EQUIVALENCE CLASSES OF BLOCK JACOBI MATRICES

ROSTYSLAV KOZHAN

(Communicated by Walter Van Assche)

ABSTRACT. The paper contains two results on the equivalence classes of block Jacobi matrices: first, that the Jacobi matrix of type 2 in the Nevai class has A_n coefficients converging to $\mathbf{1}$, and second, that under an L^1-type condition on the Jacobi coefficients, equivalent Jacobi matrices of types 1, 2 and 3 are pairwise asymptotic.

1. INTRODUCTION AND RESULTS

A block Jacobi matrix is an infinite matrix of the form

$$
J = \begin{pmatrix}
B_1 & A_1 & \mathbf{0} & \\
A_1^* & B_2 & A_2 & \ddots \\
\mathbf{0} & A_2^* & B_3 & \ddots \\
& \ddots & \ddots & \ddots
\end{pmatrix},
$$

where A_n, B_n are $l \times l$ matrices with A_n invertible. The sequences A_n and B_n are called Jacobi parameters of J.

Two block Jacobi matrices J and $\widetilde{J}$ are called equivalent if their Jacobi parameters satisfy

$$
(1.1) \qquad \widetilde{A}_n = \sigma_n^* A_n \sigma_{n+1}, \quad \widetilde{B}_n = \sigma_n^* B_n \sigma_n
$$

for unitary σ_n's with $\sigma_1 = \mathbf{1}$. The definition comes from the fact that (1.1) holds if and only if the (matrix-valued) spectral measures of J and $\widetilde{J}$ coincide (see [3] for the details).

Using the convention $p_{-1} = \mathbf{0}$, $A_0 = \mathbf{1}$, $p_0 = \mathbf{1}$, the recurrence

$$
x p_n(x) = p_{n+1}(x) A_{n+1}^* + p_n(x) B_{n+1} + p_{n-1}(x) A_n, \quad n = 0, 1, \ldots,
$$

allows one to define a sequence of matrix-valued polynomials which turn out to be (right-) orthonormal with respect to the above mentioned spectral measure.

Inductively it is easy to see that

$$
(1.2) \qquad \widetilde{p}_n(x) = p_n(x) \sigma_{n+1},
$$

where $\widetilde{p}_n$ are the orthonormal polynomials for $\widetilde{J}$.

We say that a block Jacobi matrix is of type 1 if $A_n > 0$ for all n, of type 2 if $A_1 A_2 \ldots A_n > 0$ for all n, and of type 3 if every A_n is lower triangular with strictly

positive elements on the diagonal. Each equivalence class of block Jacobi matrices contains exactly one matrix of types 1, 2 and 3 (this follows from the uniqueness of the polar and QR decompositions; see [3] for the proof).

We say that J is in the Nevai class if

$$B_n \to \mathbf{0}, \quad A_n A_n^* \to \mathbf{1}.$$

Note that this definition is invariant within the equivalence class of Jacobi matrices.

Theorem 1. *Assume J belongs to the Nevai class. If J is of type 1, 2 or 3, then $A_n \to \mathbf{1}$ as $n \to \infty$.*

This result was proven in [3] for the type 1 and 3 cases, and was left open for type 2. It is proven here in Section 2.

Note that the essence of Theorem 1 is to show that $\sigma_n^* \sigma_{n+1} \to \mathbf{1}$, where σ_n's are the unitary coefficients from (1.1) for J, $\widetilde{J}$ of type 1, 2 or 3. Looking at (1.2), it is clear that any result on the asymptotics of p_n (see e.g. [1], [4], [5]) would involve the limit $\lim_{n\to\infty} \sigma_n$. This explains the need for the following definition.

Definition. Two equivalent matrices J and $\widetilde{J}$ with (1.1) are called asymptotic to each other if the limit $\lim_{n\to\infty} \sigma_n$ exists.

