Matrix Measures, Moment Spaces and Favard's Theorem for the interval $[0, 1]$ and $[0, \infty)$

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 70 94 559

William J. Studden
Purdue University
Department of Statistics
1399 Mathematical Sci Bldg
West Lafayette, IN, 47907-1399
USA
email: studden@stat.purdue.edu

February 27, 2001

Abstract

In this paper we study the moment spaces corresponding to matrix measures on compact intervals and on the nonnegative line $[0, \infty)$. A representation for nonnegative definite matrix polynomials is obtained, which is used to characterize moment points by properties of generalized Hankel matrices. We also derive an explicit representation of the orthogonal polynomials with respect to a given matrix measure, which genereralize the classical determinant representations of the one dimensional case. Moreover, the coefficients in the recurrence relations can be expressed explicitly in terms of the moments of the matrix measure. These results are finally used to prove a refinement of the well known Favard theorem for matrix measures, which characterizes the domain of the underlying measure of orthogonality by properties of the coefficients in the recurrence relationships.

AMS Subject Classification: 42C05
Keywords and Phrases: Matrix measure, nonegative definite matrix polynomials, orthogonal polynomials, canonical moments of matrix measures, Favard's theorem

1 Introduction

Moment problems, orthogonal polynomials, continued fractions, quadrature formulas and approximation theory, etc., have been studied for a long time and have a vast literature. In recent years considerable interest has been shown in generalizing many of the results in these areas to the case of matrix polynomials and matrix measures. Among many others we refer to the early work of

Krein (1949) and to the more recent papers of Aptekarev and Nikishin (1983), Geronimo (1982), Rodman (1990), Sinap and Van Assche (1994), Duran and Van Assche (1995), Duran (1995, 1996, 1999) and Duran and Lopez-Rodriguez (1996, 1997). A $p \times p$ matrix polynomial is a $p \times p$ matrix with polynomial entries. It is of degree n if all the polynomials are of degree less or equal than n and is usually written in the form

$$P(t) = \sum_{i=0}^{n} A_i t^i. \tag{1.1}$$

where the A_i are real $p \times p$ matrices. The matrix polynomial $P(t)$ is called monic if the highest coefficient satisfies $A_n = I_p$ where I_p denotes the $p \times p$ identity matrix. A matrix measure μ is a $p \times p$ matrix $\mu = \{\mu_{ij}\}$ of finite signed measures μ_{ij} on the Borel field of the real line $\mathbb{R}$ or of an appropriate subset. It will be assumed here that for each Borel set A the matrix $\mu(A) = \{\mu_{ij}(A)\}$ is symmetric and nonnegative definite, i.e. $\mu(A) \geq 0$. The moments of the matrix measure μ are given by the $p \times p$ matrices

$$S_k = \int t^k d\mu(t) \qquad k = 0, 1, \cdots \tag{1.2}$$

Only measures for which all relevant moments exist will be considered throughout this paper. The integrals will usually be over the interval $[0, 1]$ but integrals over the half line $[0, \infty)$ will also be considered in Section 5. The $(n + 1)$th moment space is given by

$$M_{n+1} = \{ \ (S_0, S_1, \cdots, S_n) \ \mid \ \mu \ \} \tag{1.3}$$

where μ ranges over the set of all matrix measures with existing moments up to the order n.

The purpose of the present paper is to investigate properties of the moment space, mainly with the purpose of providing some generalizations of Favard's Theorem in the matrix case for the nonnegative line $[0, \infty)$ and the compact interval $[0, 1]$. It is well known in the scalar case where $p = 1$, that if a sequence of polynomials $\{P_n\}_{n \geq 0}$ (where P_n is of exact degree n) is orthogonal with respect to some measure μ on the real line, then it satisfies a three term recurrence relation of the form

$$P_{n+1}(x) = (x - \alpha_{n+1})P_n(x) - \beta_{n+1}P_{n-1}(x) \ , \quad n \ \geq \ 1 \tag{1.4}$$

with $\beta_{n+1} > 0$ [see Favard (1935)]. Conversely, if the sequence of polynomials satisfies (1.4) with $\beta_{n+1} > 0$ for all $n \in \mathbb{N}$, then there exists a measure on the real line for which the polynomials are orthogonal. It is also well known that the measure μ concentrates on the nonnegative axis $[0, \infty)$ if and only if there exists a sequence $\{\zeta_k\}_{k \geq 1}$ of positive numbers such that the coefficients in the recurrence relation (1.4) satisfy for all $k \geq 1$

$$\beta_{k+1} = \zeta_{2k-1}\zeta_{2k} \quad \text{and} \qquad \alpha_{k+1} = \zeta_{2k} + \zeta_{2k+1}. \tag{1.5}$$

[see e.g. Chihara (1978), p. 47]. It was further discovered by Wall (1940) that μ is concentrated on the interval $[0, 1]$ if and only if the coefficients ζ_k form a chain sequence; that is, they can further be decomposed as

$$\zeta_k = q_{k-1}p_k \tag{1.6}$$

where $q_k = 1 - p_k$ ($q_0 = 1$) and $0 < p_k < 1$. The sequence of constants were called canonical moments by Skibinsky (1967, 1969, 1986) and are discussed fully in the monograph Dette and Studden (1997).

Essentially complete analogs of these results will be presented below. For the whole line (and the circle) these results are known and are discussed, for example, in Sinap and Van Assche (1996). In Section 2-4, the moments spaces M_{n+1} for measures on the interval $[0,1]$ are investigated and matrix analogs of the canonical moments are introduced. One of the main theorems in Section 2 is a representation theorem for nonnegative definite matrix polynomials, which is used to characterize the points in the moment space by properties of generalized Hankel matrices. We also present a result describing the 'width' of the moment space in terms of matrix valued canonical moments. In Section 3 explicit formulas for the orthogonal polynomials on the interval $[0,1]$ in the matrix case are given, which generalize the well known determinant representation for orthogonal polynomials in the case $p = 1$ [see e.g. Szegö (1959), p. 27]. These results are used for a discussion of the recurrence formula in more detail. Section 4 contains the generalized Favard theorem for the interval $[0,1]$. Finally, the Favard theorem for the half line, with some discussion of the corresponding moment space, is given in Section 5.

2 Moment Spaces

As indicated above, the discussion in sections 2-4 is confined to the interval $[0,1]$. The moments S_k and the moment space M_{n+1} are defined as in (1.2) and (1.3), respectively. The set M_{n+1} can be viewed as an Euclidean space of dimension $(n+1)p(p+1)/2$. Hyperplanes in this space may be assumed to be of the form

(2.1) $$\sum_{i=0}^{n} tr(A_i S_i) = b$$

where $b \in \mathbb{R}$ and the $A_0, \ldots, A_n$ are real symmetric $p \times p$ matrices. Consider the set

(2.2) $$C_{n+1} = \{ \ (aa^T, taa^T, \cdots, t^n aa^T) \mid 0 \le t \le 1, \quad a \ \in \mathbb{R}^p \ \}$$

and let $\mathcal{C}(C_{n+1})$ denote the convex cone generated by this set. Without loss of generality we can view elements of this set as being of the form

(2.3) $$\sum_i \sum_j t_i^k d_{ij} a_{ij} a_{ij}^T$$

where the d_{ij} are positive and $a_{ij} \ \in \mathbb{R}^p$. Note that there is some redundency here in that we could normalize the vectors $a_{ij} \in \mathbb{R}^p$ to have length one or leave out the factors d_{ij}.

Lemma 2.1: *The set M_{n+1} defined in (1.3) is equal to the convex cone $\mathcal{C}(C_{n+1})$.*

Proof: Obviously $\mathcal{C}(C_{n+1}) \subset M_{n+1}$ and for the converse inclusion the question amounts to asking whether all the points in the moment space M_{n+1} are limits of points in the convex cone generated above. Because the generated cone is closed, the assertion then follows. If μ is composed of a matrix of smooth densities, say f, with respect to the Lebesque measure and the integrals are viewed as Riemman integrals, then one can approximate the kth moment of μ by

$$\sum t_i^k f(t_i), \tag{2.4}$$

where each $f(t_i)$ is a nonnegative definite $p \times p$ matrix. In this case we can write for each i

$$f(t_i) = \sum_j d_{ij} a_{ij} a_{ij}^T \tag{2.5}$$

and the result follows. For the general case we use the fact that a finite number of moments of any signed measure μ_{ij} can be approximated by one with a density of the above form.

