Automatica 48 (2012) 514–520

Contents lists available at SciVerse ScienceDirect

Automatica

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

Brief paper

Consistency of subspace methods for signals with almost-periodic components ☆

Martina Favaro [a], Giorgio Picci [b,1]

[a] Department of Electronics for Automation, University of Brescia, Italy

[b] Department of Information Engineering, University of Padova, Italy

ARTICLE INFO

Article history:
Received 29 June 2010
Received in revised form
11 May 2011
Accepted 7 August 2011
Available online 28 January 2012

Keywords:
Subspace identification
Frequency estimation
Statistical consistency
Almost periodic signals
Purely deterministic processes

ABSTRACT

It is sometimes claimed in the literature that subspace methods provide consistent estimates, also when the underlying observed signal has purely oscillatory modes (or the generating system has uncontrollable eigenvalues on the unit circle) but a formal proof of this assertion does not seem to exist. In this paper, we prove consistency of subspace methods with purely oscillatory modes. A well-known subspace identification procedure based on canonical correlation analysis and approximate partial realization is shown to be consistent under certain conditions on the purely deterministic part of the generating system. The algorithm uses a fixed finite regression horizon and the proof of consistency does not require that the regression horizon goes to infinity at a certain rate with the sample size N.

1. Introduction

This paper deals with subspace identification of stationary processes with oscillatory components. At a first sight this problem may look like a minor generalization of a standard identification problem which has been exhaustively treated in the literature since the early 1990s. In reality, on one hand the problem encompasses harmonic retrieval; that is, estimation of the harmonic components of a stationary signal in additive noise, a problem of paramount importance in signal processing which, in the multichannel case, cannot be approached by the standard methods like Pisarenko, MUSIC, ESPRIT etc. It seems fair to say that the specialized literature on harmonic retrieval in the case of vector signals, when the additive noise is colored, is still far from offering satisfactory solutions. For this class of signals, on the other hand, subspace system identification appears as a natural choice.

However it is well-known that stationary random processes with periodic components are not ergodic. Non-ergodicity means in particular that the limit when the sample size goes to infinity of the process sample covariance is *sample dependent*. In particular, the limit sample covariance depends on the random amplitudes of its elementary oscillatory components; see, e.g. Söderström and Stoica (1989, pp. 105–109). On the other hand, the asymptotic statistical properties of subspace methods (and, more generally, of correlation-based methods) depend essentially on the limit sample covariances, which in the presence of oscillatory or quasi-periodic components are not equal to the ensemble averages; i.e., do not coincide with the true covariances. Since parameter estimation procedures based on correlation methods require solving linear relations involving estimated sample covariances, a natural question to ask is if the parameter estimates obtained by solving these linear equations are consistent. This is generally true for signals which are second-order ergodic but sample dependence casts doubts on the validity of standard asymptotic statistical properties, like consistency, of subspace methods in this setting. In particular legitimate doubts arise on the validity of the standard proofs of consistency of subspace methods for signals of this type.

Sections 4 and 5 deal with the question of asymptotically recovering the system parameters (modulo similarity) starting from finite data by a standard subspace algorithm, formulated as an approximate partial realization problem. This setting permits to prove almost sure consistency of the algorithm without having to estimate the transient estimation errors inherent in the truncated least-squares regression approach of Peternell (1995), Peternell, Deistler, and Scherrer (1995); Peternell, Scherrer, and Deistler (1996).

Consistency of subspace methods for purely non-deterministic signals (time series) has been proved earlier in the just cited references. However, to the best of the authors' knowledge, a proof

☆ This work was supported in part by the project *New methods and algorithms for Bayesian estimation, Identification and Adaptive Control* funded by the Italian Ministry of higher education (MIUR). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Wolfgang Scherrer under the direction of Editor Torsten Söderström.

E-mail addresses: martina.favaro@ing.unibs.it, martinafavaro@yahoo.it (M. Favaro), picci@dei.unipd.it (G. Picci).

[1] Tel.: +39 049 8277705; fax: +39 049 8277614.

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.08.059

of consistency when there are quasi-periodic components due to uncontrollable eigenvalues on the unit circle, does not exist. The only paper which comes close in spirit to what concerns us here is Bissacco, Chiuso, and Soatto (2007). In this paper however consistency analysis had to be left out as being "beyond the scope of the paper". Finally, note that processes described by systems whose eigenvalues of modulus one are reachable for the driving process noise, do not concern us here as these processes are actually non-stationary and do not contain almost-periodic oscillations.

