TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 342, Number 2, April 1994

BLOCK JACOBI MATRICES AND ZEROS OF MULTIVARIATE ORTHOGONAL POLYNOMIALS

YUAN XU

ABSTRACT. A commuting family of symmetric matrices are called the block Jacobi matrices, if they are block tridiagonal. They are related to multivariate orthogonal polynomials. We study their eigenvalues and joint eigenvectors. The joint eigenvalues of the truncated block Jacobi matrices correspond to the common zeros of the multivariate orthogonal polynomials.

1. INTRODUCTION

The Jacobi matrix has been studied extensively in operator theory, and it plays an important role in the study of orthogonal polynomials of one variable. It is well known that every orthonormal polynomial sequence $\{p_n\}_{n=0}^{\infty}$ satisfies a three-term relation

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1}p_{n-1}(x),$$

and thus, can be associated to a Jacobi matrix

$$J = \begin{bmatrix} b_0 & a_0 & & 0 \\ a_0 & b_1 & a_1 & \\ & a_1 & b_2 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix}$$

that acts as an operator on l^2 (cf. Stone [15]). The truncated matrix T_n is the $n \times n$ main submatrix of T. It is known that the eigenvalues of T_n correspond to the zeros of p_n, and the eigenvalues of operator T can be characterized through p_n.

In this paper we extend this correspondence between the zeros of orthogonal polynomials and the eigenvalues to the multivariable setting. Unlike the univariate case, where p_n always has n distinct real zeros, the existence of zeros is difficult to establish, as the zeros of multivariate orthogonal polynomials mean the common zeros of a family of polynomials. These common zeros have been studied and used in the construction of minimum cubatures (cf. [10, 13, 16]). Our approach is based on our recent study of the multivariate orthogonal polynomials in [18, 19]. A system of multivariate orthogonal polynomials is

Received by the editors January 31, 1992 and, in revised form, August 21, 1992.
1991 *Mathematics Subject Classification*. Primary 42C05, 47B15, 65D30.
Key words and phrases. Block Jacobi matrices, truncated block Jacobi matrices, joint eigenvectors, common zeros of multivariate orthogonal polynomials, commuting selfadjoint operators.

characterized by a three-term relation of vector-matrix form, and connected to a commuting family of symmetric operators that have block Jacobi matrix representation. If the coefficient matrices in the three-term relation are bounded in the spectral norm, then the operators are also bounded and selfadjoint. In this case, the spectral theorem for a commuting family of selfadjoint operators can be applied to establish the existence of the orthogonality measure. The study of the eigenvalues of these block Jacobi matrices will help us to understand the structure of this measure. Moreover, the eigenvalues of the truncated block Jacobi matrices correspond to the common zeros of orthogonal polynomials.

In §2, we introduce the notation and present the earlier results that will be needed. In §3 we establish the connection between the eigenvalues and common zeros, in particular, an alternative proof and the extension of a fundamental result in Mysovskikh [9] will be given. The eigenvalues of block Jacobi matrices will be discussed in §4, where an interesting identity of orthogonal polynomials is extended to the multivariable setting and used to analyse the eigenvalues in a special case.

2. Preliminaries

Let $\mathbb{N}_0$ be the set of nonnegative integers. For $n \in \mathbb{N}_0$ we denote by Π_n^d the set of polynomials of total degree at most n in d variables, and Π^d the set of all polynomials in d variables. Let $\mathscr{L}$ be a linear functional defined on Π^d such that $\mathscr{L}(g^2) > 0$ whenever $g \in \Pi^d$ and $g \neq 0$. Such an $\mathscr{L}$ is called square positive. Two polynomials P and Q are said to be orthogonal with respect to $\mathscr{L}$ if $\mathscr{L}(PQ) = 0$. For each $k \in \mathbb{N}_0$, let $V_k^d \subset \Pi_k^d$ denote the set of polynomials of degree exactly equal to k, together with zero, that are orthogonal to all polynomials in Π_{k-1}^d. Then V_k is a vector space of dimension $r_k^d = \binom{k+d-1}{k}$, and the V_k's are mutually orthogonal. Throughout this paper, the letter d is reserved for the number of variables or the dimension. It is fixed and will be omitted sometimes.

For a sequence of polynomials $\{P_j^k\}_{j=1}^{r_k}$, where the superscript k means that P_j^k is of total degree k, we introduce the vector notation

$$\mathbb{P}_k(\mathbf{x}) = [P_1^k(\mathbf{x}), P_2^k(\mathbf{x}), \dots, P_{r_k}^k(\mathbf{x})]^T. \tag{2.1}$$

For our convenience, if $\{P_j^k\}_{j=1}^{r_k}$ is a basis for V_k^d, we shall say that $\mathbb{P}_k$ is a basis for V_k^d and that $\{\mathbb{P}_k\}_{k=0}^{\infty}$ is an orthogonal basis for Π^d. We shall consider only the case of an orthonormal basis $\{\mathbb{P}_n\}_{n=0}^{\infty}$, which is an orthogonal basis for which $\mathscr{L}(P_i^k P_j^k) = \delta_{ij}$ for $k = 0, 1, \dots, i$, $j = 1, \dots, r_k$. Throughout this paper, the $n \times n$ identity matrix is denoted by I_n, or simply I. The notation $A\colon i \times j$ means that A is a matrix of size $i \times j$. For $\mathbf{x} \in \mathbb{R}^d$ we write $\mathbf{x} = (x_1, \dots, x_d)^T$.

Let $\{\mathbb{P}_n\}_{n=0}^{\infty}$ be an orthonormal basis of Π. The properties of $\mathbb{P}_n$ that will be needed later are listed as follows (cf. [18, 19]).

1°. *Three-term relation*: For $k \geq 0$, $1 \leq i \leq d$, there exist matrices $A_{k,i}\colon r_k \times r_{k+1}$ and $B_{k,i}\colon r_k \times r_k$, such that

$$x_i \mathbb{P}_k = A_{k,i}\mathbb{P}_{k+1} + B_{k,i}\mathbb{P}_k + A_{k-1,i}^T \mathbb{P}_{k-1}, \qquad 1 \leq i \leq d, \tag{2.2}$$

where $\mathbb{P}_{-1} = 0$, $\mathbb{P}_0 = 1$, and $A_{-1,i}$ is taken to be zero.