$\square$

Corollary 2.2:
Let $A_0, \ldots, A_n$ denote symmetric $p \times p$ matrices, then
a)

$$\sum_{k=0}^{n} t^k A_k \geq 0 \quad \textit{for all} \quad t \in [0,1] \quad \Leftrightarrow \quad \sum_{k=0}^{n} tr A_k S_k \geq 0 \quad \textit{for all} \quad S = (S_0, \ldots, S_n) \in M_{n+1}.$$

b) Every point $S = (S_0, \ldots, S_n)$ in the moment space M_{n+1} has a finite representation of the form

$$S_k = \sum_{i=1}^{q} d_i a_i a_i^T t_i^k \quad k = 0, \ldots, n$$

where the number of terms in these representations is bounded by

$$q \leq (n+1)\frac{p(p+1)}{2}.$$

c) Every point in the interior of the moment space M_{n+1} has a representation of the above form using any specific pair $(t, a) \in [0,1] \times \mathbb{R}^p$.

Proof: The proof for part a) follows by noting that

$$\begin{aligned}
\sum_{k=0}^{n} t^k A_k \geq 0 \text{ for all } t \in [0,1] &\Longleftrightarrow \sum_{k=0}^{n} t^k a^T A_k a \geq 0 \qquad \text{for all } a \in \mathbb{R}^p \text{ and } t \in [0,1] \\
&\Longleftrightarrow \sum_{k=0}^{n} tr(A_k t^k a a^T) \geq 0 \qquad \text{for all } a \in \mathbb{R}^p \text{ and } \ t \in [0,1] \\
&\Longleftrightarrow \sum_{k=0}^{n} tr(A_k S_k) \geq 0 \qquad \text{for all } (S_0, \cdots, S_n) \ \in \ M_{n+1},
\end{aligned}$$

where trB denotes the trace of the matrix B and the last equivalence follows from Lemma 2.1. Part b) follows from the Caratheodory theorem [see e.g. Rockafellar (1970)].
For part c) we take the line joining the point corresponding to (t, a) and the interior point and extend it past the interior point to the boundary. Therefore we are able to write the interior point as a bonafide convex combination of the point corresponding to (t, a) and some other point.

$\square$

Lemma 2.3: *Let* $S = (S_0, \ldots, S_n) \in M_{n+1}$, *then*

a) $S \in M_{n+1}$ *if and only if*

$$\sum_{k=0}^{n} t^k A_k \geq 0 \quad \textit{for all} \ \ t \in [0,1] \Longrightarrow \ \sum_{k=0}^{n} trS_k A_k \geq 0 \tag{2.6}$$

b) $S \in Int(M_{n+1})$ *if and only if*

$$\begin{array}{c} P(t) = \sum_{k=0}^{n} t^k A_k \geq 0 \quad \text{for all} \ \ t \ \in [0,1] \ \ \text{and} \\ P(t) \ \ \text{is not the zero matrix for all} \ \ t \in [0,1] \end{array} \quad \Longrightarrow \ \sum_{k=0}^{n} trS_k A_k > 0 \tag{2.7}$$

Proof: For part a) the necessity follows directly from Corollary 2.2. We now show sufficiency proving that $S^o \notin M_{n+1}$ implies the converse of (2.6). To this end assume that $S^o = (S_0^o, \ldots, S_n^o)$ is not in the moment space M_{n+1}. By Corollary 2.2 there exists matrices $A_0, \cdots, A_n$ (symmetric) with

$$\sum_{k=0}^{n} trA_k S_k^o < 0$$

and

$$\sum_{k=0}^{n} trA_k S_k \geq 0 \quad \text{for all} \quad S = (S_0, \ldots, S_n) \in M_{n+1}.$$

These latter inequalities include

$$\sum_{k=0}^{n} trA_k t^k aa^T \geq 0 \ \ \forall \ \ a \in \mathbb{R}^p, t \in [0,1].$$

Thus we have

$$\sum_{k=0}^{n} t^k A_k \geq 0 \ \ \forall \ \ t \in [0,1] \quad \text{but} \quad \sum_{k=0}^{n} trA_k S_k^o < 0$$

or (2.6) is not true.
To prove part b) we first suppose that $S = (S_0, \ldots, S_n)$ is in the interior of moment space M_{n+1} and $P(t)$ is not identically zero for all t. Then for some t, there is a vector $a \in \mathbb{R}^p$ such that $a^T P(t) a$ is positive. We now take the representation of S from part c) of Corollary 2.2 which involves t and a to show that $\sum_{k=0}^{n} trS_k A_k > 0$. For the converse we suppose that S is not in the interior of M_{n+1}. In this case there exists a nontrivial polynomial $P(t) = \sum_{k=0}^{n} t^k A_k \geq 0$ such that $\sum_{k=0}^{n} trA_k S_k \leq 0$. $\square$

Remark 2.4: When normalizing the coefficients in the representations it seems like a good choice is to take the points $S = (S_0, \ldots, S_n) \in M_{n+1}$ satisfying $trS_0 = 1$. Thus we let C^o_{n+1} denote the subset of C_{n+1} defined in (2.2) where $a^T a = 1$ and M^o_{n+1} be the subset of M_{n+1} where $trS_0 = 1$. It is fairly easy to show that M^o_{n+1} is the convex hull of C^o_{n+1}. To do this we note that the convex hull of C^o_{n+1} is contained in M^o_{n+1}. Further, if one is using pairs (d_i, a_i) in the representation (2.3) we can get the same point by replacing d_i by $\|a_i\|_2^2 d_i$ and a_i by $a_i/\|a_i\|_2$ where $\|\cdot\|_2$ is the euclidean norm of $a_i \in \mathbb{R}^p$.
Karlin-Studden (1966) discuss taking sections of the moment cones which are the intersection of the cone with some affine space to produce normalizations. These are generated by positive polynomials. The usual normalization in the scalar case takes the polynomial $P(t) = 1$ to give $c_0 = 1$ or the measures with norm one. If we take the matrix polynomial $P(t) = I$ we get the above normalization where $trS_0 = 1$.

The next theorem gives a representation for nonnegative definite $p \times p$ matrix polynomials

$$P_n(t) = \sum_{k=0}^{n} t^k A_k \quad \geq \quad 0$$

on the interval $[0, 1]$.

Theorem 2.5: *Assume that the matrix polynomial $P_n(t)$ is nonnegative definite for all t in the interval $[0, 1]$.*
If $n = 2m$, then there exist matrix polynomials $B_n(t) = \sum_{i=0}^{m} B_i t^i, C_{m-1}(t) = \sum_{i=0}^{m-1} C_i t^i$ such that

$$\begin{aligned} P_{2m}(t) &= B_m(t) B_m(t)^T + t(1-t) C_{m-1}(t) C^T_{m-1}(t) \\ &= \sum_{k=1}^{p} \Big[\, (\sum_{i=0}^{m} b_{ik} t^i)(\sum_{i=0}^{m} b_{ik} t^i)^T + t(1-t)(\sum_{i=0}^{m-1} c_{ik} t^i)(\sum_{i=0}^{m-1} c_{ik} t^i)^T \, \Big]. \end{aligned}$$

If $n = 2m+1$, then there exist matrix polynomials $B_m(t) = \sum_{i=0}^{m} B_i t^i, C_m(t) = \sum_{i=0}^{m} C_i t^i$ such that

$$\begin{aligned} P_{2m+1}(t) &= t B_m(t) B_m(t)^T + (1-t) C_m(t) C_m(t)^T \\ &= \sum_{k=1}^{p} \Big[\, t(\sum_{i=0}^{m} b_{ik} t^i)(\sum_{i=0}^{m} b_{ik} t^i)^T + (1-t)(\sum_{i=0}^{m} c_{ik} t^i)(\sum_{i=0}^{m} c_{ik} t^i)^T \, \Big]. \end{aligned}$$

Here $(b_{i1}, \ldots, b_{ip}) = B_i$ and $(c_{i1}, \ldots, c_{ip}) = C_i$ denote the columns of the coefficients B_i and C_i in the matrix polynomials $B_m(t)$, $C_{m-1}(t)$ and $C_m(t)$, respectively.