Clearly this is an equivalence relation on the set of equivalent block Jacobi matrices. Thus, establishing Szegő asymptotics (which simply means $\lim_{n\to\infty} z^n p_n(z + z^{-1})$ exists) for any block Jacobi matrix immediately implies the corresponding asymptotics for any of the Jacobi matrices asymptotic to the original one.

Theorem 2. *Assume*

$$(1.3) \qquad\qquad \sum_{n=1}^{\infty} \left[\|\mathbf{1} - A_n A_n^*\| + \|B_n\| \right] < \infty.$$

Then the corresponding Jacobi matrices of types 1, 2 and 3 are pairwise asymptotic.

Remarks. 1. The condition (1.3) doesn't depend on the choice of the representative of the equivalence class of equivalent matrices.

2. The proof also shows that any Jacobi matrix, for which eventually each A_n has real eigenvalues, is also asymptotic to types 1, 2 and 3.

3. An example of an equivalence class of block Jacobi matrices that fails (1.3) and that has type 1 and type 2 nonasymptotic to each other can be found at the end of Section 2.

2. Proofs of the results

We will be using the following lemma from [6]. For self-containment purposes we give a proof of it in the Appendix.

Lemma 1 (Li [6])**.** *Let ϕ be the map that takes any invertible matrix T to the unitary factor U in its polar decomposition $T = |T|U$, where $|T| = \sqrt{TT^*}$. Then for any invertible $l \times l$ matrices B, D the following holds:*

$$\|\phi(B) - \phi(BD)\|_{HS} \le \sqrt{\|\mathbf{1} - D^{-1}\|_{HS}^2 + \|\mathbf{1} - D\|_{HS}^2},$$

where $\|\cdot\|_{HS}$ is the Hilbert–Schmidt norm.

Proof of Theorem 1. For types 1 and 3, the statement is proven in [3].

Assume J is of type 2. Denote by $\widehat{J}$ the type 1 Jacobi matrix equivalent to J. Denote its Jacobi parameters by $\widehat{A}_n, \widehat{B}_n$, and let

$$(2.1) \qquad A_n = \sigma_n^* \widehat{A}_n \sigma_{n+1}$$

for some unitaries σ_n. Since $\widehat{A}_n \to \mathbf{1}$, we get $A_n = \sigma_n^* \widehat{A}_n \sigma_{n+1} = \left(\sigma_n^* \widehat{A}_n \sigma_n \right) \sigma_n^* \sigma_{n+1}$ converges to $\mathbf{1}$ if and only if $\lim_{n \to \infty} \sigma_n^* \sigma_{n+1} = \mathbf{1}$.

Denote $Q_n = A_1 \ldots A_n$, which is a positive-definite matrix. Note that $\widehat{Q}_n = \widehat{A}_1 \ldots \widehat{A}_n = A_1 \ldots A_n \sigma_{n+1}^* = Q_n \sigma_{n+1}^*$, so $Q_n = |\widehat{Q}_n|$ and $\sigma_{n+1} = \phi(\widehat{Q}_n)^*$. Here ϕ is the same as in Lemma 1.

Now $\widehat{A}_{n+1} \to \mathbf{1}$, together with Lemma 1, implies that $\phi(\widehat{Q}_{n+1}) - \phi(\widehat{Q}_n) = \phi(\widehat{Q}_n \widehat{A}_{n+1}) - \phi(\widehat{Q}_n) \to \mathbf{0}$. Thus, $\sigma_{n+1} - \sigma_n \to \mathbf{0}$ and $\lim_{n \to \infty} \sigma_n^* \sigma_{n+1} = \mathbf{1}$. $\qquad \square$

For the type 3 case of Theorem 2, we will need the following lemma. Recall that the singular values of a matrix A are defined to be the eigenvalues of $|A|$.

Lemma 2. *There exists a constant c such that for all $l \times l$ matrices A*

$$(2.2) \qquad \sum_{j=1}^{l} (\sigma_j - |\lambda_j|) \leq c \sum_{j=1}^{l} (1 - \sigma_j)^2,$$

where $\{\lambda_j\}_{j=1}^l$ and $\{\sigma_j\}_{j=1}^l$ are the eigenvalues and singular values of A, ordered by $|\lambda_1| \geq \ldots \geq |\lambda_l|$, $\sigma_1 \geq \ldots \geq \sigma_l \geq 0$, where c depends on l only.