2°. *Favard's Theorem*. Let $\{\mathbb{P}_n\}_{n=0}^{\infty}$, $\mathbb{P}_0 = 1$, be a sequence in Π^d. Then the following statements are equivalent:

(1) There exists a linear functional which is square positive and makes $\{\mathbb{P}_n\}_{k=0}^{\infty}$ an orthonormal basis in Π^d.

(2) For $k \ge 0$, $1 \le i \le d$, $\mathbb{P}_n$ satisfies the three-term relation (2.2) and

$$\operatorname{rank} A_k = r_{k+1}, \qquad A_k = (A_{k,1}^T|, \dots, |A_{k,d}^T)^T. \tag{2.3}$$

The equality (2.3) implies that there exist matrices $D_{k,i}$: $r_{k+1} \times r_k$ such that

$$\sum_{i=1}^{d} D_{k,i} A_{k,i} = I. \tag{2.4}$$

For a sequence of orthonormal polynomials, the coefficient matrices in the three-term relation satisfy the following equations.

3°. *Equations for the coefficient matrices.*

$$A_{k,i}A_{k+1,j} = A_{k,j}A_{k+1,i}, \tag{2.5}$$

$$A_{k,i}B_{k+1,j} + B_{k,i}A_{k,j} = B_{k,j}A_{k,i} + A_{k,j}B_{k+1,i}, \tag{2.6}$$

$$\begin{aligned} &A_{k-1,i}^T A_{k-1,j} + B_{k,i}B_{k,j} + A_{k,i}A_{k,j}^T \\ &\quad = A_{k-1,j}^T A_{k-1,i} + B_{k,j}B_{k,i} + A_{k,j}A_{k,i}^T, \end{aligned} \tag{2.7}$$

for $i \ne j$, $1 \le i, j \le d$, and $k \ge 0$, where $A_{-1,i} = 0$.

4°. *Christoffel-Darboux formula.*

$$\sum_{k=0}^{n-1} \mathbb{P}_k^T(\mathbf{x})\mathbb{P}_k(\mathbf{y}) = \frac{[A_{n-1,i}\mathbb{P}_n(\mathbf{x})]^T\mathbb{P}_{n-1}(\mathbf{y}) - \mathbb{P}_{n-1}^T(\mathbf{x})[A_{n-1,i}\mathbb{P}_n(\mathbf{y})]}{x_i - y_i}$$

for $1 \le i \le d$.

Let $A_{k,i}$: $r_k \times r_{k+1}$ and $B_{k,i}$: $r_k \times r_k$ be given matrices such that the rank condition (2.3) is satisfied. Furthermore, assume that $A_{k,i}$ and $B_{k,i}$ satisfy equations (2.5), (2.6), and (2.7). We then define linear operators T_i on l^2 whose matrix representation are given by

$$T_i = \begin{bmatrix} B_{0,i} & A_{0,i} & & \mathbf{0} \\ A_{0,i}^T & B_{1,i} & A_{1,i} & \\ & A_{1,i}^T & B_{2,i} & \ddots \\ \mathbf{0} & & \ddots & \ddots \end{bmatrix}, \qquad 1 \le i \le d. \tag{2.8}$$

We shall call T_i block Jacobi matrices. For $d = 1$, we have $r_k = 1$ for all $k \in \mathbb{N}_0$ and the matrix T_1 is the Jacobi matrix. The truncated block Jacobi matrices are denoted by $T_{n,i}$, and are obtained from T_i by deleting block rows and block columns with numbers $\ge n$. Thus $T_{n,i}$ has n block rows and block columns, $n = 1, 2, \dots$. The size of $T_{n,i}$ is $\binom{n-1+d}{n-1} \times \binom{n-1+d}{n-1}$. We note that the size of elements $A_{n,i}$ and $B_{n,i}$ increase with n.

3. Common zeros of orthogonal polynomials

By zeros of orthogonal polynomials we mean the zeros of $\mathbb{P}_n$, i.e. the common zeros of the components in $\mathbb{P}_n$. Clearly, they can also be considered as zeros of the subspace V_k, which is why we consider only orthonormal bases. If $\mathbf{x}$ is a zero of $\mathbb{P}_n$ and at least one partial derivative of $\mathbb{P}_n$ at $\mathbf{x}$ is not zero, then we say that $\mathbf{x}$ is a simple zero of $\mathbb{P}_n$.

First we derive two elementary properties of zeros from the Christoffel-Darboux formula.

Theorem 3.1. *All zeros of $\mathbb{P}_n$ are distinct and simple. Two consecutive polynomials $\mathbb{P}_n$ and $\mathbb{P}_{n-1}$ do not have common zeros.*

Proof. From the Christoffel-Darboux formula, one has by taking limit (cf. [18]),

$$\sum_{k=0}^{n-1} \mathbb{P}_k^T(\mathbf{x})\mathbb{P}_k(\mathbf{x}) = \mathbb{P}_{n-1}^T(\mathbf{x})A_{n-1,i}\partial_i\mathbb{P}_n(\mathbf{x}) - \mathbb{P}_n^T(\mathbf{x})A_{n-1,i}^T\partial_i\mathbb{P}_{n-1}(\mathbf{x}),$$

where $\partial_i = \partial/\partial x_i$ denotes the partial derivative with respect to x_i. If $\mathbf{x}$ is a zero of $\mathbb{P}_n$, then we have

$$\sum_{k=0}^{n-1} \mathbb{P}_k^T(\mathbf{x})\mathbb{P}_k(\mathbf{x}) = \mathbb{P}_{n-1}^T(\mathbf{x})A_{n-1,i}\partial_i\mathbb{P}_n(\mathbf{x}).$$

Since the left-hand side is positive, neither $\mathbb{P}_{n-1}(\mathbf{x})$ nor $\partial_i\mathbb{P}_n(\mathbf{x})$ can be zero. □

In the following, we shall denote by T_n the tube of matrices $T_n = (T_{n,1}, \dots, T_{n,d})$. If $\mathbf{x}$ is an eigenvector for all $T_{n,i}$, i.e. $T_{n,i}\mathbf{x} = \lambda_i\mathbf{x}$ for $1 \le i \le d$, then we call $\mathbf{x}$ a joint eigenvector of $(T_{n,1}, \dots, T_{n,d})$, or simply an eigenvector of T_n, and call $\Lambda = (\lambda_1, \dots, \lambda_d)$ the eigenvalue of T_n.

