

Properties of orthogonal polynomials

Kerstin Jordaan

University of South Africa

LMS Research School

University of Kent, Canterbury

- 1 Orthogonal polynomials
 - Gram-Schmidt orthogonalisation
 - The three-term recurrence relation
 - Jacobi operator
 - Hankel determinants
 - Hermite and Laguerre polynomials
- 2 Properties of classical orthogonal polynomials
- 3 Quasi-orthogonality and semiclassical orthogonal polynomials
- 4 The hypergeometric function
- 5 Convergence of Padé approximants for a hypergeometric function

The pioneer of orthogonality



Chebyshev Chebychev Chebyshev Tchebychev Tchebycheff Tschebyscheff

Murphy [1835] first defined orthogonal functions, Tchebychev realised their importance. His work since 1855 was motivated by the analogy with Fourier Series and by the theory of continued fractions and approximation theory.

The Tchebychev polynomials

$$T_n(x) = \cos n\theta \quad \text{where } x = \cos \theta \quad \text{for } n \in \mathbb{N}.$$

Consider

$$\int_0^\pi \cos m\theta \cos n\theta \, d\theta, \quad n, m \in \mathbb{N}.$$

For $m \neq n$,

$$\begin{aligned} & \int_0^\pi \cos m\theta \cos n\theta \, d\theta \\ &= \frac{1}{2} \int_0^\pi [\cos(m+n)\theta + \cos(m-n)\theta] \, d\theta \\ &= \frac{1}{2} \left[\frac{\sin(m+n)\theta}{m+n} + \frac{\sin(m-n)\theta}{m-n} \right]_0^\pi \\ &= 0. \end{aligned}$$

The Tchebychev polynomials

$$T_n(x) = \cos n\theta \quad \text{where} \quad x = \cos \theta \quad \text{for} \quad n \in \mathbb{N}.$$

Consider

$$\int_0^\pi \cos m\theta \cos n\theta \, d\theta, \quad n, m \in \mathbb{N}.$$

For $m = n$,

$$\begin{aligned} \int_0^\pi \cos m\theta \cos m\theta \, d\theta &= \int_0^\pi \cos^2 m\theta \, d\theta \\ &= \frac{1}{2} \int_0^\pi (1 + \cos 2m\theta) \, d\theta \\ &= \frac{1}{2} \left[\theta + \frac{\sin 2m\theta}{2m} \right]_0^\pi \\ &= \frac{\pi}{2}. \end{aligned}$$

The Tchebychev polynomials

$$T_n(x) = \cos n\theta \quad \text{where} \quad x = \cos \theta \quad \text{for} \quad n \in \mathbb{N}.$$

$$\int_0^\pi \cos m\theta \cos n\theta \, d\theta = \begin{cases} 0, & n \neq m \\ \frac{\pi}{2}, & m = n. \end{cases}$$

Making the substitution $x = \cos \theta$ in this integral, then $dx = -\sin \theta \, d\theta$ or

$$d\theta = \frac{-dx}{\sin \theta} = \frac{-dx}{\sqrt{1-x^2}}.$$

Also when $\theta = 0$, $x = 1$ and $\theta = \pi$, $x = -1$ so

$$\begin{aligned} \int_0^\pi \cos m\theta \cos n\theta \, d\theta &= \int_{-1}^1 T_n(x) T_m(x) (1-x^2)^{-1/2} dx \\ &= \begin{cases} 0, & n \neq m \\ \frac{\pi}{2}, & m = n. \end{cases} \end{aligned}$$

Definition

A sequence of polynomials $\{p_n(x)\}_{n=0}^{\infty}$ where $p_n(x)$ is of exact degree n , is called orthogonal on the interval (a, b) with respect to the positive weight function $w(x)$ if, for $m, n = 0, 1, 2, \dots$

$$\int_a^b p_n(x) p_m(x) w(x) dx = \begin{cases} 0 & \text{if } n \neq m \\ h_n \neq 0 & \text{if } n = m. \end{cases}$$

For Tchebychev polynomials

$$\int_{-1}^1 T_n(x) T_m(x) (1-x^2)^{-1/2} dx = \begin{cases} 0, & n \neq m \\ \frac{\pi}{2}, & m = n. \end{cases}$$

Tchebychev polynomials $\{T_n(x)\}_{n=0}^{\infty}$ are orthogonal on the interval $[-1, 1]$ with respect to the positive weight function $(1-x^2)^{-1/2}$.

- The interval (a, b) is called the interval of orthogonality and need not be finite. With due attention to convergence, either or both endpoints of the interval of orthogonality may be taken to be infinite.
- The limits of integration are important but the form in which the interval of orthogonality is stated is not vital.
- The weight function $w(x)$ should be continuous and positive on (a, b) so that the moments

$$\mu_n := \int_a^b w(x)x^n dx, \quad n = 0, 1, 2, \dots$$

exist.

- The weight function $w(x)$
 - does not change sign on the interval of orthogonality by assumption
 - may vanish at the finite endpoints (if any) of the interval of orthogonality

$w(x) \geq 0$ for all $x \in [a, b]$ and $w(x) > 0$ for all $x \in (a, b)$ is the usual **definition** of a weight function

- Because we have taken $w(x) > 0$ on (a, b) and $p_n(x)$ real, it follows that

$$h_n = \int_a^b w(x)p_n^2(x)dx \neq 0.$$

- The sequence of polynomial is uniquely defined up to normalization.
- If $h_n = 1$ for each $n = 0, 1, 2, \dots$ the sequence of polynomials is called orthonormal.
- If

$$p_n = k_n x^n + \text{lower order terms with } k_n = 1$$

for each $n = 0, 1, 2, \dots$, the sequence is called monic.

