11

Application of a Global Inverse Function Theorem of Byrnes and Lindquist to a Multivariable Moment Problem with Complexity Constraint[*]

Augusto Ferrante[1], Michele Pavon[2], and Mattia Zorzi[3]

[1] Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy
[2] Dipartimento di Matematica Pura ed Applicata, Università di Padova, via Trieste 63, 35131 Padova, Italy
[3] Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy

Summary. A generalized moment problem for multivariable spectra in the spirit of Byrnes, Georgiou and Lindquist is considered. A suitable parametric family of spectra is introduced. The map from the parameter to the moments is studied in the light of a global inverse function theorem of Byrnes and Lindquist. An efficient algorithm is proposed to find the parameter value such that the corresponding spectrum satisfies the moment constraint.

11.1 Introduction

This paper represents an attempt to pay a tribute to two great figures of Systems and Control Theory. It would be difficult to even mention the long string of benchmark contributions that we owe to Anders and Chris. It would entail listing results in linear and nonlinear control, deterministic and stochastic systems, finite and infinite dimensional problems, etc. This string, no matter how much compactification we drew from string theory, would simply be too long. So we leave this task to the many that are better qualified than us. We like to stress, instead, two other aspects of their long lasting influence in the systems and control community. One is that both have devoted a lot of time and energy to form young researchers. Their generous help and tutoring to students and junior scientists continues unabated to this day. A second peculiar aspect of Anders and Chris is that they embody at its best an American-European scientist, having strong cultural and scientific ties on both sides of the ocean. For instance, it is not by chance that both have contributed so much over the years to MTNS, one of the few conferences that belongs equally to the US and to Europe (and to the rest of the world).

Over the past decade Anders, Chris and Tryphon Georgiou, together with a number of coworkers and students, have developed a whole new field that may be called *Moment Problems with Complexity Constraint*, see [4, 16] and references therein. Their generalized moment problems include as special cases some of the most central problems in our field such as the *covariance extension problem* (see the next section) and *Nevanlinna-Pick interpolation* of robust control. The mathematics, involving global inverse function theorems, differential geometry, analytic interpolation, convex optimization, homotopy methods, iterative numerical schemes, etc. is particularly rich and beautiful. Significant applications to spectral estimation have already been developed. One of the key to the success of this broad program has been the establishing by Anders and Chris of suitable global inverse functions theorems generalizing Hadamard's type theorems, see [3] and references therein. These can be applied in manifold ways. For the generalized moment problems with entropy-like criterions, they yield existence for the dual problem which is typically a convex optimization problem with *open*, *unbounded* domain.

In this paper, we try to exploit this result of Anders and Chris to circumvent one of the stumbling blocks in this field. We deal, namely, with the multivariable problem where the spectrum must satisfy a suitable generalized moment constraint and must be of limited complexity. We consider the situation where an "a priori" estimate Ψ of the spectrum is available. Motivated by the scalar case and multivariate, $\Psi = I$ case solutions, we introduce a suitable parametric family of spectra with bounded McMillan degree. We then establish properness of the map from the parameter to the moments. Injectivity, and hence surjectivity, of this map is then proven in a special case. A multivariate generalization of the efficient algorithm [21, 7, 9] is finally proposed.

We employ the following notation. For a complex matrix A, A^* denotes the transpose conjugate of A. We denote by $\mathfrak{H}_n$ the vector space of Hermitian matrices of dimension $n \times n$ endowed with the inner product $\langle P, Q \rangle := \mathrm{tr}(PQ)$, and by $\mathfrak{H}_{+,n}$ the subset of positive definite matrices. For a matrix-valued rational function $\chi(z) = H(zI - F)^{-1}G + J$, we define $\chi^*(z) = G^*(z^{-1}I - F^*)^{-1}H^* + J^*$. We denote by $\mathbb{T}$ the unit circle in the complex plane $\mathbb{C}$ and by $C(\mathbb{T})$ the family of complex-valued, continuous functions on $\mathbb{T}$. $C_+(\mathbb{T})$ denotes the subset of $C(\mathbb{T})$ whose elements are real-valued, positive functions. Finally, $C(\mathbb{T}; \mathfrak{H}_m)$ stands for the space of $\mathfrak{H}_m$-valued continuous functions.

11.2 A Generalized Moment Problem

Consider the rational transfer function

$$G(z) = (zI - A)^{-1}B, \qquad A \in \mathbb{C}^{n \times n}, B \in \mathbb{C}^{n \times m}, \quad n \geq m \qquad (11.1)$$

of the system

$$x(t+1) = Ax(t) + By(t),$$

where A is a stability matrix, i.e. has all its eigenvalues in the open unit disc, (A, B) is a *reachable pair*, and B is a full column rank matrix. The transfer function G

models a bank of filters fed by a stationary process $y(t)$ of unknown spectral density $\Phi(z)$. We assume that we know (or that we can reliably estimate) the steady-state covariance Σ of the state x of the filter. We have

$$\Sigma = \int G\Phi G^*,$$

where, here and in the sequel, integration occurs on the unit circle with respect to the normalized Lebesgue measure. Let $\mathscr{S}_m = \mathscr{S}_+^{m \times m}(\mathbb{T})$ be the family of $\mathfrak{H}_{+,m}$-valued functions defined on the unit circle which are bounded and coercive. We consider the following *generalized moment problem*:

