Signal Processing 99 (2014) 69–85

Contents lists available at ScienceDirect

Signal Processing

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

Subspace-based spectrum estimation in frequency-domain by regularized nuclear norm minimization

Hüseyin Akçay

Department of Electrical and Electronics Engineering, Anadolu University, Eskişehir 26470, Turkey

ARTICLE INFO

Article history:
Received 6 February 2013
Received in revised form
21 December 2013
Accepted 25 December 2013
Available online 2 January 2014

Keywords:
Identification
Power spectrum
Subspace method
Nuclear norm
Frequency-domain

ABSTRACT

Subspace-based methods have been effectively used to estimate multi-input/multi-output, discrete-time, linear-time-invariant systems from noisy spectrum samples. In these methods, a critical step is splitting of two invariant subspaces associated with causal and non-causal eigenvalues of some structured matrices built from spectrum measurements via singular-value decomposition in order to determine model order. Mirror image symmetry with respect to the unit circle between the eigenvalue sets of the two invariant spaces, required by the subspace algorithms, is lost due to low signal-to-noise ratio, unmodeled dynamics, and insufficient amount of data. Consequently, the choice of model order is not straightforward. In this paper, we propose a new model order selection scheme that is insensitive to noise and undermodeling and based on the regularized nuclear norm optimization in combination with a recently developed subspace-based spectrum estimation algorithm which uses non-uniformly spaced, in frequencies, spectrum measurements. A detailed simulation study shows the effectiveness of the proposed scheme to large amplitude noise over short data records. Examples illustrating application of the proposed scheme to real-life problems are also presented. The proposed scheme can be readily integrated into frequency-domain instrumental variable subspace algorithms to estimate auto-power spectral density or cross-power spectral density function matrices from non-uniformly spaced, in frequencies, spectrum measurements.

1. Introduction

Subspace methods are popular to obtain low-order state-space models of multi-input/multi-output (MIMO), linear-time-invariant (LTI) systems from noise corrupted time or frequency-domain measurements. There is an extensive literature on the topic, and the reader is referred to the monographs [1–3] for a review of its foundations and engineering applications.

The focus of this paper is the estimation of an auto-power spectral density (auto-PSD) function matrix from its noise corrupted samples. This problem arises, for example, in the design of linear-shape filters for noise processes [4–6]. A closely related problem is the estimation of a cross-power spectral density (cross-PSD) function matrix in frequency-domain. The cross-spectral analysis is a fundamental and powerful technique to investigate an unknown relationship between two time series in frequency-domain. It is widely used in many engineering problems; e.g., time delay estimation of spatial sensors [7], blind equalization in communications [8], analysis of feedback systems [9], and system identification of mechanical vibration systems [10].

In [6,11,12], frequency-domain subspace-based identification algorithms were presented. The main idea behind these algorithms is to recognize that range space of a matrix built from frequency-shifted and weighted noise-free samples of a rational spectrum is exactly the linear span of the extended observability matrix associated with

E-mail address: huakcay@anadolu.edu.tr

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.12.028

the causal and anti-causal components of the auto-PSD function matrix in a given state-space realization. This idea has found applications in the auto and the cross-spectrum estimation problems [13] and the retrieval of scalar transfer functions from phase measurements [14]. The algorithms proposed in [6,12] use the spectrum measurements on the uniform grids of frequencies. Furthermore, the matrix used to extract the observability range space has a Hankel structure and it is obtained by the inverse discrete-Fourier transform (IDFT) directly from the spectrum samples. Under some mild noise assumptions, these algorithms are *strongly consistent*. An extension to the irregularly spaced frequencies case was reported in [11], where the strong consistency of the proposed identification algorithm was established assuming that the noise covariance function is known *a priori*.

The identification algorithms in [6,11–14] determine the model order by inspecting the singular values of a matrix used to extract the observability range space. If n is the true order, then the $2n$ most significant singular values and the corresponding left and right singular vectors are to be retained in order to retrieve the observability range space. Implicit in this process is the assumption that there exists a mirror image symmetry with respect to the unit circle between the eigenvalue sets of the causal and the anti-causal invariant spaces of a state-transition matrix. Under this assumption, the causal eigenvalues can be obtained by a Jordan decomposition. See, [6,11–14] for details and Step 5 of Algorithm 2.1 in Section 2.