Our main result in this section is the following.

Theorem 3.2. *Let $\{\mathbb{P}_n\}$ be orthonormal polynomials, and $T_{n,i}$, $1 \le i \le d$, be the corresponding truncated block Jacobi matrices. Then $\Lambda = (\lambda_1, \dots, \lambda_d)$ is a zero of $\mathbb{P}_n$ if and only if Λ is an eigenvalue of T_n that has a joint eigenvector. Moreover, the joint eigenvector is given by $(\mathbb{P}_0^T(\Lambda), \dots, \mathbb{P}_{n-1}^T(\Lambda))^T$.*

Proof. If $\mathbb{P}_n(\Lambda) = 0$, then it follows from the three-term relation that

$$\begin{aligned} &B_{0,i}\mathbb{P}_0(\Lambda) + A_{0,i}\mathbb{P}_1(\Lambda) = \lambda_i\mathbb{P}_0(\Lambda), \\ &A_{k-1,i}^T\mathbb{P}_{k-1}(\Lambda) + B_{k,i}\mathbb{P}_k(\Lambda) + A_{k,i}\mathbb{P}_{k+1}(\Lambda) = \lambda_i\mathbb{P}_k(\Lambda), \qquad 1 \le k \le n-2, \\ &A_{n-2,i}^T\mathbb{P}_{n-2}(\Lambda) + B_{n-1,i}\mathbb{P}_{n-1}(\Lambda) = \lambda_i\mathbb{P}_{n-1}(\Lambda) \end{aligned}$$

for $1 \le i \le d$. From the definition of $T_{n,i}$ it follows then

$$T_{n,i}\mathbf{x} = \lambda_i\mathbf{x}, \qquad \mathbf{x} = [\mathbb{P}_0^T(\Lambda), \dots, \mathbb{P}_{n-1}^T(\Lambda)]^T.$$

Thus, Λ is the eigenvalue of T_n with joint eigenvector $\mathbf{x}$.

On the other hand, suppose that $\Lambda = (\lambda_1, \dots, \lambda_d)$ is an eigenvalue of T_n, and T_n has a joint eigenvector $\mathbf{x}$ for Λ. We write

$$\mathbf{x} = (\mathbf{x}_0^T, \dots, \mathbf{x}_{n-1}^T)^T, \qquad \mathbf{x}_j \in \mathbb{R}^{r_j}.$$

Since $T_{n,i}\mathbf{x} = \lambda_i\mathbf{x}$, it follows that $\{\mathbf{x}_j\}$ satisfies a three-term relation

$$\begin{aligned} &B_{0,i}\mathbf{x}_0 + A_{0,i}\mathbf{x}_1 = \lambda_i\mathbf{x}_0, \\ &A_{k-1,i}^T\mathbf{x}_{k-1} + B_{k,i}\mathbf{x}_k + A_{k,i}\mathbf{x}_{k+1} = \lambda_i\mathbf{x}_k, \qquad 1 \le k \le n-2, \\ &A_{n-2,i}^T\mathbf{x}_{n-2} + B_{n-1,i}\mathbf{x}_{n-1} = \lambda_i\mathbf{x}_{n-1} \end{aligned}$$

for $1 \le i \le d$. First we show that $\mathbf{x}_0$, thus $\mathbf{x}$, is not zero. Indeed, if $\mathbf{x}_0 = 0$, then we have from the first equation in the three-term relation that $A_{0,i}\mathbf{x}_1 = 0$. But then $A_0\mathbf{x}_1 = 0$, and A_0 at (2.3) has full rank, so it follows that $\mathbf{x}_1 = 0$. With $\mathbf{x}_0 = 0$ and $\mathbf{x}_1 = 0$, it follows from the three-term relation that $A_{1,i}\mathbf{x}_2 = 0$, which leads to $\mathbf{x}_2 = 0$. Similarly, we have $\mathbf{x}_i = 0$ for $i \ge 3$. Thus, we have $\mathbf{x} = 0$, which contradicts the assumption that $\mathbf{x}$ is an eigenvector. Let us assume that $\mathbf{x}_0 = 1 = \mathbb{P}_0$ and define $\mathbf{x}_n \in \mathbb{R}^{r_n}$ as $\mathbf{x}_n = 0$. We now prove that $\mathbf{x}_j = \mathbb{P}_j(\Lambda)$ for all $1 \le j \le n$. Since the last equation in the three-term relation of $\mathbf{x}_j$ can be written as

$$A_{n-2,i}^T\mathbf{x}_{n-2} + B_{n-1,i}\mathbf{x}_{n-1} + A_{n-1,i}\mathbf{x}_n = \lambda_i\mathbf{x}_{n-1},$$

we see that $\{\mathbf{x}_k\}_{k=0}^n$ and $\{\mathbb{P}_k(\Lambda)\}_{k=0}^n$ satisfy the same three-term relation. Thus, so does $\{\mathbf{y}_k\} = \{\mathbb{P}_k(\Lambda) - \mathbf{x}_k\}$. But since $\mathbf{y}_0 = 0$, it follows from the previous argument that $\mathbf{y}_k = 0$ for all $1 \le k \le n$. In particular, $\mathbf{y}_n = \mathbb{P}_n(\Lambda) = 0$. The proof is completed. □

For the corresponding result for the Jacobi matrix $(d = 1)$, we refer to [15, p. 532] or [2, p. 30]. From this theorem follows several interesting corollaries.