Problem 11.2.1. Let $\Sigma \in \mathfrak{H}_{+,n}$ and $G(z)(zI - A)^{-1}B$ of dimension $n \times m$ with the same properties as in (11.1). Find Φ in $\mathscr{S}_m$ that satisfies

$$\int G\Phi G^* = \Sigma. \tag{11.2}$$

The question of *existence* of $\Phi \in \mathscr{S}_m$ satisfying (11.2) and, when existence is granted, the *parametrization* of all solutions to (11.2), may be viewed as a generalized moment problem. For instance, let $C_k := E\{y(n)y^*(n+k)\}$, and take

$$A = \begin{bmatrix} 0 \ I_m & 0 & \dots & 0 \\ 0 & 0 \ I_m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots \ I_m \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix}, \quad \Sigma = \begin{bmatrix} C_0 & C_1 & C_2 & \dots & C_{n-1} \\ C_1^* & C_0 & C_1 & \ddots & C_{n-2} \\ C_2^* & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ C_{n-1}^* & C_{n-2}^* & \ddots & \ddots & C_0 \end{bmatrix},$$

so that $G(z)$ is a block-column with k-th component being $G_k(z) = z^{k-n-1}I$. This is the classical *covariance extension problem*, where the information available is the finite sequence of covariance lags $C_0, C_1, \dots, C_{n-1}$ of the process y. It is known that the set of densities consistent with the data is nonempty if $\Sigma \geq 0$ and contains infinitely many elements if $\Sigma > 0$ [17], see also [10, 1, 2, 11]. Other important problems of Systems and Control Theory, such as the Nevanlinna-Pick interpolation problem, may be cast in the frame of Problem 11.2.1, see [15].

It may be worthwhile to recall that moment problems form a special class of inverse problems that are typically *not* well-posed in the sense of Hadamard[1]. When Problem 11.2.1 is feasible, a unique solution may be obtained by minimizing a suitable criterion: We mention the Kullback-Leibler type criterion employed in [15] and a suitable multivariable Hellinger-type distance introduced in [8, 23]. The reader is referred to these papers for full motivation, and to [22] for results on the well-posedness of these optimization problems. In [5, 15, 14, 3], a different, interesting viewpoint if taken. It is namely there shown that all solutions to Problem 11.2.1 may

---

[1]A problem is said to be well-posed, in the sense of Hadamard, if it admits a solution, such a solution is unique, and the solution depends continuously on the data.

be obtained as minimizers of a suitable entropy-like (pseudo-)distance from an "a priori" spectrum Ψ as the latter varies in $\mathscr{S}_m$. There, Ψ is thought of as a parameter. This viewpoint leads to the more challenging moment problem with degree constraint. The latter consists in finding solutions to Problem 11.2.1 whose McMillan degree is "a priori" bounded.

Existence of $\Phi \in \mathscr{S}_m$ satisfying (11.2) in the general case is a nontrivial matter. It has been shown that the following conditions are equivalent [13]:

1. The family of $\Phi \in \mathscr{S}_m$ satisfying constraint (11.2) is nonempty;
2. there exists $H \in \mathbb{C}^{m \times n}$ such that

$$\Sigma - A\Sigma A^* = BH + H^*B^*; \qquad (11.3)$$

3. the following rank condition holds

$$\text{rank} \begin{pmatrix} \Sigma - A\Sigma A^* & B \\ B^* & 0 \end{pmatrix} = \text{rank} \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}. \qquad (11.4)$$

A fourth equivalent condition is based on the linear operator $\Gamma : C(\mathbb{T}; \mathfrak{H}_m) \to \mathfrak{H}_n$ that will play a crucial role in the rest of the paper:

$$\Gamma : \Phi \mapsto \int G\Phi G^*. \qquad (11.5)$$

Existence of $\Phi \in C(\mathbb{T} : \mathfrak{H}_m)$ satisfying $\int G\Phi G^* = \Sigma$ can be expressed as

$$\Sigma \in \text{Range}\,\Gamma. \qquad (11.6)$$

It is has been shown in [12] that when there is a spectrum Φ in $\mathscr{S}_m$ satisfying (11.2), then there exists also $\Phi^o \in C(\mathbb{T}; \mathfrak{H}_m)$ (the *maximum entropy* spectrum (11.15) below) satisfying (11.2). Thus, condition (11.6) will be a standing assumption in this paper. For $X \in \mathfrak{H}_n$ and $\Phi \in C(\mathbb{T}; \mathfrak{H}_m)$, we have

$$\left\langle X, \int G\Phi G^* \right\rangle = \text{tr}\left(X \int G\Phi G^* \right) = \text{tr}\left(\int (G^*XG)\Phi \right).$$

We conclude that $\Gamma^* : \mathfrak{H}_n \to C(\mathbb{T}; \mathfrak{H}_m)$, the *adjoint map* of Γ, is given by

$$\Gamma^* : X \mapsto G^*XG, \qquad (11.7)$$

and

$$(\text{Range}\,\Gamma)^\perp = \left\{ X \in \mathfrak{H}_n | G^*(e^{j\vartheta})XG(e^{j\vartheta}) = 0, \ \forall e^{j\vartheta} \in \mathbb{T} \right\}. \qquad (11.8)$$

11.3 Kullback-Leibler Approximation of Spectral Densities

In this section, we recall some important result obtained in the *scalar case*, i.e. the case when $m = 1$. In [15], a Kullback-Leibler type of distance for spectra in $\mathscr{S}_1 := \mathscr{S}_+^{1 \times 1}(\mathbb{T})$ was introduced:

$$d(\Psi\|\Phi) = \int \Psi \log\left(\frac{\Psi}{\Phi}\right).$$

As is well known, this pseudo-distance originates in hypothesis testing, where it represents the mean information for observation for discrimination of an underlying probability density from another [19, p.6]. It also plays a central role in several other fields of science such as information theory, identification, stochastic processes, statistical mechanics, etc., where it goes under different names such as *divergence*, *relative entropy*, *information distance*, etc. If $\int \Phi = \int \Psi$, we have $d(\Psi\|\Phi) \geq 0$. The choice of $d(\Psi\|\Phi)$ as a distance measure, even for spectra that have different zeroth moment, is discussed in [15, Section III]. Minimizing $\Phi \to d(\Psi\|\Phi)$ rather than $\Phi \to d(\Phi\|\Psi)$ is unusual with respect to the statistics-probability-information theory world. Minimizing with respect to the first argument, however, leads to a non-rational solution even when Ψ is rational (see below). Moreover, this atypical minimization includes as special case ($\Psi \equiv 1$) maximization of entropy. In [15], the following problem is considered:

