Automatica 48 (2012) 2145–2151

Contents lists available at SciVerse ScienceDirect

Automatica

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

Brief paper

On the estimation of structured covariance matrices[☆]

Mattia Zorzi [1], Augusto Ferrante

Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy

ARTICLE INFO

Article history:
Received 4 April 2011
Received in revised form
6 January 2012
Accepted 21 January 2012
Available online 30 June 2012

Keywords:
Covariance matrix estimation
Covariance lags sequence
Positive semi-definite matrix
Toeplitz matrices

ABSTRACT

This paper discusses a method for estimating the covariance matrix of a multivariate stationary process w generated as the output of a given linear filter fed by a stationary process y. The estimated covariance matrix must satisfy two constraints: it must be positive semi-definite and it must be consistent with the fact that w is the output of the given linear filter. It turns out that these constraints force the estimated covariance to lie in the intersection of a cone with a linear space. While imposing only the first of the two constraints is rather straightforward, guaranteeing that both are satisfied is a non-trivial issue to which quite a bit of attention has already been devoted in the literature. Our approach extends the method for estimating the *Toeplitz* covariance matrix of order M of a process y based on the *biased spectral estimator* (Stoica & Moses, 1997). This extension is based on the characterization of the output covariance matrix in terms of the filter parameters and the sequence of covariance lags of the input process.

After introducing our estimation method, we propose a comparison performance between this one and other methods proposed in the literature. Simulation results show that our approach constitutes a valid estimation procedure.

1. Introduction

In this paper, we consider the process $w = \{w_k\}_{k=-\infty}^{\infty}$ obtained as the output of a given stable rational filter $G(z)$ fed by a stationary process $y = \{y_k\}_{k=-\infty}^{\infty}$. We assume to observe a finite-length collection of sample data $y_1, \ldots, y_N$ of the stochastic process y. We want to compute an estimate $\hat{\Sigma}$ of the covariance $\Sigma := E[w_k w_k^*]$ in such a way that $\hat{\Sigma}$ is both positive semi-definite and consistent with the filter $G(z)$. Here $*$ denotes transposition plus conjugation. To analyze the features of this problem and to provide some motivations and applications, we discuss a very simple example. Let y be a real scalar second-order stationary process and let $G(z)$ be a bank of l delays:

$$G(z) := \begin{bmatrix} z^{-l} & z^{-l+1} & \cdots & z^{-1} \end{bmatrix}^\top. \tag{1}$$

In this case, the covariance matrix Σ of the output[2] w has the form of a symmetric *Toeplitz* matrix having the first l covariance lags of y on the first row:

$$\Sigma := \begin{bmatrix} r_0 & r_1 & \cdots & r_{l-1} \\ r_1 & r_0 & \ddots & r_{l-2} \\ \vdots & \ddots & \ddots & \ddots \\ r_{l-1} & \ddots & r_1 & r_0 \end{bmatrix}, \qquad r_h := E[y_{k+h} y_k^*]. \tag{2}$$

If we need to estimate Σ, it is natural to impose that the estimate $\hat{\Sigma}$ be positive semi-definite and have Toeplitz structure. On the one hand, one can consider the estimate $\hat{\Sigma}$ obtained by computing the sample covariance lags of y and constructing the corresponding Toeplitz matrix. This estimate, however, is *not* guaranteed to be positive semi-definite. On the other hand, one can compute the sample covariance $\hat{\Sigma}_C := \sum_{k=1}^{N} w_k w_k^*$ of the output process w. The latter is, by construction, positive semi-definite but is *not* guaranteed to be Toeplitz. Notice, in passing, that the orthogonal projection of this estimate onto the linear space of Toeplitz matrices is no longer guaranteed to be positive semi-definite. This problem, yet important, is very special due to the FIR structure of $G(z)$ in (1). In this case, it is well-known that the problem can be solved by computing, from $y_1, \ldots, y_N$, the estimates $\hat{r}_h$ of the r_h in (2), with the *biased correlogram spectral estimator* (Stoica & Moses, 1997). Alternatively, one can use a constrained convex optimization approach (Burg, Luenberger, & Wenger, 1982; Ferrante, Pavon, & Zorzi, 2012).

