Jacobi matrices

Alex Bloemendal

April 30, 2012

We discuss the relationships among Jacobi matrices, orthogonal polynomials, spectral measure, moments, minors, Gaussian quadrature, resolvents and continued fractions in the simplest setting, namely the finite-dimensional one. The formal structure is essentially the same as that in the infinite-dimensional setting, where it leads into the rich analytic world of orthogonal polynomials on the real line. When the underlying measure is finitely supported, however, the analysis becomes trivial and the algebraic relationships are readily apparent.

Jacobi canonical form

We work in a real (complex) inner product space of dimension n, in particular $\mathbb{R}^n$ ($\mathbb{C}^n$) with standard basis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and standard inner product $\langle \cdot, \cdot \rangle$.

Equivalence classes of symmetric (Hermitian) matrices $A \sim QAQ^*$ up to orthogonal (unitary) conjugation are canonically represented by diagonal matrices with ordered eigenvalues. Equivalence classes of *pairs* $(A, \mathbf{v}) \sim (QAQ^*, Q\mathbf{v})$, where $\mathbf{v}$ is a normalized cyclic vector for A (i.e. $\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$ are linearly independent), are canonically represented by $(J, \mathbf{e}_1)$ where J is a *Jacobi matrix*:

$$\begin{aligned} J_{ij} &> 0 \quad \text{for} \quad |i-j| = 1 \\ J_{ij} &= 0 \quad \text{for} \quad |i-j| \geq 2, \end{aligned}$$

or in other words

$$J = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & \ddots & \ddots & \\ & & \ddots & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \end{pmatrix} \tag{1}$$

with $b_i > 0$.

Theorem 1. *Let A be an $n \times n$ Hermitian matrix and let $\mathbf{v}$ be an A-cyclic unit vector. Then there exists a unique orthonormal basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ starting with $\mathbf{v}_1 = \mathbf{v}$ with respect to which A is in Jacobi form, i.e. such that $J_{ij} = \langle \mathbf{v}_i, A\mathbf{v}_j \rangle$ is a Jacobi matrix.*

Proof. To show existence, apply the Gram-Schmidt process to the cyclic basis $(\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v})$.

Exercise 1. Verify that this works. Hint: Rather than using formulas, use that the output of Gram-Schmidt,

$$(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \mathrm{GS}(\mathbf{u}_1, \ldots, \mathbf{u}_n),$$

is characterized by

(i) $\mathbf{v}_k \in \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$

(ii) $\mathbf{v}_k \perp \{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}\}$

(iii) $\langle \mathbf{v}_k, \mathbf{u}_k \rangle > 0$

(iv) $\|\mathbf{v}_k\| = 1$.

To show uniqueness we first make the following observation.

Exercise 2. If J is a Jacobi matrix, then $\mathbf{e}_1$ is cyclic and

$$\operatorname{GS}(\mathbf{e}_1, J\mathbf{e}_1, \ldots, J^{n-1}\mathbf{e}_1) \;=\; (\mathbf{e}_1, \ldots, \mathbf{e}_n)$$

reproduces the standard basis.

Now suppose J, $J' = QJQ^*$ are both Jacobi and $Q\mathbf{e}_1 = \mathbf{e}_1$, i.e. $(J, \mathbf{e}_1) \sim (J', \mathbf{e}_1)$. Then, since GS is coordinate-independent,

$$\begin{aligned}
(\mathbf{e}_1, \ldots, \mathbf{e}_n) \;&=\; \operatorname{GS}(\mathbf{e}_1, J'\mathbf{e}_1, \ldots, J'^{n-1}\mathbf{e}_1) \\
&=\; \operatorname{GS}(Q\mathbf{e}_1, QJ\mathbf{e}_1, \ldots, QJ^{n-1}\mathbf{e}_1) \\
&=\; Q\,\operatorname{GS}(\mathbf{e}_1, J\mathbf{e}_1, \ldots, J^{n-1}\mathbf{e}_1) \\
&=\; (Q\mathbf{e}_1, \ldots, Q\mathbf{e}_n).
\end{aligned}$$

□

Orthogonal polynomials

Let $\sigma \in \mathcal{P}(\mathbb{R})$ be a probability measure with finite support, $\#\operatorname{supp}\sigma = n$; write

$$\sigma \;=\; \sum_{i=1}^{n} q_i^2 \delta_{\lambda_i} \tag{2}$$

with $\lambda_1 < \cdots < \lambda_n$ and $q_i > 0$, $\sum_i q_i^2 = 1$. Then $L^2(\sigma)$ is an n-dimensional inner product space.