Corollary 3.3. *The multivariate orthogonal polynomial $\mathbb{P}_n$ has at most $N := \dim \Pi_{n-1}$ distinct zeros. All zeros of $\mathbb{P}_n$ are real.*

This follows from the fact that there are at most $\binom{n-1+d}{n-1} = \dim \Pi_{n-1}$ joint eigenvectors of T_n and that $T_{n,i}$ are symmetric matrices. These two properties of the zeros are known (at least for $d = 2$, cf. [9, 13]).

Theorem 3.4. *The polynomial $\mathbb{P}_n$ has $N = \dim \Pi_{n-1}$ distinct zeros if and only if*

$$A_{n-1,i}A_{n-1,j}^T = A_{n-1,j}A_{n-1,i}^T \tag{3.1}$$

for all $1 \le i, j \le d$, where the $A_{n-1,i}$'s are the coefficient matrices in the three-term recurrence relation.

Proof. From Theorem 3.2 it follows that $\mathbb{P}_n$ has N distinct zeros if and only if $T_{n,1}, \ldots, T_{n,d}$ have N distinct linearly independent eigenvectors, since the eigenvectors belonging to different eigenvalues are orthogonal. This is equivalent to saying that $T_{n,1}, \ldots, T_{n,d}$ can be simultaneously diagonalized by an invertible matrix. Since a family of matrices is simultaneously diagonalizable if and only if it is a commuting family, we have $T_{n,i}T_{n,j} = T_{n,j}T_{n,i}$ for all $1 \le i, j \le d$. From the definition of $T_{n,i}$, (2.5), (2.6), and (2.7), this is equivalent to the condition

$$A_{n-2,i}^T A_{n-2,j} + B_{n-1,i}B_{n-1,j} = A_{n-2,j}^T A_{n-2,i} + B_{n-1,j}B_{n-1,i}.$$

The equation (2.6) then leads to the desired result. □

For the bivariate orthogonal polynomials $(d = 2)$, this result is due to Mysovskikh [9]. His condition is given in a form that involves the moment matrix of orthogonal polynomials (see [14] for its matrix form). But this is because he used the monic basis for V_k instead of an orthonormal basis. It is easy to see that the two conditions are really equivalent. Our form of condition enables us to prove this result for all $d \ge 2$. The proof given here is quite different from the one in [9].

The importance of this theorem lies in the existence of minimal cubature formulae. A linear functional

$$I_N(f) = \sum_{k=1}^{N} \lambda_k f(\mathbf{x}_k), \qquad \lambda_k > 0,\, \mathbf{x}_k \in \mathbb{R}^d$$

is called a cubature formula of degree m, if $\mathscr{L}(f) = I_N(f)$ whenever $f \in \Pi_m^d$, and $\mathscr{L}(f^*) \neq I_N(f^*)$ for at least one $f^* \in \Pi_{m+1}^d$. For fixed m a cubature with minimal number of nodes N is called a minimal cubature, or a Gaussian cubature. A lower bound for N is [16],

$$N \geq \dim \Pi_{[m/2]}^d. \tag{3.2}$$

For $d = 2$, Mysovskikh [9] proved the following important result. In order that there exists a cubature formula which is exact for polynomials in Π_{2n-1} and uses $N = \dim \Pi_{n-1}$ knots, it is necessary and sufficient that $\mathbb{P}_n$ has N distinct real common zeros. However, it is clear from (3.1) that the condition (3.2) with equal sign is not satisfied in general. This fact, not so obvious in Mysovskikh's form of condition, has been noticed by Möller and revealed in his deep study of the minimal cubature [6, 7] where the algebraic ideal theory is used. Our next result quantitates this fact. Let

$$\sigma_n = \max_{1 \leq i < j \leq d} \operatorname{rank}(A_{n,i}A_{n,j}^T - A_{n,j}A_{n,i}^T).$$

Theorem 3.5. *The multivariate orthogonal polynomials* $\mathbb{P}_n$ *has at most* $\binom{n-1+d}{n-1} - \sigma_{n-1}$ *distinct zeros.*

Proof. If $\Lambda \in \mathbb{R}^d$ is a zero of $\mathbb{P}_n$, then there is a joint eigenvector of $T_{n,1}, \dots, T_{n,d}$ corresponding to the eigenvalues $\lambda_1, \dots, \lambda_d$. It is also clear that two different zeros correspond to linearly independent eigenvectors. Therefore, the number of distinct zeros of $\mathbb{P}_n$ is at most equal to the number of linearly independent joint eigenvectors. However, that $\mathbf{x}$ is a joint eigenvector of T implies $\mathbf{x} \in \ker(T_{n,i}T_{n,j} - T_{n,j}T_{n,i})$ for $i \neq j$. Therefore, the number of linearly independent joint eigenvectors is at most equal to

$$\dim \ker(T_{n,i}T_{n,j} - T_{n,j}T_{n,i}) = N - \operatorname{rank}(T_{n,i}T_{n,j} - T_{n,j}T_{n,i}).$$

From (2.4), (2.5), and (2.6) it follows that

$$\operatorname{rank}(T_{n,i}T_{n,j} - T_{n,j}T_{n,i}) = \operatorname{rank}(A_{n-1,i}A_{n-1,j}^T - A_{n-1,j}A_{n-1,i}^T).$$

The proof is completed. □

For $d = 2$ the number σ_n also appears in Möller's improved bounds for the cubature formula of degree $2n+1$,

$$N \geq \dim \Pi_n + \sigma_n/2. \tag{3.3}$$

Note that σ_n is an even number since $T_{n,i}T_{n,j} - T_{n,j}T_{n,i}$ is skewsymmetric. Moreover, Möller proved that if $\mathscr{L}$ is centrally symmetric, i.e. $\mathscr{L}(x^i y^{k-i}) = 0$, $0 \leq i \leq k$, for all odd $k \in \mathbb{N}$, then $\sigma_n = 2[(n+1)/2]$. In this case, the nodes for the minimum cubature on the product region that attains the lower bound (3.3) are characterized by the zeros of quasi-orthogonal polynomials. For further references in this direction we refer to [8, 12, 13, 20]. It would be interesting to see whether the result in Theorem 3.5 has any application in the cubature problem.