Problem 11.3.1. Given $\Psi \in \mathscr{S}_1$ and $\Sigma \in \mathfrak{H}_{+,n}$,

$$\begin{aligned}&\text{minimize} \quad d(\Psi\|\Phi) \\ &\text{over} \quad \left\{\Phi \in \mathscr{S}_1 \mid \int G\Phi G^* = \Sigma\right\}.\end{aligned}$$

Let

$$\mathscr{L}_+ := \{\Lambda \in \mathfrak{H}_n : G^*\Lambda G > 0, \forall e^{j\vartheta} \in \mathbb{T}\}. \tag{11.9}$$

For $\Lambda \in \mathscr{L}_+$, consider the *unconstrained* minimization of the *Lagrangian function*

$$\begin{aligned}L(\Phi,\Lambda) &= d(\Psi\|\Phi) + \text{tr}\left(\Lambda\left(\int G\Phi G^* - \Sigma\right)\right) \\ &= d(\Psi\|\Phi) + \int G^*\Lambda G\Phi - \text{tr}(\Lambda\Sigma),.\end{aligned} \tag{11.10}$$

This is a convex optimization problem. The variational analysis in [15] shows that the unique minimizer is given by

$$\hat{\Phi}_{KL} = \frac{\Psi}{G^*\Lambda G}. \tag{11.11}$$

Thus, the original Problem 11.3.1 is now reduced to finding $\hat{\Lambda} \in \mathscr{L}_+$ satisfying

$$\int G\frac{\Psi}{G^*\hat{\Lambda}G}G^* = \Sigma. \tag{11.12}$$

This is accomplished via duality theory. The dual problem turns out to be equivalent to minimizing a strictly convex function on the open and unbounded set $\mathscr{L}_+^\Gamma = \mathscr{L}_+ \cap$ Range(Γ). A global inverse function theorem of Byrnes and Lindquist is then used to establish existence and uniqueness for the dual problem under the assumption

of feasibility of the primal problem, see [3], references therein and [7]. Notice that, when Ψ is rational, (11.11) shows that the degree of the solution is "a priori" bounded by $2n$ plus the degree of Ψ.

In practical applications, the solution of the dual problem is a numerical challenge. In fact, the dual variable is an Hermitian matrix and, as discussed in [15], the reparametrization in vector form may lead to a loss of convexity. Moreover, the dual functional and its gradient tend to infinity at the boundary. To efficiently deal with the dual problem, the following algorithm has been proposed in [21] and further discussed in [7]:

$$\Lambda_{k+1} = \Theta(\Lambda_k) := \int \Lambda_k^{1/2} G \left[\frac{\Psi}{G^* \Lambda_k G} \right] G^* \Lambda_k^{1/2}, \qquad \Lambda_0 = \frac{1}{n} I. \qquad (11.13)$$

It has been shown in [21] that Θ maps *density matrices* to density matrices, i.e. if Λ is a positive semi-definite Hermitian matrix with trace equal to 1, then $\Theta(\Lambda)$ has the same properties. Moreover, Θ maintains positive definiteness, i.e., if $\Lambda > 0$, then $\Theta(\Lambda) > 0$. If the sequence $\{\Lambda_k\}$ converges to a limit point $\hat{\Lambda} > 0$ then such a $\hat{\Lambda}$ is a fixed point for the map Θ and hence satisfies (11.12). It has been recently shown in [9] that $\{\Lambda_k\}$ is locally asymptotically convergent to a limit point $\hat{\Lambda}$ that satisfies (11.12).

11.4 The Multivariable Case

Let us go back to the multivariable setting of Problem 11.2.1. Inspired by the Umegaki relative entropy of statistical quantum mechanics [20], we define $d(\Psi \| \Phi)$ for Φ and Ψ in $\mathscr{S}_m$

$$d(\Psi \| \Phi) = \int tr(\Psi(\log \Psi - \log \Phi)). \qquad (11.14)$$

Consider first the case where $\Psi = I$ the identity matrix. Then Problem 11.3.1 turns into the *maximum entropy problem*:

Problem 11.4.1. Given $\Sigma \in \mathfrak{H}_{+,n}$,

$$\text{maximize} \quad \int tr \log \Phi = \int \log \det \Phi = -d(I \| \Phi)$$

$$\text{over} \quad \left\{ \Phi \in \mathscr{S}_m \mid \int G \Phi G^* = \Sigma \right\}.$$

In [12], the following result was established that (considerably) generalizes Burg's maximum entropy spectrum [6]: Assume feasibility of Problem 11.2.1. Then, the unique solution of Problem 11.4.1 is given by

$$\hat{\Phi} = \left[G^* \Sigma^{-1} B \left(B^* \Sigma^{-1} B \right)^{-1} B^* \Sigma^{-1} G \right]^{-1}. \qquad (11.15)$$

Unfortunately, it appears quite problematic generalizing this result to the case of a general $\Psi \in \mathscr{S}_m$. Indeed, as pointed out in [14], the variational analysis cannot be carried through. To overcome this difficulty, in [8] a new metric was introduced that is induced by a sensible generalization to the multivariable case of the *Hellinger distance*. In [8, 23], the problem of computing the spectral density Φ minimizing this generalized Hellinger distance from a prior Ψ, under constraints (11.2), has been analyzed and it has been shown that the solution is still a rational function with an *a priori* bound on its McMillan degree. A different strategy is connected to *homotopy methods* that are described in [14] to find a spectrum that satisfies the constraint when such a family in nonempty.