Proof: The proof follows from the corresponding results for trigonometric polynomials given by Malyshev (1982) [see also Rosenblatt (1956)]. This results states that if the matrix trigonometric polynomial satisfies

$$(2.8) \qquad A(\varphi) \;=\; \sum_{k=-N}^{N} A_k e^{ik\varphi} \;>\; 0$$

for all $\varphi \in \mathbb{R}$, where $A_0, \ldots, A_N$ are complex $p \times p$ matrices satisfying $A_k = A_k^*$ $(k = 0, \ldots, N)$, then there exists a unique matrix polynomial $\sum_{k=0}^{N} D_k e^{ik\varphi}$ with $D_0 = D_0^* > 0$ and $\det \sum_{k=0}^{N} D_k \lambda^k \neq 0$ for all $|\lambda| \leq 1$ such that the polynomial A can be represented as

$$(2.9) \qquad A(\varphi) \;=\; D(\varphi) D(\varphi)^*.$$

Moreover, if $A(\varphi)$ is semidefinite the representaion (2.9) also exists, but the polynomial D is not necessarily unique. The proof given by Malyshev (1982) can be easily extended to show that if $A_0, \ldots A_N$ are real symmetric matrices and $A_{-k} = A_k$ $(k = 1, \ldots, N)$, then a representation of the form (2.9) also exists, where the coefficients of the matrix polynomial D are real matrices.
With this result in hand the proof of Theorem 2.5 follows by similar arguments as given in Szegö (1959), Theorem 1.21.1 or in Dette and Studden (1997), Remark 9.2.9. To be precise let $P(t)$ denote a nonnegative definite matrix polynomial of degree $2m$ on the interval $[-1, 1]$ and put $t = \cos\varphi$. Because $P(\cos\varphi)$ is a cosine polynomial of degree $2m$ it follows that it has a representation of the form

$$P(\cos\varphi) \;=\; \sum_{k=-2m}^{2m} A_k e^{ik\varphi} \;\geq\; 0,$$

where the coefficients satisfy $A_{-k} = A_k$. Consequently, the generalization of Malyshev's result yields

$$(2.10) \qquad P(\cos\varphi) \;=\; D(\varphi) D(\varphi)^* \;=\; \Big(D(\varphi) e^{-im\varphi} \Big) \Big(D(\varphi) e^{-im\varphi} \Big)^*$$

for a real matrix polynomial $D(t) = \sum_{k=0}^{2m} D_k t^k$ of degree $2m$. Now

$$D(\varphi) e^{-im\varphi} \;=\; \sum_{k=0}^{2m} D_k e^{i(k-m)\varphi} \;=\; B_m(\cos\varphi) + i \sin\varphi \; C_{m-1}(\cos\varphi) \;,$$

where B_m and C_{m-1} are real matrix polynomials of degree m and $m-1$, respectively. This gives for the polynomial in (2.10)

$$\begin{aligned} P(\cos\varphi) \;=\; & B_m(\cos\varphi) B_m^T(\cos\varphi) \;+\; (\sin\varphi)^2 C_{m-1}(\cos\varphi) C_{m-1}^T(\cos\varphi) \\ & + i \sin\varphi \Big\{ C_{m-1}(\cos\varphi) B_m^T(\cos\varphi) - B_m(\cos\varphi) C_{m-1}^T(\cos\varphi) \Big\} \\ \;=\; & B_m(\cos\varphi) B_m^T(\cos\varphi) \;+\; (\sin\varphi)^2 C_{m-1}(\cos\varphi) C_{m-1}^T(\cos\varphi) \;, \end{aligned}$$

where the last equality follows from the fact that the left hand side of the above equation is a real matrix polynomial in $\cos\varphi$. This proves the assertion of Theorem 2.5 in the case $n = 2m$ for the interval $[-1, 1]$. The transformation to the interval $[0, 1]$ is obvious and the remaining case $n = 2m + 1$ follows by similar arguments. $\square$

Theorem 2.5 together with Lemma 2.3 now gives us necessary and sufficient conditions for the point $S = (S_0, \cdots, S_n)$ to belong to the moment space M_{n+1} or to its interior. To this end we define the "Hankel" matrices

$$(2.11) \qquad \underline{H}_{2m} = \begin{pmatrix} S_0 & \cdots & S_m \\ \vdots & & \vdots \\ S_m & \ldots & S_{2m} \end{pmatrix} \qquad \overline{H}_{2m} = \begin{pmatrix} S_1 - S_2 & \cdots & S_m - S_{m+1} \\ \vdots & & \vdots \\ S_m - S_{m+1} & \ldots & S_{2m-1} - S_{2m} \end{pmatrix}$$

and

$$(2.12) \qquad \underline{H}_{2m+1} = \begin{pmatrix} S_1 & \cdots & S_{m+1} \\ \vdots & & \vdots \\ S_{m+1} & \ldots & S_{2m+1} \end{pmatrix} \qquad \overline{H}_{2m+1} = \begin{pmatrix} S_0 - S_1 & \cdots & S_m - S_{m+1} \\ \vdots & & \vdots \\ S_m - S_{m+1} & \ldots & S_{2m} - S_{2m+1} \end{pmatrix}$$

and obtain the following characterization.

Theorem 2.6:
a) *The point $(S_0, \cdots, S_n)$ is in the moment space M_{n+1} if and only if the matrices $\underline{H}_n$ and $\overline{H}_n$ are nonnegative definite.*
b) *The point $S = (S_0, \ldots, S_n)$ is in the interior of the moment space M_{n+1} if and only if the matrices $\underline{H}_n$ and $\bar{H}_n$ are positive definite.*

The nonnegativity of the matrices $\underline{H}_n$ and $\overline{H}_n$ impose limits on the moments S_k as is the one dimensional case. To be precise let

$$\underline{h}_{2m}^T = (S_{m+1}, \cdots, S_{2m})$$
$$\underline{h}_{2m-1}^T = (S_m, \cdots, S_{2m-1})$$
$$\bar{h}_{2m}^T = (S_m - S_{m+1}, \cdots, S_{2m-1} - S_{2m})$$
$$\bar{h}_{2m-1}^T = (S_m - S_{m+1}, \cdots, S_{2m-2} - S_{2m-1})$$

and define $S_1^- = 0$ and

$$(2.13) \qquad S_{n+1}^- = \underline{h}_n^T \underline{H}_{n-1}^{-1} \underline{h}_n, \quad n \geq 1 \ ,$$

and $S_1^+ = S_0$, $S_2^+ = S_1$ and

$$(2.14) \qquad S_{n+1}^+ = S_n - \bar{h}_n^T \bar{H}_{n-1}^{-1} \bar{h}_n, \quad n \geq 2 \ ,$$

whenever the inverses of the Hankel matrices exist. It is to be noted and stressed that S_n^- and S_n^+ depend on $(S_0, S_1, \cdots, S_{n-1})$ although this is not mentioned explicitly. It follows from Theorem 2.6 and a straightforward calculation with partitioned matrices that $(S_0, \ldots, S_{n-1})$ is in the interior of the moment space M_n if and only if $S_n^- < S_n^+$ (note that a matrix is positve definite if and only if its main subblock and the corresponding Schur complement are positive definite). Moreover, for $(S_0, \ldots, S_n) \in M_{n+1}$ we have

$$(2.15) \qquad S_n^- \leq S_n \leq S_n^+.$$

If $(S_0, S_1, \cdots, S_{n-1})$ is in the interior of the moment space M_n, then we define the kth matrix canonical moment as the matrix

$$P_k = D_k^{-1}(S_k - S_k^-) \ , \qquad 1 \le k \le n \ , \tag{2.16}$$

where

$$D_k = S_k^+ - S_k^- . \tag{2.17}$$

These quantities are the analog of the classical canonical moments p_n in the scalar case [see Skibinsky (1967, 1969, 1986) or Dette and Studen (1997)]. We will also make use of the quantities

$$Q_k = I - P_k = D_k^{-1}(S_k^+ - S_k) \ , \qquad 1 \le k \le n \ . \tag{2.18}$$

The canonical moments of lower order can easily be calculated. From $S_1^+ = S_0$, $S_1^- = 0$ we have $D_1 = S_0$ and $P_1 = S_0^{-1}S_1$. Similary, the definitions (2.13) and (2.14) imply

$$\begin{aligned} S_2^+ &= S_1 \\ S_2^- &= S_1 S_0^{-1} S_1 \end{aligned}$$

which gives

$$D_2 \ = \ S_1(I_p - S_0^{-1}S_1) \ = \ S_1(I_p - P_1) \ = \ S_1 Q_1$$

and

$$P_2 = (I_p - S_0^{-1}S_1)^{-1} S_1^{-1}(S_2 - S_1 S_0^{-1} S_1) \ .$$

One of our main theorems in this section is the following result, which represents the width D_{n+1} of the moment space M_{n+1} in terms of the matrix canonical moments P_k and Q_k.