2. Stationary processes with an almost-periodic component

All random variables/vectors, denoted by lowercase boldface characters, will have zero mean and finite second order moments. The symbol $\mathbb{E}$ denotes mathematical expectation. All random processes will be discrete time. It is a well-known fact that every vector-valued, say m-dimensional, second-order stationary process admits an orthogonal decomposition

$$\mathbf{y}(t) = \mathbf{y}_d(t) + \mathbf{y}_s(t), \quad t \in \mathbb{Z} \tag{2.1}$$

where $\mathbf{y}_d$ and $\mathbf{y}_s$ are the *purely deterministic (p.d.)* and the *purely non-deterministic (p.n.d.)* components, the latter with an absolutely continuous spectrum and a log-integrable spectral density; see e.g. Rozanov (1967). If $\mathbf{y}$ admits finite-dimensional realizations it can be described by a minimal state space model of the form,

$$\begin{bmatrix} \mathbf{x}(t+1) \\ \mathbf{z}(t+1) \end{bmatrix} = \begin{bmatrix} A_d & 0 \\ 0 & A_s \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{z}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ K \end{bmatrix} \mathbf{e}(t) \tag{2.2a}$$

$$\mathbf{y}(t) = \begin{bmatrix} C_d & C_s \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{z}(t) \end{bmatrix} + \mathbf{e}(t) \tag{2.2b}$$

where the undriven subsystem with p.d. output $\mathbf{y}_d(t) := C_d\mathbf{x}(t)$, described by an observable pair (A_d, C_d) has a positive definite initial state covariance matrix $P_d = \mathbb{E}\,\mathbf{x}(0)\mathbf{x}(0)^\top$. The minimal triplet (C_s, A_s, K) describing the p.n.d. component $\mathbf{y}_s(t) = C_s\mathbf{z}(t) + \mathbf{e}(t)$ originates a stable minimum phase transfer function $I + C_s(zI - A_s)^{-1}K$. Here $\mathbf{e}$ will be taken to be the innovation process of $\mathbf{y}_s$ (and hence of $\mathbf{y}$ as well), having a positive definite covariance matrix $\Delta := \mathbb{E}\,\mathbf{e}(t)\mathbf{e}(t)^\top$ which we shall write in factorized form as $\Delta = DD^\top$ with a nonsingular factor D. In the following we shall need a.s. convergence of the sample second order moments of the p.n.d. output component. To ensure this (second order ergodicity) we may assume that $\mathbf{e}$ is a stationary martingale difference with finite fourth order moments. See Hannan and Deistler (1988) or Peternell et al. (1995, Section 3).

By stationarity and minimality the two block-vector components of the initial state $[\,\mathbf{x}(0)^\top\ \mathbf{z}(0)^\top\,]^\top$ of (2.2) must be uncorrelated. Each has a positive definite covariance matrix, satisfying the Lyapunov equations

$$P_d = AP_dA^\top, \qquad P_s = A_sP_sA_s^\top + KK^\top. \tag{2.3}$$

We shall denote by d the dimension of the p.d. subsystem and by p the dimension of the p.n.d. subsystem in (2.2) and let $n := d + p$. Occasionally we shall use the more compact notations

$$C = \begin{bmatrix} C_d & C_s \end{bmatrix}, \qquad P := \operatorname{diag}\{P_d, P_s\}, \qquad A := \operatorname{diag}\{A_d, A_s\}. \tag{2.4}$$

A special class of signals (2.1) is obtained when the p.n.d. component is white noise; i.e. $\mathbf{y}(t) = \mathbf{y}_d(t) + \mathbf{e}(t)$. Due to their importance in diverse applications, especially frequency estimation, a huge literature has been devoted to the identification of these signals; see e.g. the book (Stoica & Moses, 2005) and the references therein.

Since $\mathbf{y}_d$ and $\mathbf{y}_s$ are completely uncorrelated, the covariance function of the output process $\mathbf{y}$ splits into its p.d and p.n.d. components

$$\Lambda(\tau) := E\mathbf{y}(t+\tau)\mathbf{y}(t)^\top = \Lambda_d(\tau) + \Lambda_s(\tau)$$

with the p.n.d. part having the well-known representation, see e.g. Anderson (1969),

$$\begin{cases} \Lambda_s(\tau) = C_sA_s^{\tau-1}\bar{C}_s^\top & \text{for } \tau = 1, 2, \ldots \\ \Lambda_s(0) = C_sP_sC_s^\top + DD^\top & \text{for } \tau = 0 \end{cases} \tag{2.5}$$

where $\bar{C}_s^\top = A_sP_sC_s^\top + KD^\top$. The structure of Λ_d will emerge from the analysis which follows.