In this paper, in the spirit of [14, Section IV], and motivated by the scalar case and $\Psi = I$ results, we start by introducing explicitly a parametric family of spectra $\Phi_\Lambda, \Lambda \in \mathscr{L}_+$ in which to look for a solution of Problem 11.2.1. In order to do that, we need first the following result:

Lemma 11.4.1. *Let* $G(z) = (zI - A)^{-1}B$ *with* $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times m}$, *and let* (A, B) *be a reachable pair. Let* $\Lambda \in \mathscr{L}_+$. *Then, the algebraic Riccati equation*

$$\Pi = A^* \Pi A - A^* \Pi B (B^* \Pi B)^{-1} B^* \Pi A + \Lambda, \tag{11.16}$$

admits a unique stabilizing solution $P \in \mathfrak{H}_n$. *The corresponding matrix* $B^* PB$ *is positive definite and the spectrum of closed loop matrix*

$$Z := A - B(B^* PB)^{-1} B^* PA \tag{11.17}$$

lays in the open unit disk. Let L *be the unique (lower triangular) right Choleski factor of* $B^* PB$ *(so that* $B^* PB = L^* L$).

The following factorization holds:

$$G^* \Lambda G = W_\Lambda^* W_\Lambda, \tag{11.18}$$

where

$$W_\Lambda(z) := L^{-*} B^* PA (zI - A)^{-1} B + L. \tag{11.19}$$

The rational function $W_\Lambda(z)$ *is the unique stable and minimum phase right spectral factor of* $G^* \Lambda G$, *such that* $W_\Lambda(\infty)$ *is lower triangular and with positive entries in the main diagonal.*

We are now ready to introduce our class of multivariate spectral density functions:

$$\Phi_\Lambda := W_\Lambda^{-1} \Psi W_\Lambda^{-*}, \quad \Lambda \in \mathscr{L}_+. \tag{11.20}$$

Notice that the optimal Kullback-Leibler approximant in the scalar case (11.11) and in the multivariate, $\Psi = I$ case (11.15) do belong to this class. This class, however, is different from the one proposed in [14, Section IV]. Although the latter is fully justified by general geometric considerations (Krein-Nudelmann theory [18]), our class is more suitable for implementation of the following matricial version of the efficient algorithm (11.13):

$$\Lambda_{k+1} = \Theta(\Lambda_k) := \int \Lambda_k^{1/2} G \left[W_{\Lambda_k}^{-1} \Psi W_{\Lambda_k}^{-*} \right] G^* \Lambda_k^{1/2}, \qquad \Lambda_0 = \frac{1}{n} I. \qquad (11.21)$$

It is easy to see that this map preserves trace and positivity as in the scalar case. We have performed a limited number of simulations in this general setting. In all these simulations, the sequence Λ_k converges very fast to a matrix $\hat{\Lambda}$, for which the corresponding spectral density (given by (11.20)) solves Problem 11.2.1.

Before addressing the computational aspects of the problem, we need first to investigate the following question:

Problem 11.4.2. Let $\Sigma \in \mathrm{Range}_+ \Gamma := \mathrm{Range}\, \Gamma \cap \mathfrak{H}_{+,n}$. Let $G(z) = (zI - A)^{-1} B$ with the same properties as in Problem 11.2.1, and let $\Psi \in \mathcal{S}_m$. Find $\Lambda \in \mathcal{L}_+$ such that Φ_Λ given by (11.20) satisfies

$$\int G \Phi_\Lambda G^* = \Sigma. \qquad (11.22)$$

Most of this paper is devoted to this question. In particular we show that in the case when $\Psi(z) = \psi(z) Q$ with $\psi(z) \in C_+(\mathbb{T})$ and Q is a constant positive definite matrix, Problem 11.4.2 is feasible. To this aim we need some preliminary results. Consider the map $\omega : \mathcal{L}_+^\Gamma \longrightarrow \mathrm{Range}_+ \Gamma$ given by

$$\omega : \Lambda \mapsto \int G \Phi_\Lambda G^*. \qquad (11.23)$$

Notice that ω is a continuos map between open subsets of the linear space $\mathrm{Range}\, \Gamma$. It is clear that Problem 11.4.2 is feasible if and only if the map ω is surjective. We are now precisely in the setting of Theorem 2.6 in [3]. It states that if ω is *proper* and injective than it is surjective. We first show that ω is *proper*, i.e. the preimage of every compact set in $\mathrm{Range}_+ \Gamma$ is compact in $\mathcal{L}_+^\Gamma$. For this purpose, we need the following lemma.

Lemma 11.4.2. If $G^* \Lambda G > 0$, $\forall e^{j\vartheta} \in \mathbb{T}$, then there exists $\Lambda_+ \in \mathfrak{H}_{+,n}$ such that $G^* \Lambda G = G^* \Lambda_+ G$.