When the signal-to-noise ratio is low, the true spectrum is more complex than the assumed one, and when the data record is short, the singular-value decomposition (SVD) step is inconclusive since the assumed symmetry relation between the eigenvalues of the invariant spaces does not hold. A two-stage identification algorithm was proposed in [15]. The first stage of this algorithm provides an initial estimate to a parametric optimization problem of the second stage by using an asymptotic form of the subspace identification algorithm proposed in [6]. The minimum-phase property is guaranteed in the second stage via the solution of a conic linear programming problem. This scheme avoids the need to carry out the numerically sensitive split in [6,11–14].

Nuclear norm optimization methods for structured low-rank matrix approximation have been discussed in several recent papers on system identification [16–23]. The nuclear norm of a matrix-valued function as a convex heuristic for minimizing its rank was first proposed in [16]. Minimum nuclear norm solutions often have low rank and in certain applications, for example, low-rank matrix completion problems, the quality of the heuristic can be demonstrated analytically [24,25]. This approach preserves linear structure in matrix approximation unlike the SVD. Convex constraints or regularization terms in the cost function are easily accommodated in this framework. These methods have been primarily developed for Hankel structured low-rank approximation problems in time-domain settings. An extension to frequency-domain was recently made in [19]. In this work, the subspace algorithm developed in [26] to identify multivariable systems from measured frequency response at uniformly spaced frequencies was re-examined from a model validation perspective using a nuclear norm heuristic.

The contents of this paper are as follows. In Section 2, frequency-domain auto-PSD estimation via the algorithm proposed in [12] is reviewed. In Section 3, a variation of this algorithm based on the regularized nuclear norm heuristic is presented. This presentation bears similarities to the study undertaken in [20]. The observability range space revealing matrix studied in this paper for the nuclear norm minimization does not have the Hankel matrix structure in [16–23]. In addition, it is necessary to include weights to achieve strong consistency. As a by product, along the same lines of this paper one can derive regularized nuclear norm variations of the instrumental variable based subspace algorithms to identify auto-PSD and cross-PSD function matrices. In Section 4, first a simulation example is used to demonstrate that the proposed scheme is effective in determining the order and the poles of *minimal* spectral factors over short data records and insensitive to noise. The next two examples illustrate application of the proposed scheme to real-life problems. The first example is concerned with the modeling of acoustic spectra for detecting faults in induction motors. In the second example, we use the proposed scheme to design a linear-shape filter for random road excitations. Section 5 concludes the paper.

The derivation of the Cramér–Rao lower bound (CRLB) on the variance of unbiased spectrum estimators is deferred to Appendix B where we derive the finite-sample and the asymptotic CRLBs for first-order spectral factors and noise filters assuming that the corruptions in the spectrum measurements are zero-mean real Gaussian white-noise. In a simulation example, the mean-square errors of the subspace algorithm and the regularized nuclear norm heuristic are compared against the CRLBs. In Appendix C, we derive the finite-sample CRLB for the fourth-order spectral factor studied in the simulation example of Section 4 and the measurement setup in Appendix B and use the CRLB benchmark to analyze the mean-square and the bias errors of Algorithm 2.1 and (31). Parallel conclusions to Appendix B are drawn. A preliminary version of this paper was partly presented at EUSIPCO 2013 [27].

1.1. Notation

We end this Introduction with remarks on notation. The letters $\mathbf{R}$ and $\mathbf{C}$ denote the fields of real and complex numbers, respectively. Let I_n denote the n by n identity matrix. The k by l matrix of zeros is denoted by $0_{k\times l}$. When obvious from context, this notation will be omitted. The (block) diagonal matrix formed from a given sequence of numbers (matrices) $x_1,\ldots,x_n$ is denoted by $\mathrm{diag}(x_1,\ldots,x_n)$. Let X^T, $\overline{X}$, X^H, $\mathrm{Re}\,X$, $\mathrm{Im}\,X$ denote respectively the transpose, the complex conjugate, the complex conjugate transpose, the real and the imaginary parts of a given complex matrix X. The meaning of the notation X^{-T} or X^{-H} is evident. The Moore–Penrose pseudo inverse of a given full-column rank matrix X is defined by $X^{\dagger}=(X^HX)^{-1}X^H$. The expected value of a given random variable x will be denoted by $\mathbf{E}(x)$. Let $X\otimes Y$ denote the Kronecker product of two given matrices