No.
Date.

Cues

Lecture 2.

Orthogonal polynomials: Spectral theorem, and Christoffel-Darboux kernel

- This proposition means: orthogonality $\Rightarrow$ 3-term recurrence!

E.x.1: Verify that $P_n(x) = P_{n+1}(x) + b_n P_n(x) + a_n^2 P_{n-1}(x)$ by using orthogonality.

E.x.2: Show that $b_n = \tilde{D}_{n+1}/D_{n+1}$ where
$$\tilde{D}_{n+1} = \det\begin{pmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & \vdots & & \vdots \\ m_{n-1} & m_n & \cdots & m_{2n-1} \\ m_{n+1} & m_{n+2} & \cdots & m_{2n+1} \end{pmatrix}.$$

Notes

1. Three-term recurrence and Jacobi operator

In this part, we show that orthogonality implies a 3-term recurrence relation.

__Prop.__ 2.1 (3-term recurrence): The orthogonal polynomials $\{P_n(x)\}_{n\geq 0}$ satisfy a 3-term recurrence relation $xP_n(x) = a_{n+1}P_n(x) + b_n P_n(x) + a_n P_{n-1}(x)$ where $a_0 = 0$ and
$$a_n = \int_{\mathbb{R}} x P_n(x) P_{n-1}(x)\, d\mu(x), \quad b_n = \int_{\mathbb{R}} x P_n^2(x)\, d\mu(x).$$

__Proof__: Since $\{P_n(x)\}_{n\geq 0}$ is a basis of $L^2(\mu)$, we have $xP_n(x) = \sum_{k=0}^{n} C_{n,k} P_k(x)$, where, by orthogonality, $C_{n,k} = \int_{\mathbb{R}} x P_n(x) P_k(x)\, d\mu(x)$. Moreover, from $\deg(xP_k) = k+1$, we have $C_{n,k} = 0$ if $k+1 < n$, so that only three terms remain in the recurrence. Obviously,
$$C_{n,n-1} = \int_{\mathbb{R}} x P_n(x) P_{n-1}(x)\, d\mu(x), \quad C_{n,n} = \int_{\mathbb{R}} x P_n^2(x)\, d\mu(x),$$
$$C_{n,n+1} = \int_{\mathbb{R}} x P_n(x) P_{n+1}(x)\, d\mu(x),$$

Summary

which depends on a single index n.

__Remark 2.2__: We have an explicit determinant formula for a_n. By comparing the coefficient of the highest order, we have $\gamma_n = a_{n+1}\gamma_{n+1}$, and thus $a_{n+1} = \gamma_n \gamma_{n+1}^{-1} = (D_{n+2} D_n / D_{n+1}^2)^{1/2}$, where $D_n = \det(m_{i+j})_{i,j=0}^{n-1}$ is the Hankel determinant.

No.

Date.

Cues

- Another expression for 3-term recurrence:

$$\begin{pmatrix} P_{n+1}(x) \\ P_n(x) \end{pmatrix} = L_n \begin{pmatrix} P_n(x) \\ P_{n-1}(x) \end{pmatrix},$$

where

$$L_n = \begin{pmatrix} a_{n+1}^{-1}(x-b_n) & a_{n+1}^{-1}a_n \\ 1 & 0 \end{pmatrix}.$$

This formula is useful in the construction of Lax pair in classic integrable systems.

Notes

Alternative expression for 3-term recurrence:

$$x\begin{pmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ \vdots \end{pmatrix} = \begin{pmatrix} b_0 & a_1 & 0 & \cdots \\ a_1 & b_1 & a_2 & \cdots \\ 0 & a_2 & b_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ \vdots \end{pmatrix} := J\begin{pmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ \vdots \end{pmatrix}$$

where J is called as an infinite tridiagonal Jacobi matrix.

Denote the upper left corner $(n\times n)$ of J by J_n, then we have

$$J_n\begin{pmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{pmatrix} = x\begin{pmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_n p_n(x) \end{pmatrix}.$$

Recall that $p_n(x)$ has n simple, real roots. Therefore, every zero $x_{n,j}$ of $p_n(x)$ is an eigenvalue of J_n, with corresponding eigenvector $(p_0(x_{n,j}), \cdots, p_{n-1}(x_{n,j}))^T$.

<u>Prop</u>. 2.2 (Relation between OPs and truncated Jacobi matrix): We have $P_n(x) = \det(xI_n - J_n)$.