Proof. As shown in Lemma 11.4.1, we can perform the factorization

$$G^* \Lambda G = W_\Lambda^* W_\Lambda, \qquad (11.24)$$

where the (right) spectral factor $W_\Lambda(z)$ is given by (11.19). The spectral factor $W_\Lambda(z)$ may be easily rewritten as

$$W_\Lambda = L^{-*} \left[B^* P A (zI - A)^{-1} B + B^* P B \right] = L^{-*} B^* P \left[A(zI - A)^{-1} + I \right] B. \qquad (11.25)$$

It is immediate to check that $A(zI - A)^{-1} + I = z(zI - A)^{-1}$ so that

$$W_\Lambda = z L^{-*} B^* P (zI - A)^{-1} B. \qquad (11.26)$$

and thus

$$G^* \Lambda G = W_\Lambda^* W_\Lambda = W_{\Lambda_1}^* W_{\Lambda_1}, \qquad (11.27)$$

with

$$W_{\Lambda_1} := z^{-1}W_\Lambda = L^{-*}B^*P(zI - A)^{-1}B. \tag{11.28}$$

So there exists a matrix $C_\circ = (L^{-*}B^*P)^* \in \mathbb{C}^{n \times m}$ such that

$$G^*\Lambda G = G^*C_\circ C_\circ^* G. \tag{11.29}$$

We observe that on the unit circle $\mathbb{T}$, $G^*\Lambda G$ is continuous and positive definite so that there exists a positive constant μ such that

$$G(z)^*\Lambda G(z) > \mu I, \quad \forall z \in \mathbb{T}.$$

Similarly, on the unit circle $\mathbb{T}$, G^*G is continuous and hence there exists a positive constant ν such that

$$G(z)^*G(z) < \nu I, \quad \forall z \in \mathbb{T}.$$

Let $\varepsilon := \frac{\mu}{4\nu}$. Now let $\Lambda_1 := \frac{1}{2}\Lambda - \varepsilon I$. Clearly, $\forall z \in \mathbb{T}$, we have

$$G(z)^*\Lambda_1 G(z) = \frac{1}{2}G(z)^*\Lambda G(z) - \varepsilon G(z)^*G(z) \geq \left(\frac{\mu}{2} - \frac{\mu}{4\nu}\nu\right)I = \frac{\mu}{4}I > 0. \tag{11.30}$$

Hence, by resorting to the same argument that led to (11.29), we conclude that there exists $C_1 \in \mathbb{C}^{n \times m}$ such that

$$G^*(\frac{1}{2}\Lambda - \varepsilon I)G = G^*C_1 C_1^* G.$$

Therefore we have

$$
\begin{aligned}
G^*\Lambda G &= \frac{1}{2}G^*C_\circ C_\circ^* G + \frac{1}{2}G^*C_\circ C_\circ^* G + \varepsilon G^*G - \varepsilon G^*G \\
&= G^*(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I)G + \frac{1}{2}G^*C_\circ C_\circ^* G - \varepsilon G^*G \\
&= G^*(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I)G + \frac{1}{2}G^*\Lambda G - \varepsilon G^*G \\
&= G^*(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I)G + G^*(\frac{1}{2}\Lambda - \varepsilon I)G \\
&= G^*(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I)G + G^*C_1 C_1^* G \\
&= G^*(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I + C_1 C_1^*)G = G^*\Lambda_+ G,
\end{aligned}
$$

where $\Lambda_+ := \frac{1}{2}C_\circ C_\circ^* + \varepsilon I + C_1 C_1^*$ is clearly positive definite. $\square$

Theorem 11.4.1. *The map ω is proper.*

Proof. We observe that $\mathscr{L}_+^\Gamma$ and $\text{Range}_+\Gamma$ are subsets of a finite dimensional linear space so that compact sets in $\mathscr{L}_+^\Gamma$ and $\text{Range}_+\Gamma$ are characterized by being closed and bounded. Accordingly, to prove the statement is sufficient to show that $\omega^{-1}(K)$ is closed and bounded for any compact set K. To see that $\omega^{-1}(K)$ is bounded we

choose an arbitrary sequence $\{\Lambda_n\}$ such that $\Lambda_n \in \mathcal{L}_+^\Gamma$, $\|\Lambda_n\| \to \infty$ and we show that the minimum eigenvalue of $\omega(\Lambda_n)$ approaches zero as n tends to infinity. This means that, as n tends to infinity, $\omega(\Lambda_n)$ approaches the boundary of the co-domain $\text{Range}_+ \Gamma$ which is a subset of the positive definite matrices. Therefore, there does not exist a compact set K in $\text{Range}_+ \Gamma$ such that $\omega^{-1}(K)$ contains the sequence Λ_n. Similarly, to see that $\omega^{-1}(K)$ is closed we choose an arbitrary sequence $\Lambda_n \in \mathcal{L}_+^\Gamma$ approaching the boundary of $\mathcal{L}_+$, and we show that there does not exist a compact set K in $\text{Range}_+ \Gamma$ such that $\omega^{-1}(K)$ contains the sequence Λ_n. The proof, which is detailed only for the case $\|\Lambda_n\| \to \infty$, will be divided in four steps.

Step 1: Observing that $\Psi(z)$ is bounded (i.e. $\exists\, m : \Psi \leq mI$), we have

$$0 \leq \omega(\Lambda) = \int GW_\Lambda^{-1}\Psi W_\Lambda^{-*}G^*$$
$$\leq m \int GW_\Lambda^{-1}W_\Lambda^{-*}G^* = m \int G(G^*\Lambda G)^{-1}G^*. \tag{11.31}$$

It is therefore sufficient to consider the map

$$\begin{array}{rcl} \tilde{\omega} : \mathcal{L}_+^\Gamma & \longrightarrow & \text{Range}_+ \Gamma \\ \Lambda & \mapsto & \int G(G^*\Lambda G)^{-1}G^* \end{array} \tag{11.32}$$

and to show that the minimum eigenvalue of $\tilde{\omega}(\Lambda_n)$ approaches zero.