Proof. For sufficiently large matrices A the inequality is clear. It also holds for any compact set on which the right-hand side of (2.2) does not vanish. Therefore, we only need to worry about neighborhoods of matrices with $\sum_{j=1}^l (1 - \sigma_j)^2 = 0$, that is, unitary matrices.

Consider any matrix A within distance $1/2$ from the unitary group. Let $U = \phi(A)$ be the unitary factor in the polar decomposition of A. Since $\phi(A)$ is always the closest unitary to A (see e.g. [2]), we get

$$\|A - U\| \leq 1/2 \quad \text{and} \quad \||A| - \mathbf{1}\| \leq 1/2.$$

The second inequality immediately gives $|\sigma_j - 1| \leq 1/2$, which in turn implies $||\lambda_j| - 1| \leq 1/2$ by (2.3) below. The following basic facts are well-known (see e.g. [7]):

$$(2.3) \qquad \sigma_1 \geq |\lambda_j| \geq \sigma_l \text{ for any } j,$$

$$(2.4) \qquad \prod_{j=1}^{l} |\lambda_j| = \prod_{j=1}^{l} \sigma_j.$$

Let $\varepsilon_j = \sigma_j - 1$, $\delta_j = |\lambda_j| - 1$. Then from (2.4),

$$\delta_l = \frac{\prod_{j=1}^{l} \sigma_j}{\prod_{j=1}^{l-1} |\lambda_j|} - 1 = \frac{\prod_{j=1}^{l}(1 + \varepsilon_j) - \prod_{j=1}^{l-1}(1 + \delta_j)}{\prod_{j=1}^{l-1}(1 + \delta_j)},$$

and so

$$\sum_{j=1}^{l}(\sigma_j - |\lambda_j|) = \sum_{j=1}^{l}(\varepsilon_j - \delta_j)$$

(2.5)
$$= \frac{\prod_{j=1}^{l-1}(1+\delta_j)\sum_{j=1}^{l}\varepsilon_j - \prod_{j=1}^{l-1}(1+\delta_j)\sum_{j=1}^{l-1}\delta_j - \prod_{j=1}^{l}(1+\varepsilon_j) + \prod_{j=1}^{l-1}(1+\delta_j)}{\prod_{j=1}^{l-1}(1+\delta_j)}.$$

The first-order terms (i.e. those involving only one of ε's or δ's) of the numerator cancel out:

$$\sum_{j=1}^{l}\varepsilon_j - \sum_{j=1}^{l-1}\delta_j - \left(1 + \sum_{j=1}^{l}\varepsilon_j\right) + \left(1 + \sum_{j=1}^{l-1}\delta_j\right) = 0.$$

Now note that by (2.3), $|\delta_j| \leq |\varepsilon_1| + |\varepsilon_l|$. Using this and $|\varepsilon_j\varepsilon_k| \leq (\varepsilon_j^2 + \varepsilon_k^2)/2$ we can bound all of the second-order terms (i.e. those with $\varepsilon_j\varepsilon_k$, $\varepsilon_j\delta_k$ and $\delta_j\delta_k$) by $\widetilde{c}\sum_{j=1}^{l}\varepsilon_j^2$, where $\widetilde{c}$ will depend on l only. All of the higher-order terms can be taken care of by using $|\varepsilon_j| < 1, |\delta_j| < 1$ to reduce it to the second-order. Finally, the denominator on the right-hand side of (2.5) is bounded below by $1/2^l$. Therefore, we obtain

$$\sum_{j=1}^{l}(\sigma_j - |\lambda_j|) \leq c\sum_{j=1}^{l}\varepsilon_j^2 = c\sum_{j=1}^{l}(1-\sigma_j)^2,$$

which proves our lemma. $\square$

Lemma 3. *There exists a constant c so that*

(2.6) $$\|\mathbf{1} - A\| \leq c\|\mathbf{1} - |A|\|$$

for any $l \times l$ matrix A with real positive eigenvalues, where c depends only on l.