The estimation of positive semi-definite Toeplitz matrices is just an instance of a class of problems in digital signal processing

☆ Work partially supported by the Italian Ministry for Education and Research (MIUR) under PRIN grant n. 20085FFJ2Z "New Algorithms and Applications of System Identification and Adaptive Control". The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Er-Wei Bai under the direction of Editor Torsten Söderström.

E-mail addresses: zorzimat@dei.unipd.it (M. Zorzi), augusto@dei.unipd.it (A. Ferrante).

1 Tel.: +39 049 8277702; fax: +39 049 8277799.

2 In this case the output coincides with the state process.

0005-1098/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2012.05.057

where the covariance matrix of the output process of a general linear filter has to be estimated with the knowledge of the input sample data. The importance of these problems is due to the development of a family of spectral estimation methods introduced by Byrnes, Georgiou and Lindquist in Byrnes, Georgiou, and Lindquist (2000), and Byrnes, Georgiou, and Lindquist (2001), and further developed and modified in Georgiou (2002a), Ferrante, Masiero, and Pavon (in press) and Ferrante, Pavon, and Ramponi (2008). These methods, for which $y_1, \dots, y_N$ and $G(z)$ are the given data, are based on a moment problem that requires an estimate of the covariance matrix of the output w. The first of these spectral estimation methods was called "THREE", Byrnes et al. (2000): we shall thus refer to these methods as "THREE-like".

For the special case of linear filters $G(z)$ whose output is the state of the filter, the problem of characterizing the output covariance Σ has been addressed by Georgiou in Georgiou (2001) and Georgiou (2002b). This characterization can be employed to estimate the state covariance by resorting to the maximum likelihood approach proposed in Burg et al. (1982) which, however, requires that the state covariance Σ and the sample covariance $\hat{\Sigma}_C$ are strictly positive definite. In Ferrante et al. (2012), a maximum entropy problem has been proposed that leads to a positive definite estimate $\hat{\Sigma}$ consistent with the filter structure. Notice that also this technique requires that the state covariance Σ and the sample covariance $\hat{\Sigma}_C$ are strictly positive definite and that the filter's output and state coincide. On the other hand, these techniques do not exploit the knowledge of $y_1, \dots, y_N$ that, in the THREE-like methods, are the problem data.

The purpose of this paper is to introduce a new approach—based on the knowledge of the input sample data $y_1, \dots, y_N$—to compute a positive semi-definite estimate $\hat{\Sigma}$ whose structure is consistent with an arbitrary, finite dimensional, stable, linear filter $G(z)$. Our method, which is an extension of the one for estimating the *Toeplitz* covariance matrix of order M of the process y based on the *biased spectral estimator* (Stoica & Moses, 1997), hinges on the characterization of Σ in terms of the filter $G(z)$ and the covariance lags sequence of the input process y. Thus, given an estimate of the covariance lags sequence of the input process, we can compute an estimate $\hat{\Sigma}$ consistent with the structure imposed by the filter. It will be shown that if we consider the sample covariance lags used in the biased correlogram spectral estimator we can guarantee that $\hat{\Sigma} \geq 0$.

The paper is organized as follows. In the next section, we present a more precise formulation of the problem. In Section 3, the vector space containing the covariance matrices Σ is characterized in terms of the filter $G(z)$. Section 4 is devoted to introduce our approach based on the covariance lags. In Section 5, we briefly discuss other approaches available in the literature and their possible generalizations. Section 6 is devoted to simulations: we compare covariance matrices estimated by our method with the ones obtained using alternative approaches. In Section 7, we draw our conclusions.

2. Problem formulation

Consider a linear filter

$$x_{k+1} = Ax_k + By_k$$
$$w_k = Cx_k + Dy_k, \quad k \in \mathbb{Z}, \tag{3}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $C \in \mathbb{C}^{p\times n}$, $D \in \mathbb{C}^{p\times m}$ and A has all its eigenvalues in the open unit disk. The input process y is $\mathbb{C}^m$-valued, wide sense stationary and purely nondeterministic. As mentioned in the Introduction, $\Sigma = \Sigma^* \geq 0$ denotes the covariance matrix of the (stationary) output process w and we denote by