Consider the function $x \mapsto x$, which we write just as x; consider now the linear operator of multiplication by x, which we denote also simply by x. Note that x is self-adjoint with eigenvalues λ_i and corresponding orthonormal eigenvectors $\frac{1}{q_i}\mathbf{1}_{\lambda_i}$.

Exercise 3. The constant function $\mathbf{1}$ is a cyclic vector for x. Hint: What does it mean to vanish in $L^2(\sigma)$?

Let J be the associated Jacobi matrix, i.e. $(x, \mathbf{1}) \sim (J, \mathbf{e}_1)$. Then the Jacobi basis is

$$\operatorname{GS}(\mathbf{1}, x, \ldots, x^{n-1}) \;=\; (p_0, p_1, \ldots, p_{n-1}),$$

comprising the normalized orthogonal polynomials with respect to σ.

In coordinates, (1) reads

$$J\mathbf{e}_k \;=\; b_{k-1}\mathbf{e}_{k-1} + a_k\mathbf{e}_k + b_k\mathbf{e}_{k+1};$$

in the original basis, this relation becomes the three-term recurrence

$$xp_{k-1} \;=\; b_{k-1}p_{k-2} + a_k p_{k-1} + b_k p_k. \tag{3}$$

Here we have taken $b_0 = b_n = 0$. For $k = 1$ the recurrence is initialized with $p_0 = \mathbf{1}$, and in this way $p_1, \ldots, p_{n-1}$ are determined by the coefficients, i.e. by J. For $k = n$ we find that the degree n polynomial

$$(x - a_n)p_{n-1}(x) - b_{n-1}p_{n-2}(x)$$

vanishes in $L^2(\sigma)$, i.e. its zeros coincide with the eigenvalues λ_i; this polynomial is therefore proportional to the characteristic polynomial $\det(x - J)$.

The same reasoning can be applied to the leading principal $k \times k$ minor J_k. Now the recurrence (3) truncates with b_k replaced by 0; we obtain the same $p_1, \ldots, p_{k-1}$, and find that $p_k(x)$ is proportional to $\det(x - J_k)$.

Writing

$$\hat{x} \;=\; \begin{pmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{pmatrix},$$

one can express (3) as

$$J\hat{x} \;=\; x\hat{x}.$$

This identity holds in $L^2(\sigma)$, i.e. at the eigenvalues $x = \lambda_i$. We deduce that the corresponding eigenvectors of J are $\hat{\lambda}_i$. One can also obtain this fact directly using orthonormality:

$$\left\langle p_k, \tfrac{1}{q_i}\mathbf{1}_{\lambda_i} \right\rangle \;=\; \int p_k(x)\tfrac{1}{q_i}\mathbf{1}_{\lambda_i}(x)\,\sigma(dx) \;=\; q_i\,p_k(\lambda_i),$$

so the normalized eigenvector is $q_i\hat{\lambda}_i$. In particular, the $\mathbf{e}_1$-component is just q_i.

Spectral measure

It is natural to ask whether the above procedure can be reversed: Given a Jacobi matrix, does (3) generate the orthogonal polynomials with respect to some measure of the form (2)?

Slightly more generally, given a Hermitian matrix A, one can form its *spectral measure at a vector* $\mathbf{v}$. That is, if A has spectral decomposition $A = \sum_{i=1}^n \lambda_i \mathbf{u}_i\mathbf{u}_i^*$, let $q_i = |\langle \mathbf{u}_i, \mathbf{v}\rangle|$ and define $\sigma = \sigma_{A,\mathbf{v}}$ by (2).