Theorem 2.7: *If the point $(S_0, \cdots, S_n)$ is in the interior of the moment space M_{n+1}, then*

$$D_{n+1} = S_0 P_1 Q_1 P_2 Q_2 \cdots P_n Q_n \tag{2.19}$$

and

$$P_k Q_k = Q_k P_k, \quad \text{for} \quad k = 1, \cdots, n \ .$$

Proof. The proof will follow if we can show the representations

$$S_{n+1}^+ - S_{n+1}^- = (S_n^+ - S_n)(S_{n-1}^+ - S_{n-1})^{-1}(S_{n-1}^+ - S_{n-1}^-)(S_{n-1} - S_{n-1}^-)^{-1}(S_n - S_n^-) \tag{2.20}$$

or equivalently

$$S_{n+1}^+ - S_{n+1}^- = (S_n^+ - S_n)[(S_{n-1}^+ - S_{n-1})^{-1} + (S_{n-1} - S_{n-1}^-)^{-1}](S_n - S_n^-). \tag{2.21}$$

Assuming that (2.20) is true we can obtain the assertion of Theorem 2.7 by a simple induction argument. Note that $D_1 = S_1^+ - S_1^- = S_0$ and using $P_1 = S_0^{-1}S_1$ we find that $P_1 Q_1 = Q_1 P_1$ and

$$D_2 = S_2^+ - S_2^- = S_1 - S_1 S_0^{-1} S_1 = S_0 P_1 Q_1 = S_0 Q_1 P_1.$$

Proceeding by induction we can then show that

$$
\begin{aligned}
D_{n+1} &= D_nQ_n(D_{n-1}Q_{n-1})^{-1}D_{n-1}(D_{n-1}P_{n-1})^{-1}D_nP_n \\
&= D_nQ_nQ_{n-1}^{-1}P_{n-1}^{-1}D_{n-1}^{-1}D_nP_n \\
&= D_nQ_nP_n
\end{aligned}
$$

where the first equality follows from (2.20), (2.16), (2.18) and the last is a conesequence of the induction hypothesis. Noting that all the terms in (2.13) - (2.17) are symmetric, it also follows that

$$
\begin{aligned}
D_{n+1} &= (S_n - S_n^-)(S_{n-1} - S_{n-1}^-)^{-1}(S_{n-1}^+ - S_{n-1}^-)(S_{n-1}^+ - S_{n-1})^{-1}(S_n^+ - S_n) \\
&= D_nP_n(D_{n-1}P_{n-1})^{-1}D_{n-1}(D_{n-1}Q_{n-1})^{-1}D_nQ_n \\
&= D_nP_nQ_n
\end{aligned}
$$

in which case $P_nQ_n = Q_nP_n$.

Proof of the identities (2.20) and (2.21): From the definition (2.13) and (2.14) we have

$$S_{n+1}^+ - S_{n+1}^- = S_n - \bar{h}_n^T\bar{H}_{n-1}^{-1}\bar{h}_n - \underline{h}_n^T\underline{H}_{n-1}^{-1}\underline{h}_n \tag{2.22}$$

To show (2.21) we want to factor a term $S_n - S_n^-$ from the right and a term $S_n^+ - S_n$ from the left and leave the appropriate thing in the middle to give the identity (2.21). This will follow, with some further explanation, from the following two interesting results. To be precise let $\bar{h}_n(-)$ and $\underline{\mathrm{h}}_n(-)$ be the value of $\bar{h}_n$ and $\underline{\mathrm{h}}_n$, respectively, where the last moment matrix S_n is changed to S_n^-. Similary, $\bar{h}_n(+)$ and $\underline{\mathrm{h}}_n(+)$ are obtained from $\bar{h}_n$ and $\underline{h}_n$, replacing the last moment matrix S_n by S_n^+, respectively and we define $\bar{g}_n$ and $\underline{g}_n$ in the same manner as $\bar{h}_n$ and $\underline{h}_n$ where all matrices S_k are replaced by arbitrary elements $G_k \in \mathbb{R}^{p\times p}$. We will show below that the identities

$$S_n^- - \bar{h}_n^T\bar{H}_{n-1}^{-1}\bar{h}_n(-) - \underline{h}_n^T\underline{H}_{n-1}^{-1}\underline{h}_n(-) = 0 \tag{2.23}$$

and

$$G_n - \bar{h}_n^T(+)\bar{H}_{n-1}^{-1}\bar{g}_n - \underline{h}_n^T(+)\underline{H}_{n-1}^{-1}\underline{g}_n = 0 \tag{2.24}$$

hold for all $\bar{g}_n$ and $\underline{g}_n$. Note carefully in the last equation that the leading matrix G_n of equation (2.24) is the same G_n in the last coordinate of $\bar{g}_n$ and $\underline{g}_n$.

Once we have established the validity of (2.23) and (2.24) we start with the expression for $S_{n+1}^+ - S_{n+1}^-$ in (2.22) and rewrite $S_n = S_n - S_n^- + S_n^-$ for the first matrix S_n in (2.22) and on the right side of the two quadratic forms. Using the identity (2.23) we can factor the matrix $S_n - S_n^-$ from the right side and the result is

$$S_{n+1}^+ - S_{n+1}^- = \Big[I_p - \bar{h}_n^T\bar{H}_{n-1}^{-1}(0,\cdots,0,-I_p)^T - \underline{h}_n^T\underline{H}_{n-1}^{-1}(0,\cdots,0,I_p)^T\Big](S_n - S_n^-) \tag{2.25}$$

This resulting expression can be further simplified rewriting $S_n = S_n^+ + S_n - S_n^+$ on the left side of what remains of the two quadratic forms. Using (2.24) (applied by setting on the right side of the quadratic forms $G_n = I_p$ and the other G_k values equal to zero in $\bar{g}_n$ and $\underline{g}_n$) and the definition (2.13) and (2.14) everything cancels and we are left with (2.21).

Proof of the identities (2.23) and (2.24): To see the proof of (2.23) and (2.24) we consider only the case where $n = 2m$, the odd case being similar. For $n = 2m$ we have to show that

$$
\begin{aligned}
(2.26)\qquad R_m \;=\; S_{2m}^- \;&-\; \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-1} - S_{2m} \end{pmatrix}^T \begin{pmatrix} S_0 - S_1 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-2} - S_{2m-1} \end{pmatrix}^{-1} \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-1} - S_{2m}^- \end{pmatrix} \\
&-\; \begin{pmatrix} S_{m+1} \\ \vdots \\ S_{2m} \end{pmatrix}^T \begin{pmatrix} S_1 & \cdots & S_m \\ \vdots & & \vdots \\ S_m & \cdots & S_{2m-1} \end{pmatrix}^{-1} \begin{pmatrix} S_{m+1} \\ \vdots \\ S_{2m}^- \end{pmatrix} \;=\; 0,
\end{aligned}
$$

which is equivalent to (2.23) by the definition (2.11) and (2.12). To this end we write the column vectors on the right side of the two quadratic forms in the above expression in a convenient form. From the definition of the matrix S_{2m}^- in (2.13) we obtain the representation

$$
\begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \\ S_{2m}^- \end{pmatrix} = \begin{pmatrix} S_0 & \cdots & S_{m-1} \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \\ S_m & \cdots & S_{2m-1} \end{pmatrix} \begin{pmatrix} S_0 & \cdots & S_{m-1} \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \end{pmatrix}^{-1} \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix} .
$$

Taking differences and omitting the first component it follows that

$$
(2.27)\qquad \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-1} - S_{2m}^- \end{pmatrix} = \begin{pmatrix} S_0 - S_1 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-2} - S_{2m-1} \end{pmatrix} \begin{pmatrix} S_0 & \cdots & S_{m-1} \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \end{pmatrix}^{-1} \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix}
$$

and

$$
(2.28)\qquad \begin{pmatrix} S_{m+1} \\ \vdots \\ S_{2m-1} \\ S_{2m}^- \end{pmatrix} = \begin{pmatrix} S_1 & \cdots & S_m \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \\ S_m & \cdots & S_{2m-1} \end{pmatrix} \begin{pmatrix} S_0 & \cdots & S_{m-1} \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \end{pmatrix}^{-1} \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix} .
$$

Substituting (2.27) and (2.28) into (2.26) we find that (2.26) becomes

$$
R_m \;=\; S_{2m}^- \;-\; \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-1} - S_{2m} \end{pmatrix}^T \begin{pmatrix} S_0 & \cdots & S_{m-1} \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \end{pmatrix}^{-1} \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix}
$$

$$-\begin{pmatrix} S_{m+1} \\ \vdots \\ S_{2m} \end{pmatrix}^T \begin{pmatrix} S_0 & \cdots & S_{m-1} \\ \vdots & & \vdots \\ S_{m-1} & \cdots & S_{2m-2} \end{pmatrix}^{-1} \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix}.$$

Now separate the vector on the left in the first quadratic form. One part will cancel with the second quadratic form and the rest will be zero from the definition of S_{2m}^-, which proves the identity (2.26).