$$G(z) = C(zI - A)^{-1}B + D \tag{4}$$

the filter transfer function. Let $\mathfrak{H}_m$ be the m^2-dimensional, *real* vector space of Hermitian matrices of dimension $m\times m$ and $\mathfrak{H}_{m,+}$ be the intersection between $\mathfrak{H}_m$ and the closed cone of positive semi-definite matrices. We denote by $C(\mathbb{T}, \mathfrak{H}_m)$ the family $\mathfrak{H}_m$-valued, continuous functions on the unit circle $\mathbb{T}$. Consider now the linear operator

$$\Gamma : C(\mathbb{T}, \mathfrak{H}_m) \to \mathfrak{H}_p, \quad \Psi \mapsto \int G\Psi G^*, \tag{5}$$

where integration takes place on $\mathbb{T}$ with respect to the normalized Lebesgue measure $d\vartheta/2\pi$. It follows that Σ belongs to the linear space

$$\text{Range}\,\Gamma := \left\{ M \in \mathfrak{H}_p \mid \exists \Psi \in C(\mathbb{T}, \mathfrak{H}_m) \text{ such that } \int G\Psi G^* = M \right\}. \tag{6}$$

Suppose now that A, B, C, D are known and a sample data $\{y_k\}_{i=1}^N$ is given. We want to compute an estimate $\hat{\Sigma}$ of Σ such that

$$\hat{\Sigma} \in [\text{Range}\,\Gamma]_+ := \text{Range}\,\Gamma \cap \mathfrak{H}_{p,+}. \tag{7}$$

If we feed $G(z)$ with the data $\{y_k\}_{k=1}^N$ and we collect the output data $\{w_k\}_{k=1}^N$, an estimate of Σ is given by the sample covariance $\hat{\Sigma}_C := \frac{1}{N}\sum_{k=1}^N w_k w_k^* \geq 0$. This estimate, as it happened in the example discussed in the Introduction, normally fails to belong to Range Γ. In fact, Range Γ is a linear vector subspace usually strictly contained in $\mathfrak{H}_p$. One could project $\hat{\Sigma}_C$ onto Range Γ obtaining a new Hermitian matrix $\hat{\Sigma}_\Gamma$. This matrix $\hat{\Sigma}_\Gamma$, however, may be indefinite and this is particularly likely when N is not large. In addition, when the linear filter $G(z)$ does not satisfy particular properties, the computation of a basis for Range Γ is not trivial.

3. Characterization of Range Γ

We start by considering a particular, yet very relevant, situation. We will later deal with the general case.

3.1. State covariance matrices

Next we restrict attention to the case when $C = I_n$ and $D = 0_{n\times m}$, with $m < n$, so that Σ is a state covariance matrix. Under the additional assumptions that (A, B) is a reachable pair and B has full column rank, it was shown in Georgiou (2001) and Georgiou (2002b) (see also Ramponi, Ferrante, & Pavon, 2010), that an $n \times n$ matrix $M \in \mathfrak{H}_n$ belongs to Range Γ if and only if there exists $H \in \mathbb{C}^{m\times n}$ such that

$$M - AMA^* = BH + H^*B^*. \tag{8}$$

Moreover, it is possible to prove that Range Γ has *real dimension* equal to $m(2n - m)$, Ferrante et al. (2012).

We now want to relax the reachability assumption. To this end, we derive a preliminary result. Consider an (A, B) pair and the operator Γ corresponding to $G(z) = (zI - A)^{-1}B$. We perform a state space transformation induced by an invertible matrix $T \in \mathbb{C}^{n\times n}$,

$$\tilde{A} := T^{-1}AT, \qquad \tilde{B} := T^{-1}B. \tag{9}$$

We define the corresponding operator

$$\tilde{\Gamma} : C(\mathbb{T}, \mathfrak{H}_m) \to \mathfrak{H}_n, \qquad \Psi \mapsto \int \tilde{G}\Psi\tilde{G}^* \tag{10}$$

with $\tilde{G}(z) = (zI - \tilde{A})^{-1}\tilde{B} = T^{-1}G(z)$. Note that

$$\int G\Psi G^* = \int T\tilde{G}\Psi\tilde{G}^*T^*, \quad \forall \Psi \in C(\mathbb{T}, \mathfrak{H}_m). \tag{11}$$