4. The eigenvalues of block Jacobi matrix

Let $\mathscr{H}$ be a separable Hilbert space with fixed orthonormal basis $\{\psi_n\}_{n=0}^{\infty} = \{\phi_j^k\}_{j=1}^{r_k}{}_{k=0}^{\infty}$. We introduce the formal vector notation $\Phi_k = [\phi_1^k, \dots, \phi_{r_k}^k]^T$, $k \in \mathbb{N}_0$. For our convenience we shall say that $\{\Phi_n\}_{n=0}^{\infty}$ is orthonormal, and for every $f \in \mathscr{H}$ we can then write in the vector-matrix notation that

$$f = \sum_{k=0}^{\infty} \mathbf{a}_k^T \Phi_k, \qquad \mathbf{a}_k \in \mathbb{R}^{r_k}.$$

If $T: \mathscr{H} \to \mathscr{H}$ is a linear operator, we mean by $T\Phi_k$ the vector $T\Phi_k = [T\phi_1^k, \dots, T\phi_{r_k}^k]^T$. With this notation, the usual matrix representation of the linear operator takes the form $T = (F_{ij})$, where $F_{ij} = (\langle T\Phi_i, \Phi_j^T \rangle)$ are matrices of order $r_i \times r_j$. We have used the same symbol for both the operator and its matrix representation.

Let T_i, $1 \le i \le d$, be a family of operators defined on $\mathscr{H}$ that has the matrix representation at (2.8). We shall restrict ourself to the bounded operators. Let $\|\cdot\|_2$ be the spectral norm for matrices. It is induced by the Euclidean norm for vectors.

$$\|A\|_2 = \max\{\sqrt{\lambda}: \lambda \text{ is an eigenvalue of } A^T A\}.$$

In [19] it is proved that

Lemma 4.1. *The operator* T_i *is bounded if and only if*

$$\sup_{k\ge 0} \|A_{k,i}\|_2 < +\infty, \qquad \sup_{k \ge 0} \|B_{k,i}\|_2 < +\infty. \tag{4.1}$$

Moreover, when T_i, $1 \le i \le d$, *are bounded, they form a commuting family of selfadjoint operators.*

Using the spectral theorem for a commuting family of selfadjoint operators [1, 11], the following theorem is proved in [19].

Theorem 4.2. *Let* $\{\mathbb{P}_n\}_{n=0}^{\infty}$, $\mathbb{P}_0 = 1$, *be a sequence in* Π^d. *Then the following statements are equivalent*:

(i) *There exists a nonnegative Borel measure* μ *with compact support in* $\mathbb{R}^d$ *such that* $\{\mathbb{P}_n\}_{n=0}^{\infty}$ *is orthonormal with respect to* μ.

(ii) $\{\mathbb{P}_n\}$ *satisfies three-term relation* (2.2), *rank condition* (2.3), *and* (4.1).

Since the support set of μ is the joint spectrum of $T_1, \dots, T_d$, the study of the eigenvalues will help us to understand the structure of μ. For the univariate results see [1, 5, 15]. Let $T = (T_1, \dots, T_d)$. A vector $\Lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d$ is called an eigenvalue of T, or $\{T_i\}$, if there exists $\mathbf{x} \in \mathscr{H}$, $\mathbf{x} \ne 0$, such that $T_i\mathbf{x} = \lambda_i \mathbf{x}$ for $1 \le i \le d$. We now consider the eigenvalues of the operators $\{T_i\}$. Our first result is a generalization of a univariate result [15, p. 546].

Theorem 4.3. *Let* $\{\mathbb{P}_k\}$ *be the orthonormal polynomials corresponding to the operators* $T_1, \dots, T_d$. *Let* $\Lambda \in \mathbb{R}^d$. *Then* Λ *is an eigenvalue of* T *if and only if*

$$\sum_{k=0}^{\infty} \mathbb{P}_k^T(\Lambda)\mathbb{P}_k(\Lambda) < \infty.$$

Moreover, if $\mathbf{x}$ *is an eigenvector of* T, *then*

$$\mathbf{x} = \sum_{k=0}^{\infty} \mathbf{a}_k^T \Phi_k, \qquad \mathbf{a}_k = \mathbf{a}_0 \mathbb{P}_k(\Lambda),$$

where $\mathbf{a}_0 \in \mathbb{R}$ *is a constant.*

Proof. If $\mathbf{x} = \sum \mathbf{a}_k^T \Phi_k \in \mathscr{H}$, then from the definition of T_i it follows that

$$\begin{aligned} T_i \mathbf{x} &= \sum \mathbf{a}_k^T (A_{k-1,i}^T \Phi_{k-1} + B_{k,i} \Phi_k + A_{k,i} \Phi_{k+1}) \\ &= \sum (\mathbf{a}_{k+1}^T A_{k,i}^T + \mathbf{a}_k^T B_{k,i} + \mathbf{a}_{k-1}^T A_{k-1,i}) \Phi_k. \end{aligned}$$

Therefore, $T_i \mathbf{x} = \lambda_i \mathbf{x}$ implies that the vectors $\mathbf{a}_k$ satisfy the three-term relation

$$A_{k-1,i}^T \mathbf{a}_{k-1} + B_{k,i} \mathbf{a}_k + A_{k,i} \mathbf{a}_{k+1} = \lambda_i \mathbf{a}_k.$$

From Favard's theorem, we then have $\mathbf{a}_k = \mathbb{P}_k(\lambda_1, \dots, \lambda_d) \mathbf{a}_0$. Therefore, if $\mathbf{x}$ is an eigenvector of T, then $\mathbf{x} = \mathbf{a}_0 \sum \mathbb{P}_k^T(\Lambda) \Phi_k \neq 0$. Clearly, $\mathbf{x} \in \mathscr{H}$ if and only if $\sum \mathbb{P}_k(\Lambda)^T \mathbb{P}_k(\Lambda) < \infty$. □

In the following we derive a necessary condition for Λ to be an eigenvalue. The univariate analogy of this part is in [4]. We define now a family of new operators J_i on $\mathscr{H}$ that has matrix representation

$$J_i = \frac{1}{2} \begin{bmatrix} 0 & -A_{0,i} & & \mathbf{0} \\ A_{0,i}^T & 0 & -A_{1,i} & \\ & A_{1,i}^T & 0 & \ddots \\ \mathbf{0} & & \ddots & \ddots \end{bmatrix}, \qquad 1 \le i \le d. \tag{4.2}$$