Step 2: By (11.8), $(\text{Range}\,\Gamma)^\perp = \ker\Gamma^*$. Hence, the minimum singular value ρ of the map Γ^* restricted to $\text{Range}\,\Gamma$ is strictly positive. Accordingly, since $\text{Range}_+ \Gamma \subset \text{Range}\Gamma$, we have

$$\|G^*\Lambda_n G\| \geq \rho\|\Lambda_n\| \to \infty. \tag{11.33}$$

Step 3: By Lemma 11.4.2, we know that there exists $\Lambda_{n+} = \Lambda_{n+}^* > 0$ such that

$$G^*\Lambda_{n+}G = G^*\Lambda_n G, \quad \forall\, n. \tag{11.34}$$

We have $\|\Lambda_{n+}\| \to \infty$. In fact, let μ_n be the maximum eigenvalue of Λ_{n+}, so that $\Lambda_{n+} < \mu_n I$. It follows that

$$\mu_n\|G^*G\| \geq \|G^*\Lambda_{n+}G\| = \|G^*\Lambda_n G\| \longrightarrow +\infty. \tag{11.35}$$

Since $\|G^*G\| > 0$, the latter implies $\mu_n \to +\infty$ and hence $\|\Lambda_{n+}\| \to \infty$.

Step 4: By Lemma 11.4.2 and recalling that $\Pi \leq I$ for any orthogonal projection matrix Π, we have

$$\tilde{\omega}(\Lambda_n) = \int G(G^*\Lambda_n G)^{-1}G^*$$

$$= \int G(G^*\Lambda_{n+}G)^{-1}G^*$$

$$= \Lambda_{n+}^{-1/2}\left[\int \Lambda_{n+}^{1/2}G(G^*\Lambda_{n+}G)^{-1}G^*\Lambda_{n+}^{1/2}\right]\Lambda_{n+}^{-1/2}$$

$$= \Lambda_{n+}^{-1/2}\left[\int \Pi_{\Lambda_{n+}^{1/2}G}\right]\Lambda_{n+}^{-1/2}$$

$$\leq \Lambda_{n+}^{-1}, \tag{11.36}$$

where we denote by $\Pi_{\Lambda_{n+}^{1/2}G}$ the orthogonal projection on $\Lambda_{n+}^{1/2}G$. Finally, as shown in Step 3, $\|\Lambda_{n+}\| \to \infty$ so that the minimum eigenvalue of Λ_{n+}^{-1} and, *a fortiori*, the minimum eigenvalue of $\tilde{\omega}(\Lambda_n)$, approaches zero.   □

As already mentioned, if the map ω were also injective, then we could conclude that ω is surjective and hence Problem (11.4.2) is feasible. As a preliminary result, we show injectivity in the case when $\Psi(z)$ is a scalar spectral density (i.e. $\Psi(z) = \psi(z)I_m$ with $\psi(z) \in C_+(\mathbb{T})$).

Theorem 11.4.2. *Let $\Psi(z)$ be a scalar spectral density. Then the map ω is injective and hence surjective.*

Proof. Let

$$\Lambda_1, \Lambda_2 \in \mathscr{L}_+^\Gamma \subset \operatorname{Range}\Gamma, \tag{11.37}$$

and assume that

$$\omega(\Lambda_1) - \omega(\Lambda_2) = 0. \tag{11.38}$$

Define

$$\Phi_1 := \psi W_{\Lambda_1}^{-1}W_{\Lambda_1}^{-*} = \psi(G^*\Lambda_1 G)^{-1}, \tag{11.39}$$

and

$$\Phi_2 := \psi W_{\Lambda_2}^{-1}W_{\Lambda_2}^{-*} = \psi(G^*\Lambda_2 G)^{-1}. \tag{11.40}$$

Thus,

$$0 = \omega(\Lambda_1) - \omega(\Lambda_2) = \Gamma(\Phi_1) - \Gamma(\Phi_2) = \Gamma(\Phi_1 - \Phi_2) \tag{11.41}$$

so that $(\Phi_1 - \Phi_2) \in \ker\Gamma$.

The adjoint transform of Γ is easily seen to be given by

$$\begin{aligned}\Gamma^* : \mathfrak{H}_n &\longrightarrow C(\mathbb{T},\mathfrak{H}_m)\\ M &\mapsto G^*MG.\end{aligned} \tag{11.42}$$

Thus, condition $(\Phi_1 - \Phi_2) \in \ker\Gamma = (\operatorname{Range}\Gamma^*)^\perp$ reads

$$\langle G^*MG, \Phi_1 - \Phi_2\rangle = \operatorname{tr}\int G^*MG(\Phi_1 - \Phi_2) = 0, \quad \forall M \in \mathfrak{H}_n. \tag{11.43}$$

In particular, by choosing $M = \Lambda_2 - \Lambda_1$, we get

$$0 = \operatorname{tr} \int \left[G^*(\Lambda_2 - \Lambda_1)G \right] (\Phi_1 - \Phi_2)$$

$$= \operatorname{tr} \int \left[G^*(\Lambda_2 - \Lambda_1)G \right] \psi \left[(G^*\Lambda_1 G)^{-1} - (G^*\Lambda_2 G)^{-1} \right]$$

$$= \operatorname{tr} \int \psi \left[G^*(\Lambda_2 - \Lambda_1)G \right] (G^*\Lambda_1 G)^{-1} \left[G^*\Lambda_2 G - G^*\Lambda_1 G \right] (G^*\Lambda_2 G)^{-1}$$

$$= \operatorname{tr} \int \psi \left[G^*(\Lambda_2 - \Lambda_1)G \right] (G^*\Lambda_1 G)^{-1} \left[G^*(\Lambda_2 - \Lambda_1)G \right] W_{\Lambda_2}^{-1} W_{\Lambda_2}^{-*}$$

$$= \operatorname{tr} \int \psi W_{\Lambda_2}^{-*} \left[G^*(\Lambda_2 - \Lambda_1)G \right] (G^*\Lambda_1 G)^{-1} \left[G^*(\Lambda_2 - \Lambda_1)G \right] W_{\Lambda_2}^{-1}. \quad (11.44)$$