Exercise 4. This measure has the characterizing property that, for any polynomial p,

$$\int p(x)\,\sigma_{A,\mathbf{v}}(dx) \;=\; \langle \mathbf{v}, p(A)\mathbf{v}\rangle\,. \tag{4}$$

An advantage of (4) is that it makes no reference to the spectral decomposition.

Exercise 5. Show that $\mathbf{v}$ is A-cyclic if and only if $\#\operatorname{supp}\sigma_{A,\mathbf{v}} = n$. Hint: Deduce this assertion from the identity $\int p(x)^2\,\sigma_{A,\mathbf{v}}(dx) = \|p(A)\mathbf{v}\|^2$, derived from (4).

In the latter situation A has distinct eigenvalues and can be viewed as the multiplication x on $L^2(\sigma_{A,\mathbf{v}})$; by (4) the pair $(A, \mathbf{v})$ is then unitarily equivalent to $(x, \mathbf{1})$. In particular, one can begin with a Jacobi matrix J and form $\sigma = \sigma_{J,\mathbf{e}_1}$, its spectral measure at $\mathbf{e}_1$. Then J may be viewed as representing the multiplication x on $L^2(\sigma)$ in the basis of orthogonal polynomials. In summary:

Theorem 2. *Jacobi matrices J as in* (1) *are in natural bijective correspondence with finitely supported measures σ as in* (2)*. Given σ, one obtains J as representing the multiplication x on $L^2(\sigma)$ in the basis of orthogonal polynomials. Given J, one obtains σ as the spectral measure of J at $\mathbf{e}_1$.*

Moments

Given σ as in (2), let

$$m_k \;=\; \int x^k \,\sigma(dx) \tag{5}$$

be its kth moment. Notice that knowledge of $m_1, \ldots, m_{2n-2}$ is sufficient to reconstruct the orthogonal polynomials $(p_0, p_1, \ldots, p_{n-1}) = \mathrm{GS}(1, x, \ldots, x^{n-1})$. If we also know m_{2n-1}, we can recover J by

$$J_{ij} \;=\; \langle \mathbf{e}_i, J\mathbf{e}_j \rangle \;=\; \int p_{i-1}(x)\, x\, p_{j-1}(x)\, \sigma(dx).$$

One can then recover $\sigma = \sigma_{J,\mathbf{e}_1}$. In particular, all moments may be obtained directly from J using (4):

$$m_k \;=\; \left(J^k\right)_{11}. \tag{6}$$

Exercise 6. Show directly that the first k elements of the list of moments $(m_1, m_2, m_3, \ldots)$ and the first k elements of the list of Jacobi coefficients $(a_1, b_1, a_2, b_2, \ldots)$ uniquely determine one another. Hint: Proceed by induction, expanding the right-hand side of (6).

Either way we conclude the following.

Theorem 3. *A probability measure supported on n real points is determined by its first $2n-1$ moments. (The normalization requirement may be removed by including the zeroth moment.)*

The previous exercise shows a little more: If J_k is the leading $k \times k$ principal minor of J and $\sigma_k = \sigma_{J_k,\mathbf{e}_1}$, then the first $2k-1$ moments of σ_k match those of σ. Thus σ_k is the optimal approximation of σ among measures supported on k points, in the sense that integrals of polynomials of degree up to $2k-1$ agree exactly. We have stumbled upon Gaussian quadrature, sometimes used as a rule in numerical analysis. By an earlier observation we find that the support of σ_k is precisely the zero set of p_k, a result known as the "fundamental theorem of Gaussian quadrature".

A lingering question is when some given numbers are actually the moments of a measure.

Theorem 4. *A sequence of real numbers $(m_0, \ldots, m_{2n-1})$ is given by (5) for some σ as in (2) if and only if the Hänkel matrix*

$$M \;=\; \left(m_{i+j}\right)_{i,j=0}^{n-1} \;=\; \begin{pmatrix} m_0 & m_1 & \cdots & m_{n-1} \\ m_1 & m_2 & \cdots & m_n \\ \vdots & \vdots & \ddots & \vdots \\ m_{n-1} & m_n & \cdots & m_{2n-2} \end{pmatrix}$$

is positive definite.