Let now K_i be the operators defined by the equation

$$K_i = T_i J_i - J_i T_i, \qquad 1 \le i \le d.$$

It is easy to verify that K_i also has a block Jacobi matrix representation,

$$K_i = \begin{bmatrix} K_{0,i} & L_{0,i} & & \mathbf{0} \\ L_{0,i}^T & K_{1,i} & L_{1,i} & \\ & L_{1,i}^T & K_{2,i} & \ddots \\ \mathbf{0} & & \ddots & \ddots \end{bmatrix}, \qquad 1 \le i \le d, \tag{4.3}$$

where for $k \ge 0$

$$K_{k,i} = A_{k,i} A_{k,i}^T - A_{k-1,i}^T A_{k-1,i}, \qquad L_{k,i} = \tfrac{1}{2}(A_{k,i} B_{k+1,i} - B_{k,i} A_{k,i}).$$

Theorem 4.4. *If* Λ *is an eigenvalue of* T, *then for* $1 \le i \le d$,

$$\begin{aligned} &\mathbb{P}_0^T(\Lambda) A_{0,i} A_{0,i}^T \mathbb{P}_0(\Lambda) + \sum_{k=1}^{\infty} \mathbb{P}_k^T(\Lambda)(A_{k,i} A_{k,i}^T - A_{k-1,i}^T A_{k-1,i}) \mathbb{P}_k(\Lambda) \\ &\qquad + \sum_{k=1}^{\infty} \mathbb{P}_k^T(\Lambda)(A_{k,i} B_{k+1,i} - B_{k,i} A_{k,i}) \mathbb{P}_{k+1}(\Lambda) = 0. \end{aligned} \tag{4.4}$$

Proof. Since T_i is assumed to be bounded, and thus, selfadjoint, so is $T_i - \lambda_i I$. From this fact and $K_i = (T_i - \lambda_i I) J_i - J_i (T_i - \lambda_i I)$ it follows that $\langle K_i \mathbf{x}, \mathbf{x} \rangle = 0$

for an eigenvector $\mathbf{x}$, $1 \le i \le d$. If $\mathbf{x}$ is an eigenvector corresponding to Λ, then $\mathbf{x} = \mathbf{a}_0 \sum \mathbb{P}_k^T(\Lambda)\Phi_k$. Our claim follows from $\langle K_i\mathbf{x}, \mathbf{x}\rangle = 0$ and the definition of K_i at (4.3). □

The partial sum of the infinite series (4.4) satisfies an interesting identity (4.5). For the univariate result we refer to [3] and the references given there. Since its proof follows from the three-term relation, this identity is actually true for all orthogonal polynomials.

Theorem 4.5. *Let* $\{\mathbb{P}_k\}$ *be a system of orthonormal polynomials satisfying the three-term relation* (2.2). *Let*

$$S_{n,i}(\Lambda) = \mathbb{P}_0^T(\Lambda)A_{0,i}A_{0,i}^T\mathbb{P}_0(\Lambda) + \sum_{k=1}^{n} \mathbb{P}_k^T(\Lambda)(A_{k,i}A_{k,i}^T - A_{k-1,i}^T A_{k-1,i})\mathbb{P}_k(\Lambda)$$
$$+ \sum_{k=0}^{n-1} \mathbb{P}_k^T(\Lambda)(A_{k,i}B_{k+1,i} - B_{k,i}A_{k,i})\mathbb{P}_{k+1}(\Lambda).$$

Then for $1 \le i \le d$,

$$\begin{aligned} S_{n,i}(\Lambda) &= \mathbb{P}_{n-1}^T(\Lambda)A_{n-1,i}A_{n-1,i}^T\mathbb{P}_{n-1}(\Lambda) \\ &\quad - \mathbb{P}_{n-1}^T(\Lambda)A_{n-1,i}(\lambda_i I - B_{n,i})\mathbb{P}_n(\Lambda) + \mathbb{P}_n^T(\Lambda)A_{n,i}A_{n,i}^T\mathbb{P}_n(\Lambda). \end{aligned} \tag{4.5}$$

Proof. We use induction. For $n = 0$, we have for both expressions $S_{0,i}(\Lambda) = \mathbb{P}_0^T(\Lambda)A_{0,i}A_{0,i}^T\mathbb{P}_0(\Lambda)$ since $\mathbb{P}_{-1} = 0$. Assume now that the identity has been proved for $S_{n,i}(\Lambda)$, then

$$\begin{aligned} S_{n+1,i} &= S_{n,i} + [S_{n+1,i} - S_{n,i}] \\ &= \mathbb{P}_{n-1}^T A_{n-1,i}A_{n-1,i}^T\mathbb{P}_{n-1} - \mathbb{P}_{n-1}^T A_{n-1,i}(\lambda_i I - B_{n,i})\mathbb{P}_n + \mathbb{P}_n^T A_{n,i}A_{n,i}^T\mathbb{P}_n \\ &\quad + \mathbb{P}_{n+1}^T(A_{n+1,i}A_{n+1,i}^T - A_{n,i}^T A_{n,i})\mathbb{P}_{n+1} + \mathbb{P}_n^T(A_{n,i}B_{n+1,i} - B_{n,i}A_{n,i})\mathbb{P}_{n+1} \\ &= \mathbb{P}_{n-1}^T A_{n-1,i}A_{n-1,i}^T\mathbb{P}_{n-1} + \mathbb{P}_n^T A_{n,i}A_{n,i}^T\mathbb{P}_n + \mathbb{P}_{n+1}^T A_{n+1,i}A_{n+1,i}^T\mathbb{P}_{n+1} \\ &\quad - \mathbb{P}_{n-1}^T A_{n-1,i}(\lambda_i I - B_{n,i})\mathbb{P}_n + \mathbb{P}_n^T(A_{n,i}B_{n+1,i} - B_{n,i}A_{n,i})\mathbb{P}_{n+1} \\ &\quad - [(\lambda_i I - B_{n,i})\mathbb{P}_n - A_{n-1,i}^T\mathbb{P}_{n-1}]^T[(\lambda_i I - B_{n,i})\mathbb{P}_n - A_{n-1,i}^T\mathbb{P}_{n-1}] \end{aligned}$$