Since $\psi \in C_+(\mathbb{T})$, and $(G^*\Lambda_1 G)^{-1}$ is positive definite on $\mathbb{T}$, the integrand function is positive semi-definite. Therefore, (11.44) implies

$$\left[G^*(\Lambda_2 - \Lambda_1)G \right] (G^*\Lambda_1 G)^{-1} \left[G^*(\Lambda_2 - \Lambda_1)G \right] \equiv 0, \quad (11.45)$$

that, in turn, yields

$$G^*(\Lambda_2 - \Lambda_1)G \equiv 0. \quad (11.46)$$

By (11.8), $\Lambda_2 - \Lambda_1 \in (\operatorname{Range}\Gamma)^\perp$. The latter, together with (11.37), yields

$$\Lambda_2 - \Lambda_1 \in \operatorname{Range}\Gamma \cap (\operatorname{Range}\Gamma)^\perp = \{0\}, \quad (11.47)$$

so that $\Lambda_1 = \Lambda_2$.   □.

We are now ready to prove our main result

Theorem 11.4.3. *Let* $\Psi(z) = \psi(z)Q$ *with* $\psi(z) \in C_+(\mathbb{T})$ *and* $Q \in \mathfrak{H}_{+,n}$. *Then the map* ω *is surjective.*

Proof. We first observe that, since B is assumed to be full column rank, we may perform a change of basis and assume, without loss of generality, that $B = \begin{bmatrix} I \\ 0 \end{bmatrix}$. Secondly, notice that, it is sufficient to extend the domain of ω to the whole set $\mathscr{L}_+$ and prove the result for the map with extended domain. In fact, if $\omega(\Lambda) = \Sigma$ for a certain $\Lambda \in \mathscr{L}_+$, and $\Lambda_\Gamma \in \mathscr{L}_+^\Gamma$ is the orthogonal projection of Λ in $\operatorname{Range}\Gamma$, then also $\omega(\Lambda_\Gamma) = \Sigma$. Next, we need to compute GW_Λ^{-1}. We observe that:

$$W_\Lambda^{-1} = L^{-1} - (B^*PB)^{-1}B^*PA(zI - Z)^{-1}BL^{-1}, \quad (11.48)$$

where Z, defined in (11.17), is a stability matrix. Hence,

$$GW_\Lambda^{-1} = -(zI - A)^{-1}B(B^*PB)^{-1}B^*PA(zI - Z)^{-1}BL^{-1} + (zI - A)^{-1}BL^{-1} \quad (11.49)$$

Notice that

$$B(B^*PB)^{-1}B^*PA = A - Z = (zI - Z) - (zI - A).$$

Plugging this expression into (11.49) we get

$$GW_\Lambda^{-1} = (zI - Z)^{-1}BL^{-1} = (zI - Z)^{-1} \begin{bmatrix} L^{-1} \\ 0 \end{bmatrix}, \quad (11.50)$$

where we have used the fact that $B = \begin{bmatrix} I \\ 0 \end{bmatrix}$. We now partition P conformably with B as

$$P = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^* & P_2 \end{bmatrix}. \tag{11.51}$$

Then we immediately see that $B^*PB = P_1$ so that $L^{-1} = L_{P_1^{-1}}$ is the Choleski factor[2] of P_1^{-1}. Moreover, the matrix Z has the following expression

$$Z = \begin{bmatrix} 0 & -P_1^{-1}P_{12} \\ 0 & I \end{bmatrix} A. \tag{11.52}$$

Consider now an arbitrary $\bar{\Sigma} \in \mathrm{Range}_+ \Gamma$ and let $\bar{\Psi}(z) = \psi(z)I$. In wiew of Theorem 11.4.2, the map $\bar{\omega} : \Lambda \mapsto \int G W_{\Lambda}^{-1} \bar{\Psi} W_{\Lambda}^{-*} G^*$ is surjective. Hence there exists $\bar{\Lambda} \in \mathscr{L}_+$ such that $\bar{\omega}(\Lambda) = \bar{\Sigma}$. Let $\bar{P}$ be the corresponding stabilizing solution of the ARE (11.16) and $\bar{Z}$ be the associated closed-loop matrix whose spectrum is contained in the open unit disk.