Proof. By the equivalence of norms, we can prove (2.6) for the Hilbert–Schmidt norm instead. Let $\lambda_1 \geq \ldots \geq \lambda_l > 0$ be the eigenvalues of A, and let $\sigma_1 \geq \ldots \geq \sigma_l > 0$ be the singular values of A. Note that

$$\|\mathbf{1} - A\|_{HS}^2 = \mathrm{Tr}\left[(\mathbf{1}-A)(\mathbf{1}-A)^*\right] = l - 2\sum_{j=1}^{l}\mathrm{Re}\,\lambda_j + \mathrm{Tr}AA^*$$

$$= l - 2\sum_{j=1}^{l}\lambda_j + \sum_{j=1}^{l}\sigma_j^2,$$

$$\|\mathbf{1} - |A|\|_{HS}^2 = \mathrm{Tr}\left[(\mathbf{1}-|A|)^2\right] = l - 2\sum_{j=1}^{l}\sigma_j + \sum_{j=1}^{l}\sigma_j^2,$$

and so $\|\mathbf{1} - A\|_{HS}^2 \leq M\|\mathbf{1} - |A|\|_{HS}^2$ holds if and only if

$$2\sum_{j=1}^{l}(\sigma_j - \lambda_j) \leq (M-1)\sum_{j=1}^{l}(1-\sigma_j)^2.$$

Since $\lambda_j = |\lambda_j|$, the previous lemma proves the result. $\square$

Proof of Theorem 2. As in Theorem 1, let $\widehat{A}_n$ be of type 1, and let A_n be of type 2 with the equivalence (2.1). Then keeping the notation of Theorem 1 and using Lemma 1, we have

$$\sum_{n=1}^{\infty} \|\sigma_n - \sigma_{n+1}\|_{HS} = \sum_{n=1}^{\infty} \|\phi(\widehat{Q}_{n-1}) - \phi(\widehat{Q}_n)\|_{HS}$$

$$\leq \sum_{n=1}^{\infty} \sqrt{\|\mathbf{1} - \widehat{A}_n^{-1}\|_{HS}^2 + \|\mathbf{1} - \widehat{A}_n\|_{HS}^2}$$

$$\leq \sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n^{-1}\|_{HS} + \sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n\|_{HS}$$

$$\leq (\sup_n \|\widehat{A}_n\|_{HS} + 1) \sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n\|_{HS}$$

$$\leq (\sup_n \|\widehat{A}_n\|_{HS} + 1) \sup_n \|(\mathbf{1} + \widehat{A}_n)^{-1}\|_{HS} \sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n^2\|_{HS} < \infty,$$

since $\widehat{A}_n \to \mathbf{1}$, and so $\sup_n \|\widehat{A}_n\|_{HS} < \infty$ and $\sup_n \|(\mathbf{1} + \widehat{A}_n)^{-1}\|_{HS} < \infty$. This implies that σ_n is Cauchy and so converges.

An alternative indirect way of proving that type 1 and type 2 are asymptotic to each other is as follows: it is proven in [4] that under condition (1.3) Szegő asymptotics for the type 2 block Jacobi matrix holds. In [5] the same fact is obtained for the type 1 Jacobi matrix. Therefore (1.2) implies that the limit $\lim_{n \to \infty} \sigma_n$ exists.