where we have used the three-term relation. Simplify the above formula, we get

$$\begin{aligned} S_{n+1,i} &= \mathbb{P}_n^T A_{n,i}A_{n,i}^T\mathbb{P}_n + \mathbb{P}_{n+1}^T A_{n+1,i}A_{n+1,i}^T\mathbb{P}_{n+1} \\ &\quad - \mathbb{P}_n^T(\lambda_i I - B_{n,i})^T A_{n,i}\mathbb{P}_{n+1} + \mathbb{P}_n^T(A_{n,i}B_{n+1,i} - B_{n,i}A_{n,i})\mathbb{P}_{n+1} \\ &= \mathbb{P}_n^T A_{n,i}A_{n,i}^T\mathbb{P}_n - \mathbb{P}_n^T A_{n,i}(\lambda_i I - B_{n+1,i})\mathbb{P}_{n+1} + \mathbb{P}_{n+1}^T A_{n+1,i}A_{n+1,i}^T\mathbb{P}_{n+1}, \end{aligned}$$

since $B_{n,i}$ is symmetric. The induction is completed. □

For univariate orthogonal polynomials, this identity has been used to derive information about the eigenvalues of the operator, see [3, 4]. It has also been used for deriving properties of orthogonal polynomials with respect to a given weight function. The expression of $S_{n,i}$ at (4.5) can be written in another form, which is particularly attractive in the case of $B_{n,i} = 0$.

Corollary 4.6. *We have*

$$S_{n,i}(\Lambda) = \left(A_{n-1,i}^T\mathbb{P}_{n-1}(\Lambda) - \frac{\lambda_i}{2}\mathbb{P}_n(\Lambda)\right)^T \left(A_{n-1,i}^T\mathbb{P}_{n-1}(\Lambda) - \frac{\lambda_i}{2}\mathbb{P}_n(\Lambda)\right)$$
$$+ \mathbb{P}_n^T(\Lambda)\left(A_{n,i}A_{n,i}^T - \frac{\lambda_i^2}{4}I\right)\mathbb{P}_n(\Lambda) + \mathbb{P}_{n-1}^T(\Lambda)A_{n-1,i}B_{n,i}\mathbb{P}_n(\Lambda).$$

In particular, if $B_{n,i} = 0$,

$$S_{n,i}(\Lambda) \geq \mathbb{P}_n^T(\Lambda)\left(A_{n,i}A_{n,i}^T - \frac{\lambda_i^2}{4}I\right)\mathbb{P}_n(\Lambda). \tag{4.6}$$

We now use Theorems 4.4 and 4.5 to analyse the eigenvalues of T in one particular case. Let $\{p_n\}$ be a sequence of univariate orthonormal polynomials that satisfies the three-term relation

$$xp_n(x) = a_n p_{n+1}(x) + a_{n-1}p_{n-1}(x),$$

and $\{q_n\}$ be a sequence of univariate orthonormal polynomials with

$$xq_n(x) = b_n q_{n+1}(x) + b_{n-1}q_{n-1}(x).$$

Let $P_j^k(x_1, x_2) = p_{j-1}(x_1)q_{k-j+1}(x_2)$, $k \geq 0$, $1 \leq j \leq k+1$, be bivariate orthonormal polynomials. It is easy to verify that $\mathbb{P}_k$ satisfies the three-term relation

$$x_i\mathbb{P}_k = A_{k,i}\mathbb{P}_{k+1} + A_{k-1,i}^T\mathbb{P}_{k-1}, \qquad k \geq 0,\ i = 1, 2,$$

where

$$A_{k,1} = [O|\operatorname{diag}(a_0, a_1, \dots, a_k)], \qquad A_{k,2} = [\operatorname{diag}(b_k, b_{k-1}, \dots, b_0)|O].$$

Let T_1 and T_2 be operators defined by the corresponding block Jacobi matrices.

Theorem 4.7. *Suppose both sequences* $\{a_n\}$ *and* $\{b_n\}$ *are monotone decreasing and satisfy the conditions*

$$a_{n-1}^2 - 2a_n^2 + a_{n+1}^2 \leq 0, \qquad b_{n-1}^2 - 2b_n^2 + b_{n+1}^2 \leq 0. \tag{4.7}$$

Moreover, suppose $\lim a_n = a$ *and* $\lim b_n = b$. *Then* $T = (T_1, T_2)$ *has no eigenvalues in* $(-2a, 2a) \times (-2b, 2b)$.

Proof. Suppose that $\Lambda = (\lambda_1, \lambda_2)$ is an eigenvalue of T and $\Lambda \in (-2a, 2a) \times (-2b, 2b)$. Since $\sum \mathbb{P}_k^T(\Lambda)\mathbb{P}_k(\Lambda) < \infty$ by Theorem 4.3, we let N be the integer such that

$$\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda) = \max_{k\geq 0} \mathbb{P}_k^T(\Lambda)\mathbb{P}_k(\Lambda).$$

From the definition of $A_{k,1}$ it follows that

$$A_{k,1}A_{k,1}^T = \operatorname{diag}(a_0^2, a_1^2, \dots, a_k^2)$$

and

$$A_{k,1}A_{k,1}^T - A_{k-1,1}^T A_{k-1,1} = \operatorname{diag}(a_0^2, a_1^2 - a_0^2, \dots, a_k^2 - a_{k-1}^2).$$

Since it follows from (4.7) that $0 \geq a_j^2 - a_{j-1}^2 \geq a_{j+1}^2 - a_j^2$, we have

$$\begin{aligned}
&\mathbb{P}_k^T(\Lambda)(A_{k,1}A_{k,1}^T - A_{k-1,1}^T A_{k-1,1})\mathbb{P}_k(\Lambda) \\
&\quad = \sum_{j=0}^{k} p_j^2(\lambda_1)q_{k-j}^2(\lambda_2)(a_j^2 - a_{j-1}^2) \geq (a_k^2 - a_{k-1}^2)\mathbb{P}_k^T(\Lambda)\mathbb{P}_k(\Lambda).
\end{aligned}$$