We are now ready to address the case when $\Psi(z) = \psi(z)Q$: Define

$$\tilde{P}_1 := (L_{\bar{P}_1^{-1}} L_Q^{-1} L_Q^{-*} L_{\bar{P}_1^{-1}}^*)^{-1}, \quad \tilde{P}_{12} := \tilde{P}_1 \bar{P}_1^{-1} \bar{P}_{12}, \quad \tilde{P}_2 := \bar{P}_2, \tag{11.53}$$

and let $\tilde{P}$ be the corresponding 2×2 block matrix. Moreover, let

$$\tilde{\Lambda} := \tilde{P} - (A^* \tilde{P} A - A^* \tilde{P} B (B^* \tilde{P} B)^{-1} B^* \tilde{P} A) \tag{11.54}$$

We have the following facts:

1. If $\Lambda = \tilde{\Lambda}$ then $\tilde{P}$ is, by construction, solution of the ARE (11.16).
2. The corresponding closed-loop matrix $\tilde{Z}$ is immediately seen to be equal to $\bar{Z}$ whose spectrum is contained in the open unit disk.
3. Since $\tilde{P}_1 = B^* \tilde{P} B$ is, by construction positive definite and $\tilde{Z}$ is a stability matrix, we can associate to $\tilde{P}$ a spectral factorization of $G^* \tilde{\Lambda} G$ of the form (11.18) so that $G^* \tilde{\Lambda} G$ is positive definite on $\mathbb{T}$ or, equivalently, $\tilde{\Lambda} \in \mathscr{L}_+$.
4. Since a product of Choleski factors and the inverse of a Choleski factor are Choleski factors, and taking into account that the Choleski factor is unique, from the definition of $\tilde{P}_1$, we get $L_{\tilde{P}_1^{-1}} = L_{\bar{P}_1^{-1}} L_Q^{-1}$.
5. As a consequence of the previous observation we get

$$G W_{\tilde{\Lambda}}^{-1} L_Q = G W_{\bar{\Lambda}}^{-1} \tag{11.55}$$

In conclusion, $\tilde{\Lambda} \in \mathscr{L}_+$ and, as it follows immediately from (11.55),

$$\omega(\tilde{\Lambda}) = \bar{\omega}(\bar{\Lambda}) = \bar{\Sigma} \tag{11.56}$$

which concludes the proof. $\square$

---

[2]We denote by L_Ξ the lower triangular left Choleski factor of a positive definite matrix Ξ, i.e. the unique lower triangular matrix having positive entries in the main diagonal and such that $\Xi = LL^*$.

References

1. Byrnes, C.I., Gusev, S., Lindquist, A.: A convex optimization approach to the rational covariance extension problem. SIAM J. Control and Opimization 37, 211–229 (1999)
2. Byrnes, C.I., Gusev, S., Lindquist, A.: From finite covariance windows to modeling filters: A convex optimization approach. SIAM Review 43, 645–675 (2001)
3. Byrnes, C.I., Linquist, A.: Interior point solutions of variational problems and global inverse function theorems. International Journal of Robust and Nonlinear Control 17, 463–481 (2007)
4. Byrnes, C.I., Linquist, A.: Important moments in systems and control. SIAM J. Control and Optimization 47(5), 2458–2469 (2008)
5. Byrnes, C.I., Linquist, A.: A convex optimization approach to generalized moment problems. In: Control and Modeling of Complex Systems: Cybernetics in the 21st Century, pp. 3–21. Birkhäuser, Boston (2003)
6. Cover, T.M., Thomas, J.A.: Information Theory. Wiley, New York (1991)
7. Ferrante, A., Pavon, M., Ramponi, F.: Further results on the Byrnes-Georgiou-Lindquist generalized moment problem. In: Chiuso, A., Ferrante, A., Pinzoni, S. (eds.) Modeling, Estimation and Control: Festschrift in honor of Giorgio Picci on the occasion of his sixty-fifth birthday, pp. 73–83. Springer, Heidelberg (2007)
8. Ferrante, A., Pavon, M., Ramponi, F.: Hellinger vs. Kullback-Leibler multivariable spectrum approximation. IEEE Trans. Aut. Control 53, 954–967 (2008)
9. Ferrante, A., Ramponi, F., Ticozzi, F.: On the convergence of an efficient algorithm for Kullback-Leibler approximation of spectral densities. IEEE Trans. Aut. Control, Submitted for publication (2009)
10. Georgiou, T.: Realization of power spectra from partial covariance sequences. IEEE Trans. on Acoustics, Speech, and Signal Processing 35, 438–449 (1987)
11. Georgiou, T.: The interpolation problem with a degree constraint. IEEE Trans. Aut. Control 44, 631–635 (1999)
12. Georgiou, T.: Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parameterization. IEEE Trans. Aut. Control 47, 1811–1823 (2002)
13. Georgiou, T.: The structure of state covariances and its relation to the power spectrum of the input. IEEE Trans. Aut. Control 47, 1056–1066 (2002)
14. Georgiou, T.: Relative entropy and the multivariable multidimensional moment problem. IEEE Trans. Inform. Theory 52, 1052–1066 (2006)
15. Georgiou, T., Lindquist, A.: Kullback-Leibler approximation of spectral density functions. IEEE Trans. Inform. Theory 49, 2910–2917 (2003)
16. Georgiou, T., Lindquist, A.: A convex optimization approach to ARMA modeling. IEEE Trans. Aut. Control 53, 1108–1119 (2008)
17. Grenander, U., Szegö, G.: Toeplitz Forms and Their Applications. University of California Press, Berkeley (1958)
18. M. G. Kreĭn and A. A. Nudel'man. *The Markov Moment Problem and Extremal Problems*. Amer. Math. Soc., Providence, RI, 1977.
19. Kullback, S.: Information Theory and Statistics, 2nd edn. Dover, Mineola (1968)
20. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge Univ. Press, Cambridge (2000)
21. Pavon, M., Ferrante, A.: On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities. IEEE Trans. Aut. Control 51, 639–644 (2006)

22. Ramponi, F., Ferrante, A., Pavon, M.: On the well-posedness of multivariate spectrum approximation and convergence of high- resolution spectral estimators. Systems and Control Letters, to appear (March 2009)
23. Ramponi, F., Ferrante, A., Pavon, M.: A globally convergent matricial algorithm for multivariate spectral estimation. IEEE Trans. Aut. Control 54, 2376–2388 (2009)