Proof. To see why the condition is necessary, observe that $m_{i+j} = \left\langle x^i, x^j \right\rangle_{L^2(\sigma)}$. That is, M is the Gram matrix of the monomials $1, x, \ldots, x^{n-1}$. Thus $M \geq 0$ automatically; in fact $M > 0$ since the monomials are linearly independent.

Towards sufficiency, consider an n-dimensional vector space with basis $\mathbf{x}_0, \ldots, \mathbf{x}_{n-1}$. Since $M > 0$ we can define an inner product through $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = m_{i+j}$, extended bilinearly. Let T be the shift operator, defined by $T\mathbf{x}_k = \mathbf{x}_{k+1}$ for $k = 0, \ldots, n-2$.

Exercise 7. There is a unique way to define $T\mathbf{x}_{n-1}$ so that $\langle \mathbf{x}_i, T\mathbf{x}_{n-1} \rangle = m_{i+n}$ for $0 \leq i \leq n-1$.

Now let $\sigma = \sigma_{T,\mathbf{x}_0}$. Then for $0 \leq i, j \leq n-1$,

$$\int x^{i+j+1}\, \sigma(dx) \;=\; \left\langle \mathbf{x}_0, T^{i+j+1}\mathbf{x}_0 \right\rangle \;=\; \langle \mathbf{x}_i, T\mathbf{x}_j \rangle \;=\; m_{i+j+1}. \qquad \square$$

Resolvents and continued fractions

There is a certain rational function that unifies the above representations.

Once again, take A Hermitian with spectral decomposition $\sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^*$. Define the resolvent $R(z) = (A-z)^{-1}$ for $z \in \mathbb{C} \setminus \{\lambda_1, \ldots, \lambda_n\}$. Fix a vector $\mathbf{v}$ and let $q_i = |\langle \mathbf{u}_i, \mathbf{v} \rangle|$.

Exercise 8. The Stieltjes transform of the spectral measure $\sigma = \sigma_{A,\mathbf{v}}$ is given by

$$S(z) \;=\; \int \frac{\sigma(dx)}{x-z} \;=\; \sum_{i=1}^{n} \frac{q_i^2}{\lambda_i - z} \;=\; \mathbf{v}^* R(z) \mathbf{v}. \tag{7}$$

With $\mathbf{v} = \mathbf{e}_1$ we find that $S(z) = R_{11}(z)$, the $(1,1)$-entry of the resolvent. Let $A^{(1)}$ be the $(1,1)$-principal minor of A and let $R^{(1)}(z) = (A^{(1)} - z)^{-1}$ be its resolvent; write

$$A \;=\; \begin{pmatrix} a & \mathbf{b}^* \\ \mathbf{b} & A^{(1)} \end{pmatrix}.$$

Exercise 9. Prove the identity

$$R_{11}(z) \;=\; \frac{1}{a - z - \mathbf{b}^* R^{(1)}(z) \mathbf{b}}. \tag{8}$$

Hint: It is a special case of the easily verified Schur complement formula, which gives the upper-left block of $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)^{-1}$ as $\left(A - BD^{-1}C\right)^{-1}$.

Exercise 10. The vector $\mathbf{e}_1$ is A-cyclic if and only if the rational function S has n distinct poles $\lambda_1 < \cdots < \lambda_n$. In this case it has $n-1$ distinct zeros $\mu_1 < \cdots < \mu_{n-1}$ interlaced strictly with its poles, i.e. $\lambda_1 < \mu_1 < \lambda_2 < \cdots < \mu_{n-1} < \lambda_n$.

Assuming this situation, the zeros of S have a natural interpretation as well. From (8) they coincide with the poles of $\mathbf{b}^* R^{(1)} \mathbf{b}$; but $\mathbf{b}^* R^{(1)} \mathbf{b}$ is the Stieltjes transform of $\sigma_{A^{(1)},\mathbf{b}}$, so its poles are the eigenvalues of $A^{(1)}$. Furthermore, given $\lambda_1, \ldots, \lambda_n$ together with $\mu_1, \ldots, \mu_{n-1}$ one can reconstruct

$$S(z) \;=\; \frac{(\mu_1 - z) \cdots (\mu_{n-1} - z)}{(\lambda_1 - z) \cdots (\lambda_n - z)}.$$

(The multiplicative constant is correct since both sides are $\sim -1/z$ at ∞.)