<u>proof</u>: Firstly, $\det(xI_n - J_n)$ is a polynomial of degree n, with zeros $x_{n,j}$, $j=1,\cdots,n$. Since $P_n(x) = \gamma_n^{-1} p_n(x)$, it is known that $P_n(x)$ has the same zeros with $p_n(x)$, i.e. $x_{n,j}$, $j=1,\cdots,n$. Moreover, $P_n(x)$ and $\det(xI_n - J_n)$ are monic polynomials. Therefore, they are identical. #

Summary

In the first part, we have shown that orthogonality means 3-term recurrence relation.

An important question is:

Whether could we get orthogonality from 3-term recurrence relation?

No.
Date.

Cues	Notes

2. Spectral theorem

Favard's theorem tells: 3-term recurrence $\Longrightarrow$ orthogonality!

Thm. 2.3 (Favard's theorem):
Suppose that $\{P_n\}_{n=0,1,2,\cdots}$ is a sequence of polynomials that satisfy a 3-term recurrence relation

$$xP_n = a_{n+1}P_{n+1} + b_nP_n + a_nP_{n-1}, \quad n\geq 0$$

with $a_{n+1}>0$ and $b_n\in\mathbb{R}$ for every $n\geq 0$, and initial conditions $P_{-1}(x)=0$ and $P_0(x)=1$. Then there is a probability measure μ on the real line, s.t.

$$\int_{\mathbb{R}} P_n(x)P_m(x)\,d\mu(x) = \delta_{nm}.$$

proof:

Recall: Let ~~If~~ $A=A^T$, and λ_1, λ_2 are two distinct eigenvalues of A. If v_1, v_2 are corresponding eigenvectors, then $v_1^Tv_2=0$.

(Step 1). We first show that for polynomials whose degrees are less than or equal to n, we can construct a discrete measure μ_n, s.t. $\int_{\mathbb{R}} P_i(x)P_j(x)\,d\mu_n(x)=\delta_{ij}$, $0\leq i,j\leq n-1$. Consider the truncated Jacobi matrix J_n. It is known that zeros of $P_n(x)$ (i.e. $\{x_{n,j}\}_{j=1,\cdots,n}$) coincide with eigenvalues of J_n, with eigenvector $v_j=(P_0(x_{n,j}),\cdots,P_{n-1}(x_{n,j}))^T$. Therefore, we have $J_n(v_1,\cdots,v_n)=(v_1,\cdots,v_n)D$, where $D=\mathrm{diag}(x_{n,1},\cdots,x_{n,n})$. Since J_n is symmetric, we know that $\{v_j\}_{j=1,\cdots,n}$ form an orthogonal system. By normalization $\tilde{v}_j=v_j/\|v_j\|_2$, we can construct $\tilde{V}=(\tilde{v}_1,\cdots,\tilde{v}_n)$, which

Summary

is an orthogonal matrix, s.t. $\tilde{V}^T\tilde{V}=I$. Moreover, we have

$$\sum_{k=1}^n \frac{1}{\|v_k\|_2^2}P_i(x_{n,k})P_j(x_{n,k})=\delta_{ij}.$$

If we define a discrete measure

$$\mu_n=\sum_{k=1}^n \frac{1}{\|v_k\|_2^2}\delta_{x_{n,k}},$$

then this is a probability measure supported on the zeros of $P_n(x)$, for which $\int_{\mathbb{R}} P_i(x)P_j(x)\,d\mu_n(x)=\delta_{ij}$.
The first part of the proof is complete.

No.
Date.

Cues

This part is left to be reading materials.

Christoffel-Darboux kernel could be regarded as an application of 3-term recurrence relation, but it plays an important role in random matrix theory.

Notes

(step 2) We show that $\{\mu_n\}$ has a subsequence that converges weakly to a probability measure μ, s.t.

$$\lim_{k\to\infty}\int_{\mathbb{R}} p_i(x)p_j(x)\,d\mu_{n_k}(x)=\int_{\mathbb{R}} p_i(x)p_j(x)\,d\mu(x).$$

① Show the sequence of probability measure $\{\mu_n\}_{n=0,1,2,\cdots}$ has a subsequence that converges weakly to a measure μ (by Helly's selection principle);

② Show that μ is also a probability measure.
(Equivalent to show the sequence $\{\mu_n\}_{n=0,1,\cdots}$ is tight.
i.e. $\forall \varepsilon>0$, $\exists$ a compact set $K\subseteq\mathbb{R}$, s.t. $\mu_n(K)\geq 1-\varepsilon$ for every $n=0,1,2,\cdots$). #

3. Christoffel-Darboux kernel (C-D kernel)

Def. 2.4 (C-D kernel):

The Christoffel-Darboux kernel is defined by

$$\hat{K}_n(x,y)=\sum_{j=0}^{n-1} p_j(x)p_j(y), \quad (*)$$

where $\{p_j(x)\}_{j=0,1,2,\cdots}$ is the set of orthonormal polynomials.