To verify the equation (2.24) the argument is essentially the same but slightly more complicated. We need to verify that

$$(2.29)\quad \tilde{R}_m = G_{2m} - \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-1} - S_{2m}^+ \end{pmatrix}^T \begin{pmatrix} S_0 - S_1 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-2} - S_{2m-1} \end{pmatrix}^{-1} \begin{pmatrix} G_m - G_{m+1} \\ \vdots \\ G_{2m-1} - G_{2m} \end{pmatrix}$$
$$- \begin{pmatrix} S_{m+1} \\ \vdots \\ S_{2m}^+ \end{pmatrix}^T \begin{pmatrix} S_1 & \cdots & S_m \\ \vdots & & \vdots \\ S_m & \cdots & S_{2m-1} \end{pmatrix}^{-1} \begin{pmatrix} G_{m+1} \\ \vdots \\ G_{2m} \end{pmatrix} = 0$$

holds for all matrices $G_m, \ldots, G_{2m}$. From the definition of the matrix S_{2m}^+ in (2.14) we get

$$(2.30)\quad \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-1} - S_{2m}^+ \end{pmatrix}^T = \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-2} - S_{2m-1} \end{pmatrix}^T \begin{pmatrix} S_1 - S_2 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-3} - S_{2m-2} \end{pmatrix}^{-1} \times$$
$$\begin{pmatrix} S_1 - S_2 & \cdots & S_{m-1} - S_m & S_m - S_{m+1} \\ \vdots & & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-3} - S_{2m-2} & S_{2m-2} - S_{2m-1} \end{pmatrix}$$

which can be rewritten as

$$(2.31)\quad \begin{pmatrix} S_{m+1} \\ \vdots \\ S_{2m}^+ \end{pmatrix}^T = \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix}^T - \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-2} - S_{2m-1} \end{pmatrix}^T \begin{pmatrix} S_1 - S_2 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-3} - S_{2m-2} \end{pmatrix}^{-1}$$
$$\times D(m-1,m) \begin{pmatrix} S_1 & \cdots & S_m \\ S_2 & \cdots & S_{m+1} \\ \vdots & & \vdots \\ S_m & \cdots & S_{2m-1} \end{pmatrix}$$

where

$$D(m-1,m) = \begin{pmatrix} I_p & -I_p & 0 & 0 & \ldots \\ 0 & I_p & -I_p & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \ldots & 0 & 0 & I_p & -I_p \end{pmatrix}$$

is a block differencing matrix of size $(m-1)p^2 \times mp^2$. Substituting (2.30) and (2.31) into (2.29) yields

$$
\begin{aligned}
\tilde{R}_m \;=\; G_{2m} \;-\; & \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-2} - S_{2m-1} \end{pmatrix}^T \begin{pmatrix} S_1 - S_2 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-3} - S_{2m-2} \end{pmatrix}^{-1} \\
& \times \begin{pmatrix} 0 & I_p & 0 & \cdots \\ 0 & 0 & I_p & \\ \vdots & & & \end{pmatrix} \begin{pmatrix} G_m - G_{m+1} \\ \vdots \\ G_{2m-1} - G_{2m} \end{pmatrix} \\
& - \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix}^T \begin{pmatrix} S_1 & \cdots & S_m \\ \vdots & & \vdots \\ S_m & \cdots & S_{2m-1} \end{pmatrix}^{-1} \begin{pmatrix} G_{m+1} \\ \vdots \\ G_{2m} \end{pmatrix} \\
& + \begin{pmatrix} S_m - S_{m+1} \\ \vdots \\ S_{2m-2} - S_{2m-1} \end{pmatrix}^T \begin{pmatrix} S_1 - S_2 & \cdots & S_{m-1} - S_m \\ \vdots & & \vdots \\ S_{m-1} - S_m & \cdots & S_{2m-3} - S_{2m-2} \end{pmatrix}^{-1} \begin{pmatrix} G_{m+1} - G_{m+2} \\ \vdots \\ G_{2m-1} - G_{2m} \end{pmatrix}
\end{aligned} \tag{2.32}
$$

This is zero since the first and second terms cancel with the third and fourth, respectively. This proves the identity (2.24) and completes the proof of the theorem.

□

For later use we note that Theorem 2.7 implies, using an induction argument, that the following important result holds.

Theorem 2.8: *If for some $n \geq 1$, the point $(S_0, S_1, \cdots, S_n)$ is in the interior of the moment space M_{n+1} then ($S_0^- = 0$, $Q_0 = I_p$)*

$$
(S_{n-1} - S_{n-1}^-)^{-1}(S_n - S_n^-) = Q_{n-1}P_n \;=:\; \zeta_n \tag{2.33}
$$

$$
(S_{n-1}^+ - S_{n-1})^{-1}(S_n^+ - S_n) = P_{n-1}Q_n \;=:\; \gamma_n
$$

Proof. The result follows since from Theorem 2.7 and (2.16) we have

$$
\begin{aligned}
S_n - S_n^- &= D_n P_n \\
&= S_0 P_1 Q_1 \cdots P_{n-1} Q_{n-1} P_n \;>\; 0
\end{aligned} \tag{2.34}
$$

and

$$
\begin{aligned}
S_n^+ - S_n &= D_n Q_n \\
&= S_0 P_1 Q_1 \cdots Q_{n-1} P_{n-1} Q_n \;>\; 0
\end{aligned} \tag{2.35}
$$

□

Example 2.9: Consider the matrix measure

$$\mu = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{12} & \mu_{22} \end{pmatrix}$$

where

$$\frac{d\mu_{11}}{dt} = \frac{d\mu_{22}}{dt} = \frac{1}{\pi}\frac{1}{\sqrt{t(1-t)}}$$

$$\frac{d\mu_{12}}{dt} = \frac{1}{\pi}\frac{2t-1}{\sqrt{t(1-t)}}$$

[see also VanAssche (1993), who considered this example on the interval $[-1,1]$. A straightforward calculation shows

$$S_k = \binom{2k}{k}\frac{1}{2^{2k}(k+1)}\begin{pmatrix} k+1 & k \\ k & k+1 \end{pmatrix} \quad , \quad k \geq 0 \ ,$$

which gives for the first canonical moments

$$P_1 = S_0^{-1}S_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix} \ , \ P_2 = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \ .$$

It can be shown by tedious computations that the canonical moments of the matrix measure μ are given by

$$P_{2k} = \frac{k}{2k+1}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \ P_{2k-1} = \frac{1}{2}\begin{pmatrix} 1 & \frac{1}{2k} \\ \frac{1}{2k} & 1 \end{pmatrix} \ ,$$

for $k \in \mathbb{N}$.

3 Orthogonal Polynomials

The inner product of two matrix polynomials is defined by

(3.1) $$< P, Q > = \int P^T(t)\mu(dt)Q(t).$$

Sinap and Van Assche (1996) call this the 'right' inner product. The left inner product would put the transpose of the Q polynomial. The orthogonal polynomials are defined by orthogonalizing the sequence $I_p, tI_p, t^2I_p, \cdots$ with respect to the above inner product. As the main result of this section we will write down explicitly the orthogonal polynomials somewhat in the fashion of the one-dimensional case. In the case $p = 1$ the last row of the determinant

(3.2) $$\underline{D}_{2m} = \begin{vmatrix} S_0 & \cdots & S_m \\ \vdots & & \vdots \\ S_m & \cdots & S_{2m} \end{vmatrix}$$

is changed by replacing it by $(1, t, \cdots, t^m)$ [see e.g. Szegö (1959)]. The general case $p \geq 2$ is substantially more complicated, because $P_m(t)$ is a matrix of polynomials each of degree m. To get the element in the position (i, j) of the matrix of $P_m(t)$ we write each of these by again modifing the above determinant by changing the last set of blocks $(S_m, \cdots, S_{2m})$.

To be precise define a matrix polynomial by

$$P_m(t) = \underline{H}_{2m-1}(t) = (H_{ij}(t))_{ij=1}^p$$

where the elements $H_{ij}(t)$ are determinants given by

$$H_{ij}(t) = \begin{vmatrix} S_0 & S_1 & \cdots & S_m \\ \vdots & & & \vdots \\ S_{m-1} & S_m & \cdots & S_{2m-1} \\ S_m^{ij} & S_{m+1}^{ij} & \cdots & S_{2m}^{ij} \end{vmatrix} \qquad i, j = 1, \ldots, p$$

and the matrix S_{m+k}^{ij} is obtained from S_{m+k} replacing the jth row by $e_i^T t^k$. Here $e_i \in \mathbb{R}^p$ denotes the vector with a one in the ith component and zero elsewhere.
As a simple example consider the case $m = 1$ and $p = 2$ and let $S_k = (s_{k,ij})_{i,j=1}^p$ denote the elements of the kth moment matrix. In this case the linear polynomial is given by the 2×2 matrix of determinants