Now assume that $\widehat{A}_n$ is of type 1, and A_n of type 3 with the equivalence (2.1). Since all eigenvalues of A_n are real and positive, Lemma 3 applies, and we get

$$\sum_{n=1}^{\infty} \|\mathbf{1} - A_n\| \leq c \sum_{n=1}^{\infty} \|\mathbf{1} - |A_n|\| = c \sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n\|$$

since $|A_n| = \sigma_n^* \widehat{A}_n \sigma_n$ by (2.1). Now

$$\sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n\| \leq \sup_n \|(\mathbf{1} + \widehat{A}_n)^{-1}\| \sum_{n=1}^{\infty} \|\mathbf{1} - \widehat{A}_n^2\| < \infty,$$

which implies

$$\sum_{n=1}^{\infty} \|\sigma_n - \sigma_{n+1}\| = \sum_{n=1}^{\infty} \|\mathbf{1} - \sigma_n^* \sigma_{n+1}\| \leq \sum_{n=1}^{\infty} \|\mathbf{1} - A_n\| + \sum_{n=1}^{\infty} \|A_n - \sigma_n^* \sigma_{n+1}\|$$

$$= \sum_{n=1}^{\infty} \|\mathbf{1} - A_n\| + \sum_{n=1}^{\infty} \|\widehat{A}_n - \mathbf{1}\| < \infty.$$

This shows that σ_n is Cauchy and so converges. $\qquad \square$

Example 1. Let $D_k = \left(\begin{smallmatrix} (k+1)/k & 0 \\ 0 & 1 \end{smallmatrix} \right)$ for $k \geq 1$. Note that $D_k \to \mathbf{1}$.

Pick some unitary τ, and define the sequence $\widehat{A}_n$ as follows: $\widehat{A}_1 = \tau^* D_1 \tau$, $\widehat{A}_2 = D_1$, $\widehat{A}_3 = D_1^{-1}$, $\widehat{A}_4 = D_2$, $\widehat{A}_5 = D_3$, $\widehat{A}_6 = D_3^{-1}$, $\widehat{A}_7 = D_2^{-1}$, $\widehat{A}_8 = D_4$, and so on: we define $\widehat{A}_k$'s for $2^j \leq k < 2^{j+1}$ in terms of further and further chunks of the

sequence D_k as

$$\widehat{A}_{2^j} = D_{2^{j-1}}, \ldots, \widehat{A}_{3 \cdot 2^{j-1}-1} = D_{2^j-1},$$
$$\widehat{A}_{3 \cdot 2^{j-1}} = D_{2^j-1}^{-1}, \ldots, \widehat{A}_{2^{j+1}-1} = D_{2^{j-1}}^{-1}.$$

Note that $\widehat{A}_n > 0$; i.e. the sequence corresponds to a block Jacobi matrix of type 1. Using the notation from Section 2, let $\widehat{Q}_n = \widehat{A}_1 \ldots \widehat{A}_n$. Then

$$\widehat{Q}_{2^j-1} = \widehat{A}_1, \quad \widehat{Q}_{3 \cdot 2^{j-1}-1} = \widehat{A}_1 D_{2^{j-1}} \ldots D_{2^j-1} = \widehat{A}_1 D_1,$$

and $\sigma_{2^j} = \phi(\widehat{Q}_{2^j-1})^* = \mathbf{1}$, $\sigma_{3 \cdot 2^{j-1}} = \phi(\widehat{Q}_{3 \cdot 2^{j-1}-1})^* = \phi(\tau^* D_1 \tau D_1)^*$. Now choose τ such that $\phi(\tau^* D_1 \tau D_1)$ is not positive definite. This gives us the fact that $\lim_{n \to \infty} \sigma_n$ does not exist; i.e. type 1 and type 2 are not asymptotic to each other.

Of course, the reason is that (1.3) fails here: $\sum \|\mathbf{1} - A_n A_n^*\|$ diverges as $\sum \frac{1}{n}$.