Prop. 2.5 (Alternative form for C-D kernel):
The kernel $(*)$ satisfies the C-D identity

$$\hat{K}_n(x,y)=\begin{cases} a_n\dfrac{1}{x-y}\left(p_n(x)p_{n-1}(y)-p_{n-1}(x)p_n(y)\right), & x\neq y\\ a_n\left(p_n'(x)p_{n-1}(x)-p_n(x)p_{n-1}'(x)\right), & x=y\end{cases}$$

Summary

where a_n is the coefficient of 3-term recurrence relation.

No.
Date.

Cues

E.x.: Regarding with monic orthogonal polynomials $\{P_n(x)\}$, we have the C-D kernel

$$\tilde{K}_n(x,y)=\sum_{j=0}^{n-1}\frac{D_j}{D_{j+1}}P_j(x)P_j(y),$$

where $\int_{\mathbb{R}}P_j^2(x)\,d\mu(x)=\frac{D_{j+1}}{D_j}$.

Show that

$$(x-y)\tilde{K}_n(x,y)=\frac{D_{n-1}}{D_n}\times(P_n(x)P_{n-1}(y)-P_{n-1}(x)P_n(y)).$$

The proof of Prop. 2.6 is left to be an exercise.

In fact, if $q(x)\in\operatorname{span}\{P_0(x),\cdots,P_{n-1}(x)\}$, then $\int_{\mathbb{R}}\hat{K}_n(x,y)q(y)\,d\mu(y)=q(x)$.

Notes

proof: $$x\hat{K}_n(x,y)=\sum_{j=0}^{n-1}x\mathcal{P}_j(x)\mathcal{P}_j(y)$$
$$=\sum_{j=0}^{n-1}(a_{j+1}\mathcal{P}_{j+1}(x)+b_j\mathcal{P}_j(x)+a_j\mathcal{P}_{j-1}(x))\mathcal{P}_j(y).$$

To simplify this formula, we know

$$x\hat{K}_n(x,y)=a_n\mathcal{P}_n(x)\mathcal{P}_{n-1}(y)+\sum_{j=0}^{n-1}a_j(\mathcal{P}_j(x)\mathcal{P}_{j-1}(y)+\mathcal{P}_j(y)\mathcal{P}_{j-1}(x))+\sum_{j=0}^{n-1}b_j\mathcal{P}_j(x)\mathcal{P}_j(y),$$

where the last two terms are symmetric in x & y. Noting that $y\hat{K}_n(x,y)=x\hat{K}_n(x,y)|_{x\leftrightarrow y}$, we know that $(x-y)\hat{K}_n(x,y)=a_n(\mathcal{P}_n(x)\mathcal{P}_{n-1}(y)-\mathcal{P}_n(y)\mathcal{P}_{n-1}(x))$, and the result follows by the above formula. #

Prop. 2.6 (properties of C-D kernel):

The C-D kernel $\hat{K}_n(x,y)$ has the following properties:

① $\int_{\mathbb{R}}\hat{K}_n(\overset{x}{\cancel{y}},\cancel{y})\,d\mu(x)=n$;

② $\int_{\mathbb{R}}\hat{K}_n(x,y)\hat{K}_n(y,z)\,d\mu(y)=\hat{K}_n(x,z)$; (**)

③ $\int_{\mathbb{R}}\hat{K}_n(x,y)\mathcal{P}_k(y)\,d\mu(y)=\mathcal{P}_k(x)$, $0\le k\le n-1$.

We call $\hat{K}_n(x,y)$ as reproducing kernel in $L^2(d\mu)$ due to the equation (**).

Summary

In general, $\hat{K}_n(x,y)$ can induce an integral operator $\hat{K}_n: L^2(d\mu)\to L^2(d\mu)$ s.t.

$$\hat{K}_nf(x)=\int_{\mathbb{R}}\hat{K}_n(x,y)f(y)\,d\mu(y).$$

Remark 2.7: If $d\mu(x)=\omega(x)dx$, then $K_n(x,y)=\sqrt{\omega(x)}\sqrt{\omega(y)}\hat{K}_n(x,y)=\sqrt{\omega(x)}\sqrt{\omega(y)}\sum_{j=0}^{n-1}\mathcal{P}_j(x)$ which has the reproducing property in $L^2(\mathbb{R},d$ s.t.

$$\int_{-\infty}^{+\infty}K_n(x,y)K_n(y,z)\,dy=K_n(x,z).$$