning algorithm to automate the interpretation of stabilization diagrams. In their study, a hierarchical clustering algorithm groups similar modes of a stabilization diagram together. Then, the final decision on the nature of a cluster, being either physical or a noise artifact, is made by a self-learning Support Vector Machine (SVM) algorithm. Their hypothesis is that once the SVM algorithm is sufficiently trained from sets of data obtained from designed synthetic experiments, the algorithm will automatically classify physical modal parameters for real test data.

Special attention has been given to unsupervised learning algorithms. Hierarchical and centroid-based clustering[1] algorithms are two examples of this type of learning algorithms. Hierarchical clustering starts by assigning one cluster to each data point in a stabilization diagram. Then, it proceeds by merging the closest clusters together until the distance between the resulting clusters exceeds a user-defined threshold. Finally, physical modes are defined by clusters in which the number of modes is larger than a user-specified threshold. A considerable research effort has been made to develop appropriate distance measure for the hierarchical clustering. Magalhães et al. [18] suggested a distance measure which is based on the eigenfrequency difference and the Modal Assurance Criterion (MAC) value. Allemang et al. [19] used the MAC value between pole-weighted mode shapes as the distance measure between clusters. Goethals et al. [12] proposed a distance measure based on the difference of damping ratios and eigenfrequencies.

Other examples of unsupervised learning algorithms are centroid-based clustering schemes such as k-means and fuzzy c-means. In these clustering approaches, a central point/vector (centroid) serves as a representative for each cluster, although it may not be a member of data points in a stabilization diagram. Then the data points are grouped into k clusters such that the squared distances from the cluster centroids are minimized. The main drawback of this approach is that the number of clusters is assumed to be known *a priori*, which is often not the case in practice. Scionti and Lanslots [20] employed fuzzy c-means clustering to directly group the existing modes in the stabilization diagram into a predefined number of clusters. Vanlanduit et al. [6] and Verboven et al. [8] proposed a frequency-domain Maximum Likelihood Estimator (MLE) to estimate the modal parameters using a single high model order n. Subsequently, they grouped the estimated modes into two classes of physical and noise modes using a fuzzy c-means clustering algorithm. Reynders et al. [7] automates the analysis of stabilization diagrams using three steps of clustering. First, a centroid-based clustering algorithm is employed to remove the noise modes from the stabilization diagram. Secondly, a hierarchical clustering algorithm is employed to group similar modes that

[1] In general, clustering refers to the task of subdividing a set of data points into subsets such that the (in some sense) similar data points are grouped together.