Appendix A. Proof of Li's lemma

Proof of Lemma 1. Let $B = U\Sigma V^*$ and $BD = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^*$ be the singular value decompositions of B and BD (i.e. $U, \widetilde{U}, V, \widetilde{V}$ are unitary, and $\Sigma, \widetilde{\Sigma}$ are positive and diagonal). Denote

$$Y = \widetilde{U}^*(B - BD)V = \widetilde{U}^* U \Sigma - \widetilde{\Sigma}\widetilde{V}^* V,$$
$$Z = U^*(B - BD)\widetilde{V} = \Sigma V^* \widetilde{V} - U^* \widetilde{U}\widetilde{\Sigma}.$$

Then

(A.1) $\qquad Y - Z^* = (\widetilde{U}^* U - \widetilde{V}^* V)\Sigma + \widetilde{\Sigma}(\widetilde{U}^* U - \widetilde{V}^* V) = X\Sigma + \widetilde{\Sigma}X,$

where $X = \widetilde{U}^* U - \widetilde{V}^* V$. On the other hand,

(A.2) $\quad \begin{aligned} Y - Z^* &= \widetilde{U}^*(B - BD)V - \widetilde{V}^*(B^* - D^* B^*)U \\ &= \widetilde{\Sigma}\widetilde{V}^*(D^{-1} - \mathbf{1})V - \widetilde{V}^*(\mathbf{1} - D^*)V\Sigma = \widetilde{\Sigma}E - F\Sigma, \end{aligned}$

where $E = \widetilde{V}^*(D^{-1} - \mathbf{1})V$, $F = \widetilde{V}^*(\mathbf{1} - D^*)V$. Note that Σ and $\widetilde{\Sigma}$ are diagonal, and therefore the solution of (A.1)=(A.2) is

$$x_{ij} = \frac{\widetilde{\sigma}_{ii} e_{ij} - f_{ij} \sigma_{jj}}{\sigma_{jj} + \widetilde{\sigma}_{ii}},$$

where $X \equiv (x_{ij})$, $E \equiv (e_{ij})$, $F \equiv (f_{ij})$, $\Sigma \equiv (\sigma_{ij})$, $\widetilde{\Sigma} \equiv (\widetilde{\sigma}_{ij})$. Note that $\sigma_{jj} > 0$ and $\widetilde{\sigma}_{ii} > 0$, and thus by the Schwarz inequality,

$$|x_{ij}|^2 \le \frac{\sigma_{jj}^2 + \widetilde{\sigma}_{ii}^2}{(\sigma_{jj} + \widetilde{\sigma}_{ii})^2}(|e_{ij}|^2 + |f_{ij}|^2) \le |e_{ij}|^2 + |f_{ij}|^2,$$

which implies

$$\|X\|_{HS}^2 \le \|E\|_{HS}^2 + \|F\|_{HS}^2 = \|\mathbf{1} - D^{-1}\|_{HS}^2 + \|\mathbf{1} - D\|_{HS}^2.$$

Finally, note that $\phi(B) = UV^*$ and $\phi(BD) = \widetilde{U}\widetilde{V}^*$, so $\|\phi(B) - \phi(BD)\|_{HS} = \|\widetilde{U}XV^*\|_{HS} = \|X\|_{HS}$, and we are done. $\qquad \square$

References

1. A. Aptekarev, E. Nikishin, The scattering problem for a discrete Sturm–Liouville operator, Mat. Sb. 121 (163) (1983), 327–358. MR708000 (85d:39004)
2. R. Bhatia, Matrix Analysis, Springer-Verlag, New York, 1997. MR1477662 (98i:15003)
3. D. Damanik, A. Pushnitski, B. Simon, The analytic theory of matrix orthogonal polynomials, Surveys in Approximation Theory 4 (2008), 1–85. MR2379691 (2010d:42038)
4. R. Kozhan, Szegő asymptotics for matrix-valued measures with countably many bound states, J. Approx. Theory 162 (2010), no. 6, 1211–1224.
5. R. Kozhan, Jost function for matrix orthogonal polynomials, preprint.
6. R.-C. Li, Relative perturbation bounds for the unitary polar factor, BIT 37 (1997), no. 1, 67–75. MR1431359 (97k:15026)
7. H. Weyl, Inequalities between the two kinds of eigenvalues of a linear transformation, Proc. Natl. Acad. Sci. USA 35 (7) (1949), 408–411. MR0030693 (11:37d)

Department of Mathematics 253-37, California Institute of Technology, Pasadena, California 91125

E-mail address: `rostysla@caltech.edu`