If A is in Jacobi form (1) we can go further. In this case $\mathbf{b}^* R^{(1)} \mathbf{b} = b_1^2 R_{11}^{(1)}$, and we can iterate (8) to obtain the continued fraction expansion

$$S(z) = \cfrac{1}{a_1 - z - \cfrac{b_1^2}{a_2 - z - \cfrac{b_2^2}{\ddots - \cfrac{b_{n-1}^2}{a_n - z}}}}.$$

In particular, A can be recovered directly from S. Explicitly one computes

$$\begin{aligned} a_1 \;&=\; 1/S(z) + z \,\big|_{z=\infty} \\ b_1^2 \;&=\; z\big(1/S(z) + z - a\big)\big|_{z=\infty} \\ S^{(1)}(z) \;&=\; -\big(1/S(z) + z - a\big)/b_1^2 \end{aligned}$$

and iterates with $S^{(1)}$ in place of S. This procedure gives a way to obtain the Jacobi form of any pair $(A, \mathbf{v})$ from the rational function S.

Finally, the moments $m_k = \sum_{i=1}^n q_i^2 \lambda_i^k$ are related to the Taylor expansion of S at infinity. Provided that $|z| > \max_{1 \le i \le n} |\lambda_i|$ we can expand (7) around $z = \infty$ to obtain

$$S(z) \;=\; -\frac{1}{z} \sum_{k=0}^{\infty} \frac{m_k}{z^k}.$$

This representation incidentally gives yet another way to see that the first $2n-1$ moments determine S. For suppose S and S' as above have $m_k = m'_k$ for $k = 1, \ldots, 2n-1$; then their difference $S - S'$ has a zero of order at least $2n+1$ at ∞. But $S - S'$ can have at most $2n$ poles, contradicting the fact that a rational function has the same number of zeros as poles—unless of course it vanishes identically.

In summary, the following data are equivalent ways to represent the pair $(A, \mathbf{v})$:

- the spectral measure of A at $\mathbf{v}$;
- the eigenvalues of A together with those of its principal minor with respect to $\mathbf{v}$;
- the Jacobi form of A with respect to $\mathbf{v}$;
- the first $2n-1$ moments of the spectral measure of A at $\mathbf{v}$.

There are four corresponding representations of the rational function $S(z)$:

$$\frac{q_1^2}{\lambda_1 - z} + \cdots + \frac{q_n^2}{\lambda_n - z} \;=\; \frac{(\mu_1 - z)\cdots(\mu_{n-1} - z)}{(\lambda_1 - z)\cdots(\lambda_n - z)} \;=\; \cfrac{1}{a_1 - z - \cfrac{b_1^2}{\ddots - \cfrac{b_{n-1}^2}{a_n - z}}} \;=\; -\frac{1}{z}\sum_{k=0}^{\infty} \frac{m_k}{z^k},$$

i.e. as partial fraction decomposition, polynomial quotient, continued fraction expansion and Taylor series. These relations induce bijections (in fact diffeomorphisms) between the domains

$$\begin{aligned}
\big\{(\lambda_1, \ldots, \lambda_n, q_1, \ldots, q_n) \;&:\; \lambda_1 < \cdots < \lambda_n,\, q_i > 0,\, q_1^2 + \cdots + q_n^2 = 1\big\},\\
\big\{(\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_{n-1}) \;&:\; \lambda_1 < \mu_1 < \lambda_2 < \cdots < \mu_{n-1} < \lambda_n\big\},\\
\big\{(a_1, \ldots, a_n, b_1, \ldots, b_{n-1}) \;&:\; b_i > 0\big\},\\
\big\{(m_1, \ldots, m_{2n-1}) \;&:\; (m_{i+j})_{i,j=0}^{n-1} > 0\big\}.
\end{aligned}$$