Since a_k is decreasing, from Corollary 4.6, we have

$$S_{N,1} \geq \mathbb{P}_N^T(\Lambda)\left(A_{N,1}A_{N,1}^T - \frac{\lambda_1^2}{4}I\right)\mathbb{P}_N(\Lambda) \geq \left(a_N^2 - \frac{\lambda_1^2}{4}\right)\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda).$$

We then have for any $m > N$,

$$\begin{aligned} S_{m,1}(\Lambda) &= S_{N,1}(\Lambda) + \sum_{k=N+1}^{m} \mathbb{P}_k^T(\Lambda)(A_{k,1}A_{k,1}^T - A_{k-1,1}^T A_{k-1,1})\mathbb{P}_k(\Lambda) \\ &\geq S_{N,1}(\Lambda) + \sum_{k=N+1}^{m} (a_k^2 - a_{k-1}^2)\mathbb{P}_k^T(\Lambda)\mathbb{P}_k(\Lambda) \\ &\geq \left(a_N^2 - \frac{\lambda_1^2}{4}\right)\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda) - (a_N^2 - a_m^2)\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda) \\ &= \left(a_m^2 - \frac{\lambda_1^2}{4}\right)\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda) \\ &\geq \left(a^2 - \frac{\lambda_1^2}{4}\right)\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda), \end{aligned}$$

where in the second inequality we have used the assumption that $a_k^2 - a_{k-1}^2 \leq 0$. However, $\mathbb{P}_N^T(\Lambda)\mathbb{P}_N(\Lambda) > 0$ by our choice of N and $a^2 - \lambda_1^2/4 > 0$, this inequality contradicts Theorem 4.4. Therefore, $\lambda_1 \notin (-2a, 2a)$. Similarly, by using $S_{n,2}$, we can show $\lambda_2 \notin (-2b, 2b)$. The proof is completed. □

We note that if both $\{p_n\}$ and $\{q_n\}$ are taken to be Tchebycheff polynomials of the second kind, then the conditions on a_n and b_n are trivially satisfied as then $a_n = 1/2$ for all n.

In the univariate case, results of this nature have been proved under rather weak conditions on a_n (cf. [4]). The structure of eigenvalues of Jacobi matrix has been studied extensively and understood in large part. The multivariate case is far more complicated, partly due to the geometric complexity. For example, the classical orthogonal polynomials have several extensions to polynomials of two variables, depending on the geometric region of the support set of the measure (cf. [5]). The case discussed in Theorem 4.7 is the easier one that corresponds to the product region, which allows us to deal with one variable at a time. In general, variables are interrelated, while the relation forms the geometric structure of the measure. To fully understand the eigenvalue structure of T, or the spectrum of T, it seems that certain conditions have to be imposed on the coefficient matrices to reflect the geometric structures.

Acknowledgment

The author thanks a referee for the great care with which the details of the paper have been checked and corrected.

References

1. N. I. Akhiezer and I. M. Glazman, *Theory of linear operators in Hilbert space*, Ungar, New York, 1961.
2. T. S. Chihara, *An introduction to orthogonal polynomials*, Mathematics and its Applications, vol. 13, Gordon and Breach, New York, 1978.
3. J. Dombrowski, *Orthogonal polynomials and functional analysis*, Orthogonal Polynomials (P. Nevai, ed.), Kluwer Academic, Amsterdam, 1990, pp. 147–161.
4. ——, *Tridiagonal matrix representations of cyclic selfadjoint operators.* I, Pacific J. Math. **114** (1984), 325–334; II, **120** (1985), 47–53.

5. T. Koornwinder, *Two-variable analogues of the classical orthogonal polynomials*, Theory and Applications of Special Functions (R. A. Askey, ed.), Academic Press, 1975.

6. H. Möller, *Polynomideale und Kubaturformeln*, Thesis, Univ. of Dortmund, 1973.

7. ——, *Kubaturformeln mit minimaler knotenzahl*, Numer. Math. **25** (1976), 185–200.

8. C. R. Morrow and T. N. L. Patterson, *Construction of algebraic cubature rules using polynomial ideal theory*, SIAM J. Numer. Anal. **15** (1978), 953–976.

9. I. P. Mysovskikh, *Numerical characteristics of orthogonal polynomials in two variables*, Vestnik Leningrad Univ. Math. **3** (1976), 323–332.

10. ——, *The approximation of multiple integrals by using interpolatory cubature formulae*, Quantitative Approximation (R. A. DeVore and K. Scherer, eds.), Academic Press, New York, 1980, pp. 217–243.

11. F. Riesz and B. Sz.-Nagy, *Functional analysis*, Ungar, New York, 1955.

12. H. Schmid, *On cubature formulae with a minimal number of knots*, Numer. Math. **31** (1978), 282–297.

13. ——, *Interpolatorische Kubaturformeln*, Dissertationes Math. **220** (1983), 1–122.

14. ——, *Minimal cubature formulae and matrix equation*, preprint.

15. M. Stone, *Linear transformations in Hilbert space and their applications to analysis*, Amer. Math. Soc. Colloq. Publ., vol. 15, Amer. Math. Soc., Providence, R.I., 1932; reprinted1983.

16. A. Stroud, *Approximate calculation of multiple integrals*, Prentice-Hall, Englewood Cliffs, N.J., 1971.

17. G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, R.I., 4th ed., 1975.

18. Y. Xu, *On multivariate orthogonal polynonials*, SIAM J. Math. Anal. **24** (1993), 783–794.

19. ——, *Multivariate orthogonal polynomials and operator theory*, Trans. Amer. Math. Soc. (to appear).

20. ——, *Gaussian cubature and bivariate polynomial interpolation*, Math. Comp. **59** (1992), 547–555.

Department of Mathematics, University of Oregon, Eugene, Oregon 97403-1222
E-mail address: `yuan@bright.uoregon.edu`