$$P_1(t) = \underline{\mathrm{H}}_1(t) = \begin{pmatrix} \begin{vmatrix} \begin{pmatrix} s_{0,11} & s_{0,12} \\ s_{0,21} & s_{0,22} \end{pmatrix} & \begin{pmatrix} s_{1,11} & s_{1,12} \\ s_{1,21} & s_{1,22} \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ s_{1,21} & s_{1,22} \end{pmatrix} & \begin{pmatrix} t & 0 \\ s_{2,21} & s_{2,22} \end{pmatrix} \end{vmatrix} & \begin{vmatrix} \begin{pmatrix} s_{0,11} & s_{0,12} \\ s_{0,21} & s_{0,22} \end{pmatrix} & \begin{pmatrix} s_{1,11} & s_{1,12} \\ s_{1,21} & s_{1,22} \end{pmatrix} \\ \begin{pmatrix} s_{1,11} & s_{1,12} \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} s_{2,11} & s_{2,12} \\ t & 0 \end{pmatrix} \end{vmatrix} \\ \\ \begin{vmatrix} \begin{pmatrix} s_{0,11} & s_{0,12} \\ s_{0,21} & s_{0,22} \end{pmatrix} & \begin{pmatrix} s_{1,11} & s_{1,12} \\ s_{1,21} & s_{1,22} \end{pmatrix} \\ \begin{pmatrix} 0 & 1 \\ s_{1,21} & s_{1,22} \end{pmatrix} & \begin{pmatrix} 0 & t \\ s_{2,21} & s_{2,22} \end{pmatrix} \end{vmatrix} & \begin{vmatrix} \begin{pmatrix} s_{0,11} & s_{0,12} \\ s_{0,21} & s_{0,22} \end{pmatrix} & \begin{pmatrix} s_{1,11} & s_{1,12} \\ s_{1,21} & s_{1,22} \end{pmatrix} \\ \begin{pmatrix} s_{1,11} & s_{1,12} \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} s_{2,11} & s_{2,12} \\ 0 & t \end{pmatrix} \end{vmatrix} \end{pmatrix}$$

The following result shows that the polynomials $P_j(t)$ are the right orthogonal matrix polynomials with respect to the inner product defined by (3.1) and identifies the leading coefficient of $P_m(t)$, i.e. the matrix of the term t^m.

Theorem 3.1: *The polynomial $P_m(t) = \underline{H}_{2m-1}(t)$ is the mth matrix orthogonal polynomial with respect to the inner product (3.1) and has leading coefficient*

$$L_m = \underline{D}_{2m}(S_{2m} - S_{2m}^-)^{-1},$$

where the determinant $\underline{D}_{2m}$ is defined in (3.2).

Proof: It is required to show that

$$B = \int_0^1 t^k d\mu(t) \underline{H}_{2m-1}(t) = 0 \in \mathbb{R}^{p\times p} \ , \qquad \text{for all} \quad k = 0, 1, \cdots, m-1.$$

For the element in the position (i,j) of this matrix we obtain

$$B_{ij} = \sum_{l=1}^{p} \int_0^1 H_{lj}(t) t^k d\mu_{il}(t)$$

and a straightforward calculation shows that this is equal to the determinant $\underline{D}_{2m}$ with the jth row replaced by

$$\int_0^1 t^k (d\mu_{i1}, \cdots, d\mu_{ip}), \cdots, \int_0^1 t^{k+m} (d\mu_{i1}, \cdots, d\mu_{ip}).$$

For this reason B_{ij} is zero because this is the ith row of $(S_k, \cdots, S_{k+m})$, provided $k < m$. If $k = m$, we have by the same argument that if $i \neq j$, $B_{ij} = 0$ (because here the block corresponding to $(S_m, \cdots, S_{2m})$ has two equal rows) and in the case $i = j$, $B_{ii} = \underline{D}_{2m}$ for $i = 1, \cdots, m$.
Note that we are using the right inner product (3.1) which gives for all $p \times p$ matrices A and B

$$< PA, QB > \ = \ A^T < P, Q > B.$$

Therefore, if L_m is the leading coefficient of $P_m(t)$ we have from the previous calculations

$$< P_m, P_m > \ = \ L_m^T diag(\underline{D}_{2m}, \cdots, \underline{D}_{2m}) \ = \ L_m^T \underline{\Delta}_{2m} \tag{3.3}$$

where the last line defines the diagonal matrix $\underline{\Delta}_{2m} \in \mathbb{R}^{p\times p}$. The leading coefficient of the polynomial $\underline{H}_{2m-1}(t)$ can be now calculated by noting that the leading coefficient L_m of H_{ij} is obtained by deleting the row and column corresponding to t^m [i.e. the $(mp + j)$th row and $(mp + i)$th column] and multiplying by $(-1)^{i+j}$. The same value is obviously obtained by deleting the same row and column in the matrix $\underline{\mathrm{H}}_{2m}$ defined in (2.11) and calculating the resulting determinant, which is denoted by $\underline{\mathrm{D}}_{2m}^{pm+j,pm+i}$ $(i, j = 1, \ldots, p)$. Observing that the matrix

$$((-1)^{i+j} \underline{\mathrm{D}}_{2m}^{pm+j,pm+i})_{i,j=1}^{p}$$

is proportional to the block in the position $(m + 1, m + 1)$ of the matrix $\underline{\mathrm{H}}_{2m}^{-1}$ (by Cramer's rule) we obtain from the definition of S_{2m}^{-} in (2.13)

$$\begin{aligned} L_m &= ((-1)^{i+j} \underline{D}_{2m}^{pm+j,pm+i})_{ij=1}^{p} \\ &= \underline{\mathrm{D}}_{2m} \Big[S_{2m} - (S_m, \ldots, S_{2m-1}) \underline{\mathrm{H}}_{2m-2}^{-1} \begin{pmatrix} S_m \\ \vdots \\ S_{2m-1} \end{pmatrix} \Big]^{-1} \\ &= \underline{D}_{2m} (S_{2m} - S_{2m}^{-})^{-1}, \end{aligned} \tag{3.4}$$

which proves the remaining assertion of Thorem 3.1. □

It follows from Theorem 3.1 that the monic orthogonal polynomials are given by

$$(3.5) \qquad \underline{P}_m(x) \;=\; P_m(x)L_m^{-1} \;=\; P_m(x)(S_{2m}-S_{2m}^{-})/\underline{D}_{2m}.$$

where we have to multiply from the right, because we use the right inner product. The matrix orthogonal polynomials with respect to the measure $td\mu(t)$ are obtained similarly and given in the following Theorem. The proof is left to the reader.

Theorem 3.2: *The polynomials $Q_m(t)$ defined by*

$$(3.6) \qquad Q_m(t) = \underline{H}_{2m}(t) = (K_{ij}(t))_{ij=1}^{p}$$

with

$$(3.7) \qquad K_{ij}(t) \;=\; \begin{vmatrix} S_1 & S_2 & \cdots & S_{m+1} \\ S_2 & S_3 & \cdots & S_{m+2} \\ \vdots & \vdots & & \vdots \\ S_m & S_{m+1} & \cdots & S_{2m} \\ S_{m+1}^{ij} & S_{m+2}^{ij} & \cdots & S_{2m+1}^{ij} \end{vmatrix}.$$

are the right orthogonal polynomials with respect to the matrix measure $td\mu(t)$. Moreover, the leading coefficient of $Q_m(t)$ is given by

$$(3.8) \qquad K_m = ((-1)^{i+j}\underline{D}_{2m+1}^{mp+i,mp+j})_{ij=1}^{p} = \underline{D}_{2m+1}(S_{2m+1}-S_{2m+1}^{-})^{-1}.$$

where $\underline{D}_{2m+1} = |\underline{H}_{2m+1}|$ and $\underline{D}_{2m+1}^{mp+i,mp+j}$ denotes the determinant of the matrix which is obtained from $\underline{H}_{2m+1}$ by deleting the $(mp+i)$th column and $(mp+j)$th row.

We finally note that it follows from the proof of Theorem 3.2

$$(3.9) \qquad \begin{aligned} < Q_m, Q_m >^* &= K_m^T diag(\underline{D}_{2m+1}, \cdots, \underline{D}_{2m+1}) \\ &= K_m^T \underline{\Delta}_{2m+1}, \end{aligned}$$

where, $<>^*$ is the right inner product with respect to the matrix measure $td\mu(t)$, and that the monic orthogonal polynomials are given by

$$(3.10) \qquad \underline{Q}_m(x) = Q_m(x)K_m^{-1} = Q_m(x)(S_{2m+1}-S_{2m+1}^{-})/\underline{D}_{2m+1}\ .$$

Lemma 3.3:
The sequence of monic orthogonal polynomials $\{\underline{P}_k(x)\}_{k\geq 0}$ with respect to the matrix measure μ satisfies the recurrence formula $\underline{P}_0(x) = I_p, \underline{P}_{-1}(x) = 0$ and for $m \geq 0$

$$(3.11) \qquad x\underline{P}_m(x) = \underline{P}_{m+1}(x) + \underline{P}_m(x)(\zeta_{2m+1}+\zeta_{2m}) + \underline{P}_{m-1}(x)\zeta_{2m-1}\zeta_{2m},$$

where the quantities $\zeta_j \in \mathbb{R}^{p\times p}$ are defined by $\zeta_0 = 0$, $\zeta_1 = P_1$, $\zeta_j = Q_{j-1}P_j$ if $j \geq 2$.