We can choose an orthonormal basis in which $P_d = I$, and A_d is an orthogonal (and hence diagonalizable) matrix with complex eigenvalues $e^{\pm i\theta_k}$, $k = 1, \ldots, \nu$ and possibly real eigenvalues at $\theta_0 = 0$ and $\theta_{\nu+1} = \pi$. Hence A_d is similar to a block-diagonal real matrix

$$A_d = \operatorname{diag}\left\{ I_{n_0}, \begin{bmatrix} \cos\theta_1 I_{n_1} & -\sin\theta_1 I_{n_1} \\ \sin\theta_1 I_{n_1} & \cos\theta_1 I_{n_1} \end{bmatrix}, \ldots \begin{bmatrix} \cos\theta_\nu I_{n_\nu} & -\sin\theta_\nu I_{n_\nu} \\ \sin\theta_\nu I_{n_\nu} & \cos\theta_\nu I_{n_\nu} \end{bmatrix}, -I_{n_{\nu+1}} \right\} \quad \theta_k \neq \theta_j \tag{2.6}$$

where $n_1, \ldots, n_\nu$ are the multiplicities of the complex eigenvalues $e^{i\theta_k}$ appearing in conjugate pairs and n_0 and $n_{\nu+1}$ are the multiplicities of the real eigenvalues $\lambda = 1$ and $\lambda = -1$, some or both of which may possibly be absent. Observability implies that the output dimension m must be an upper bound for the multiplicity of the eigenvalues. Hence for a scalar process n_0 and $n_{\nu+1}$ are ≤ 1 and there are just ν elementary 2×2 oscillatory blocks each corresponding to one of the ν distinct angular frequencies θ_k, $k = 1, \ldots, \nu$, which are strictly between $\theta = 0$ and $\theta = \pi$.

The $m \times d$ (where $d = 2\sum n_k + n_0 + n_{\nu+1}$) matrix C_d splits into blocks $\begin{bmatrix} C_0 & C_1 & \cdots & C_\nu & C_{\nu+1} \end{bmatrix}$ where C_0 and $C_{\nu+1}$ are $m \times n_0$ and $m \times n_{\nu+1}$ and the $C_k, = 1, 2, \ldots, \nu$ are $m \times 2n_k$. The diagonal block elements in A_d are denoted by A_k. Starting from the complex representation where the matrix A_d is diagonal, each corresponding matrix C_k is (complex and) of full row rank. Thus there exists a collection of rows such that the corresponding submatrix is nonsingular. Using this matrix to transform the basis one can achieve a unity matrix in these rows. Converting back to real matrices then achieves the specific form

$$C_k = \Pi_k \begin{bmatrix} I_{n_k} & 0 \\ H_{k,1} & H_{k,2} \end{bmatrix}, \quad k = 1, \ldots, \nu$$

where Π_k is a permutation matrix and the row-block $[H_{k,1}\ H_{k,2}]$ is $(m - n_k) \times 2n_k$. In the scalar case $c_k = [\,1, 0\,]$ for $k = 1, 2, \ldots, \nu$ and 1 otherwise. Returning to the (complex) basis in which A_d is diagonal, it is easy to see that stationarity implies that all the n_k-dimensional complex state subvectors $\mathbf{z}_k(t) := \mathbf{x}_{1,k}(t) + i\bar{\mathbf{x}}_{2,k}(t)$ and $\bar{\mathbf{z}}_k(t) := \mathbf{x}_{1,k}(t) - i\bar{\mathbf{x}}_{2,k}(t)$, $k = 1, 2, \ldots, \nu$ and the random vectors $\mathbf{x}_k(t)$, $k = 1, \nu + 1$ must be mutually uncorrelated. This implies in particular that

$$\begin{aligned} \mathbb{E}\{\mathbf{z}_k(0)\,\bar{\bar{\mathbf{z}}}_k(0)^\top\} &= \mathbb{E}\{\mathbf{x}_{1,k}(0)\mathbf{x}_{1,k}(0)^\top\} - \mathbb{E}\{\mathbf{x}_{2,k}(0)\mathbf{x}_{2,k}(0)^\top\} \\ &\quad + i\left(\mathbb{E}\{\mathbf{x}_{1,k}(0)\mathbf{x}_{2,k}(0)^\top\} \right. \\ &\quad \left. + \mathbb{E}\{\mathbf{x}_{2,k}(0)\mathbf{x}_{1,k}(0)^\top\}\right) = 0 \end{aligned}$$

so that the covariance of $\mathbf{x}_k(0) = \begin{bmatrix} \mathbf{x}_{1,k}(0)^\top & \mathbf{x}_{2,k}(0)^\top \end{bmatrix}^\top$ must have the following structure

$$P_k = \mathbb{E}\,\mathbf{x}_k(0)\mathbf{x}_k(0)^\top = \begin{bmatrix} \Sigma_k & M_k \\ M_k^\top & \Sigma_k \end{bmatrix} \quad M_k = -M_k^\top$$