Similary, the sequence of monic orthogonal polynomials $\{\underline{Q}_k(x)\}_{k\geq 0}$ *with respect to the matrix measure* $xd\mu(x)$ *satisfies the recurrence formula* $\underline{Q}_0(x) = I_p, \underline{Q}_{-1}(x) = 0$ *and for* $m \geq 0$

$$xQ_m(x) = \underline{Q}_{m+1}(x) + \underline{Q}_m(x)(\zeta_{2m+1} + \zeta_{2m+2}) + \underline{Q}_{m-1}(x)\zeta_{2m}\zeta_{2m+1} \ . \tag{3.12}$$

Proof: The recurrence formula for the orthogonal polynomials will be obtained by using the two sets of polynomials $\underline{P}_m(x)$ and $\underline{Q}_m(x)$ and the two recurrence formula

$$\underline{P}_{m+1}(x) = x\underline{Q}_m(x) - \underline{P}_m(x)A_m \tag{3.13}$$

and

$$\underline{Q}_{m+1}(x) = \underline{P}_{m+1}(x) - \underline{Q}_m(x)B_m. \tag{3.14}$$

These identities follow by multiplying the right hand sides with x^k and integrateing with respect to the measures $d\mu(x)$ and $xd\mu(x)$, respectively. By an appropriate definition of the matrices A_m and B_m in the case $k = m$ it follows that the multiplied and integrated right hand sides of (3.13) and (3.14) vanish and consequently these terms are the monic orthogonal matrix polynomials with respect to the measures $d\mu(x)$ and $xd\mu(x)$, respectively.
From (3.13) we get that

$$0 \ = \ <\underline{P}_m, \underline{P}_{m+1}> \ = \ <\underline{P}_m, x\underline{Q}_m> \ - \ <\underline{P}_m, \underline{P}_m> A_m.$$

Observing the identity (3.3) in the proof of Theorem 3.1 we obtain

$$\begin{aligned} <\underline{P}_m, \underline{P}_m> \ &= \ (L_m^{-1})^T < P_m, P_m > L_m^{-1} \\ &= \ (L_m^{-1})^T L_m^T \underline{\Delta}_{2m} L_m^{-1} \ = \ \underline{\Delta}_{2m} L_m^{-1} \\ &= \ S_{2m} - S_{2m}^- \end{aligned}$$

and similarly [recall (3.8) - (3.10)]

$$\begin{aligned} <\underline{P}_m, x\underline{Q}_m> \ &= \ <\underline{P}_m, xQ_m> K_m^{-1} \ = \ (K_m^{-1})^T < Q_m, Q_m >^* K_m^{-1} \\ &= \ \underline{\Delta}_{2m+1} K_m^{-1} = S_{2m+1} - S_{2m+1}^- \end{aligned}$$

which yields [observing the definition of ζ_j in (2.33)]

$$A_m = (S_{2m} - S_{2m}^-)^{-1}(S_{2m+1} - S_{2m+1}^-) = \zeta_{2m+1} \tag{3.15}$$

Similarly, we can multiply the representation (3.14) by $\underline{Q}_m$ and take the inner product with respect to the matrix measure $td\mu(t)$ to get

$$B_m = (S_{2m+1} - S_{2m+1}^-)^{-1}(S_{2m+2} - S_{2m+2}^-) = \zeta_{2m+2} \tag{3.16}$$

To combine the two recurrence formula (3.13) and (3.14) into one recurrence formula for the sequence of monic orthogonal polynomials $\underline{P}_m$ we write

$$\underline{Q}_m(x) = x^{-1}(\underline{P}_{m+1}(x) + \underline{P}_m(x)A_m)$$

which gives

$$\begin{aligned}\underline{P}_m(x) &= \underline{Q}_m(x) + \underline{Q}_{m-1}(x)B_{m-1} \\ &= x^{-1}(\underline{P}_{m+1}(x) + \underline{P}_m(x)A_m) + x^{-1}(\underline{P}_m(x) + \underline{P}_{m-1}(x)A_{m-1})B_{m-1}\end{aligned}$$

or equivalently

$$x\underline{P}_m(x) = \underline{P}_{m+1}(x) + \underline{P}_m(x)(A_m + B_{m-1}) + \underline{P}_{m-1}(x)A_{m-1}B_{m-1} \tag{3.17}$$

From (2.33), (3.15) and (3.16) we have $A_m = \zeta_{2m+1}$ and $B_{m-1} = \zeta_{2m}$ where $\zeta_k = Q_{k-1}P_k$, which proves the first assertion of the lemma. The second part follows by similar arguments and the corresponding proof is therefore omitted. □

4 Favard's Theorem on the interval $[0, 1]$

Using the results of sections 2 and 3 we have the following result.

Theorem 4.1:
a) If $\{\underline{P}_n\}_{n\geq 1}$ is the sequence of monic polynomials which are orthogonal with respect to the matrix measure μ on the interval $[0, 1]$ then the sequence satisfies the recurrence formula (3.11) where the coefficients $\zeta_n = Q_{n-1}P_n$ in the recursion satisfy (2.34) and (2.35).

b) Conversely, if the sequence of monic polynomials $\{\underline{P}_n\}_{n\geq 1}$ satisfies the recurrence formula (3.11) where the ζ_n form a chain sequence (i.e. there exist matrices P_n and Q_n such that $\zeta_n = Q_{n-1}P_n$) and satisfy (2.34) and (2.35) with $S_0 = I$, then there exists a matrix measure μ supported on the interval $[0, 1]$ for which the polynomials $\{\underline{P}_n\}_{n\geq 1}$ are the corresponding monic orthogonal matrix polynomials.

Proof: Part a) has already been shown above. For part b) we define the corresponding moment sequence to obtain the measure μ. To do this we let $S_0 = I_p$, $S_1 = P_1$ and successively define S_n for $n \geq 2$ by setting

$$S_n = S_n^- + D_nP_n. \tag{4.1}$$

Recall that S_n^- is defined in (2.13) and depends only on $S_0, S_1, \cdots, S_{n-1}$. If the matrix canonical moments P_n are such that (2.34) and (2.35) hold, then it follows that

$$S_n^- < S_n < S_n^+. \tag{4.2}$$

In this case the defined sequence is such that $(S_0, S_1, \cdots, S_n)$ is in the interior of M_{n+1} for each n. Any corresponding sequence of measures will converge to a measure μ with the moments S_n for all $n \in \mathbb{N}_0$.
The sequence of monic orthogonal polynomials from this measure μ will satisfy the same recurrence formulas so these must be the sequence $\{\underline{P}_n\}_{n\geq 1}$ because the monic orthogonal polynomials are unique. $\square$

Example 4.2: Consider the matrix measure discussed in Example 2.9. We have

$$\zeta_{2k-1} = \frac{k}{2(2k-1)} \begin{pmatrix} 1 & \frac{1}{2k} \\ \frac{1}{2k} & 1 \end{pmatrix} , \quad \zeta_{2k} = \frac{k}{2(2k+1)} \begin{pmatrix} 1 & -\frac{1}{2k} \\ -\frac{1}{2k} & 1 \end{pmatrix} ,$$

which gives for the monic orthogonal polynomials with repsect to the matrix measure $d\mu(x)$ the following recursions

$$\underline{P}_0(x) = I_2 ,$$

$$\underline{P}_1(x) = x\underline{P}_0(x) - \zeta_1 = \begin{pmatrix} x-\frac{1}{2} & x-\frac{1}{4} \\ x-\frac{1}{2} & x-\frac{1}{2} \end{pmatrix} ,$$

$$\underline{P}_{k+1}(x) = (x-\frac{1}{2})\underline{P}_k(x) - \frac{1}{16}\underline{P}_{k-1}(x) , \; k \geq 1 .$$

5 Favard's Theorem for the interval $[0, \infty)$

Most of the material above can be extended to the interval $[0,\infty)$. All the results in Section 2 remain valid up to Theorem 2.5. If the matrix polynomial satisfies $P(t) \geq 0$ on $[0,\infty)$ then corresponding to Theorem 2.5 we have the following representation theorem for nonnegative definite matrix polynomials.

Theorem 5.1 : *If $P_n(t)$ is a nonnegative definite matrix polynomial of degreee n for all $t \in [0,\infty)$, then there exist matrix polynomials $B_j(t)$ and $C_j(t)$ of degreee j such that*

$$P_{2m}(t) = B_m(t)B_m(t)^T + tC_{m-1}(t)C_{m-1}(t)^T \tag{5.1}$$

if $n = 2m$ is even, and

$$P_{2m+1}(t) = tB_m(t)B_m(t)^T + C_m(t)C_m(t)^T \tag{5.2}$$

if $n = 2m+1$ is odd.

Proof: If $P(t) \geq 0$ on $[0,\infty)$ then certainly $P(t) \geq 0$ on $[0,b]$. We can rewrite the result from Theorem 2.5 by a change of variable to get, for the case $n = 2m$ for example, that

$$P_{2m}(t) = B_m(t)B_m(t)^T + \frac{t}{b}(1-\frac{t}{b})C_{m-1}(t)C_{m-1}(t)^T$$

where the polynomials B_m and C_{m-1} depend on b. We want to consider the case $b \to \infty$. To do this we note that

$$c^T P_{2m}(t)c \geq c^T B_m(t) B_m(t)^T c$$

for all vectors $c \in \mathbb{R}^p$. Letting c be the vector with a one in a single component it follows that all the coefficients of the matrix polynomial $B_m(t)$ are bounded and we can take convergent subsequences. Similarly, we obtain a subsequence for the part $b^{-1} C_{m-1}(t) C_{m-1}(t)^T$ and considering the limit for this subsequence yields the representation (5.1), which proves the assertion for $n = 2m$ even. The odd case $n = 2m + 1$ is similar and therefore omitted. □

Theorem 5.2:
a) The point $(S_0, S_1, \cdots, S_n)$ is contained in the moment space M_{n+1} on the interval $[0, \infty)$ if and only if $\underline{H}_n$ is nonnegative definite.

b) The point $(S_0, S_1, \cdots, S_n)$ is contained in the interior of moment space M_{n+1} on the interval $[0, \infty)$ if and only if $\underline{H}_n$ is positive definite.

If the space of consideration is the interval $[0, \infty)$ a lower bound for the moments S_k is available but not an upper bound. The lower bound is precisely as before, that is

$$S_{n+1} - S_{n+1}^- = \underline{h}_n^T \underline{H}_{n-1}^{-1} \underline{h}_n \ \geq \ 0 \tag{5.3}$$

provided that $\underline{H}_{n-1}$ is nonsingular. The quantity

$$\zeta_k = (S_{k-1} - S_{k-1}^-)^{-1} (S_k - S_k^-)$$

is defined as in (2.33) and we can trivially write

$$S_n - S_n^- = S_0 \zeta_1 \zeta_1 \cdots \zeta_n > 0 \tag{5.4}$$

if $(S_0, \ldots, S_n)$ is in the interior of the moment space M_{n+1}. The systems of polynomials $\underline{P}_m(t)$ and $\underline{Q}_m(t)$ are defined as before and the analysis regarding these is the same, except no mention is now made of the canonical moments P_k. Thus we have the following refinement of Favards's theorem on the nonnegative line $[0, \infty)$.

Theorem 5.3:
a) If $\{\underline{P}_n\}_{n\geq 1}$ is the sequence of monic polynomials which are orthogonal with respect to the matrix measure μ on the interval $[0, \infty)$, then the sequence satisfies the recurrence formula (3.11) where the coefficients ζ_n satisfy (5.4)

b) Conversely, if the sequence of monic polynomials $\{\underline{P}_n\}_{n\geq 1}$ satisfies the recurrence formula (3.11), where the coefficients ζ_n satisfy (5.4) with $S_0 = I$, then there exists a matrix measure μ supported on the interval $[0, \infty)$ for which the polynomials $\{\underline{P}_n\}_{n\geq 1}$ are the corresponding monic orthogonal polynomials.

Proof: The proof is very similar to the proof of Theorem 4.1. Part a) is again immediate. For part b) we define the moment sequence by letting $S_0 = I$, $S_1 = \zeta_1$ and successively define

$$S_n = S_n^- + \zeta_1 \zeta_2 \cdots \zeta_n$$

The conditions in (5.4) insure that for each n, $S_n > S_n^-$ and the proof proceeds as in Theorem 4.1. □

Acknowledgements. This work was done while H. Dette was visiting the Department of Statistics, Purdue University, and this author would like to thank the department for its hospitality. The work of H. Dette was supported by the Deutsche Forschungsgemeinschaft (De 502/14-1, SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen). The authors would also like to thank I. Gottschlich, who typed parts of this paper with considerable technical expertise.

References

A.I. Aptekarev, E.M. Nikishin (1983). The scattering problem for a discrete Sturm-Liouville operator. Mat. Sb. 121 (163), 327-358; Math. U. Sb. 49, 325-355, 1984.

G-N. Chen, Z-Q. Li (1999). The Nevanlinna-Pick interpolation problems and power moment problems for matrix-valued functions, Linear Algebra and its Applications 288, 123-148.

T.S. Chihara (1978). An Introuction to Orthogonal Polynomials. Gordon and Breach, NY.

E. Defez, L. Jódar. A. Law, E. Ponsoda (2000). Three-term recurrences and matrix orthogonal polynomials. Utilitas Mathematic 57, 129-146.

H. Dette, W.J. Studden (1997). The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis, Wiley.

A.J. Duran (1995). On orthogonal polynomials with respect to a positive definite matrix of measures. Can. J. Math. 47, 88-112.

A.J. Duran (1996). Markov's Theorem for orthogonal matrix polynomials, Can. J. Math. 48, 1180-1195.

A.J. Duran (1999). Ratio asymptotics for orthogonal matrix polynomials. Journal of Approximation Theory 100, 304-344.

A.J. Duran, P. Lopez-Rodriguez (1996). Orthogonal matrix polynomials: zeros and Blumenthal's Theorem. Journal of Approximation Theory 84, 96-118.

A.J. Duran, P. Lopez-Rodriguez (1997). Density questions for the truncated matrix moment problem. Can. J. Math. 49, 708-721.

A. J. Duran, W. Van Assche (1995). Orthogonal matrix polynomials and higher-order recurrence relations. Linear Algebra and its Applications 219, 261-280.

J. Favard (1935). Sur les polynomes de Tchebicheff. C.R. Acad. Sci. Paris 200, 2052-2053.

J.S. Geronimo (1982). Scattering theory and matrix orthogonal polynomials on the real line. Circuits Systems Signal Process 1, 471-495.

I. Gohberg, P. Lancaster, L. Rodman (1986). Quadratic matrix polynomials with a parameter. Adv. Appl. Math. 7, 253-281.

I. Gohberg, P. Lancaster, L. Rodman (1982). Matrix polynomials. Computer Science and Applied Mathematics.Academic Press, New York.

S. Karlin, W.J. Studden (1966). Tchebycheff Systems: with Applications in Analysis and Statistics. Wiley, New York.

M. G. Krein (1949). Fundamental aspects of the representation theory of Hermitian operators with dificiency index (m,m), Ukrain. Mat. Sh. 1, 3-66; Amer. Math. Soc. Transl. (2) 97, 75-143(1971).

P. Lancaster, M. Tismenetsky (1985). The Theory of Matrices. Academic Press, Inc., New York.

P. Lopez-Rodriguez (1999). Riesz's Theorem for orthogonal matrix polynomials. Constr. Approx. 15, 135-151.

A.N. Malyshev (1982). Factorization of matrix polynomials. Sibirskii Matematicheskii Zhurnal 23, 136-146.

F. Marcellán, G. Sansigre (1993). On a class of Matrix orthogonal polynomials on the real line. Linear Algebra and its Applications 181, 97-109.

L. Rodman (1990). Orthogonal matrix polynomials. In: P. Nevai (ed.), Orthogonal polynomials: theory and practice. NATO ASI Series C, Vol. 295; Kluwer, Dordrecht.

R.T. Rockafellar (1970). Convex Analysis. Princeton University Press; Princeton, NJ.

M. Rosenblatt (1956). A multidimensional prediction problem. Arkiv. Mat. 37, 407-424.

A. Sinap, W. Van Assche (1994). Polynomial interpolation and Gaussian quadrature for matrix-valued functions, Linear Algebra and its Applications 207, 71-114.

A. Sinap, W. Van Assche (1996). Orthogonal Matrix polynomials and applications. Jour. Computational and Applied Math. 66: 27-52.

M. Skibinsky (1967). The range of the $(n+1)$th moment for distributions on $[0,1]$. J. Appl. Prob. 4, 543-552.

M. Skibinsky (1969). Some striking properties of Binomial and Beta moments. Ann. Math. Statist. 40, 1753-1764.

M. Skibinsky (1986). Principal representations and canonical moment sequences for distributions on an interval. J. Math. Anal. Appl. 120, 95-120.

G. Szegö (1959). Orthogonal olynomials. Amer. Math. Soc. Colloqu. Public. 23. Providence, RI.

M. Tismenetsky (1993). Matrix generalizations of a moment problem theorem, I. The Hermitian case. SIAM J. Matrix Anal. Appl. 14, 92-112.

W. Van Assche (1993). Weak convergence of orthogonal polynomials. Indag. Mathem., N.S., 6(1), 7-23.

H.S. Wall (1948). Analytic Teory of Continued Fractions. VanNostrand, NY.