- The integral

$$\langle P_n, P_m \rangle := \int_a^b P_n(x)P_m(x)w(x)dx$$

denotes an inner product of the polynomials P_n and P_m .

More generally

Let μ be a positive Borel measure with support S defined on \mathbb{R} for which moments of all orders exist, i.e.

$$\mu_k = \int_S x^k d\mu(x), \quad k = 0, 1, 2, \dots \quad (1)$$

Definition

A sequence of real polynomials $\{P_n(x)\}_{n=0}^N$, $N \in \mathbb{N} \cup \{\infty\}$, where $P_n(x)$ is of exact degree n , is orthogonal with respect to μ on S , if

$$\langle P_n, P_m \rangle = \int_S P_n(x) P_m(x) d\mu(x) = h_n \delta_{mn}, \quad m, n = 0, 1, 2, \dots, N \quad (2)$$

where S is the support of μ and h_n is the square of the weighted L^2 -norm of P_n given by

$$h_n := \langle P_n, P_n \rangle = \|P_n\|^2 = \int_S (P_n(x))^2 d\mu(x) > 0.$$

If the measure is absolutely continuous and the distribution $d\mu(x) = w(x) dx$, then (2) reduces to

$$\int_a^b p_n(x) p_m(x) w(x) dx = h_n \delta_{mn}, \quad m, n = 0, 1, 2, \dots, N \quad (3)$$

or equivalently (see Assignment 1, Exercise 2),

$$\int_a^b x^m P_n(x) w(x) dx = 0, \text{ for } n = 1, 2, \dots; \quad m < n.$$

If the weight function $w(x)$ is discrete and $\rho_i > 0$ are the values of the weight at the distinct points x_i , $i = 0, 1, 2, \dots, M$, $M \in \mathbb{N} \cup \{\infty\}$, then (3) takes the form

$$\sum_{i=0}^M P_n(x_i) P_m(x_i) \rho_i = h_n \delta_{mn}, \quad m, n = 0, 1, 2, \dots, N$$

Gram-Schmidt orthogonalisation

Since the Hilbert space $L^2(S, \mu)$ contains the set of polynomials, Gram-Schmidt orthogonalisation applied to the canonical basis $\{1, x, x^2, \dots\}$, yields a set of orthogonal polynomials on the real line.

Example

Take $w(x) = 1$ and $(a, b) = (0, 1)$.

Start with the sequence $\{1, x, x^2, \dots\}$.

Choose $p_0(x) = 1$.

Then we have

$$p_1(x) = x - \frac{\langle x, p_0(x) \rangle}{\langle p_0(x), p_0(x) \rangle} p_0(x) = x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} = x - \frac{1}{2},$$

since

$$\langle 1, 1 \rangle = \int_0^1 1 \, dx = 1 \quad \text{and} \quad \langle x, 1 \rangle = \int_0^1 x \, dx = \frac{1}{2}.$$

Example

Further we have

$$\begin{aligned} p_2(x) &= x^2 - \frac{\langle x^2, p_0(x) \rangle}{\langle p_0(x), p_0(x) \rangle} - \frac{\langle x^2, p_1(x) \rangle}{\langle p_1(x), p_1(x) \rangle} p_1(x) \\ &= x^2 - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle} - \frac{\langle x^2, x - \frac{1}{2} \rangle}{\langle x - \frac{1}{2}, x - \frac{1}{2} \rangle} \left(x - \frac{1}{2} \right) \\ &= x^2 - \frac{1}{3} - \left(x - \frac{1}{2} \right) \\ &= x^2 - x + \frac{1}{6}, \end{aligned}$$

The polynomials $p_0(x) = 1$, $p_1(x) = x - \frac{1}{2}$ and $p_2(x) = x^2 - x + \frac{1}{6}$ are the first three monic orthogonal polynomials on the interval $(0, 1)$ with respect to the weight function $w(x) = 1$.

Example

Repeating this process we obtain

$$p_3(x) = x^3 - \frac{3}{2}x^2x - \frac{1}{20}$$

$$p_4(x) = x^4 - 2x^3 + \frac{9}{7}x^2 - \frac{2}{7}x + \frac{1}{70}$$

$$p_5(x) = x^5 - \frac{5}{2}x^4 + \frac{20}{9}x^3 - \frac{5}{6}x^2 + \frac{5}{42}x - \frac{1}{252},$$

and so on.

The orthonormal polynomials would be $q_0(x) = p_0(x)/\sqrt{h_0} = 1$,

$$q_1(x) = \frac{p_1(x)}{\sqrt{h_1}} = 2\sqrt{3}(x - 1/2)$$

$$q_2(x) = \frac{p_2(x)}{\sqrt{h_2}} = 6\sqrt{5} \left(x^2 - x + \frac{1}{6} \right)$$

$$p_3(x) = \frac{p_3(x)}{\sqrt{h_3}} = 20\sqrt{7} \left(x^2 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20} \right),$$

etcetera.

The three-term recurrence relation

The fact that $\langle xp, q \rangle = \langle p, xq \rangle$ gives rise to the following fundamental property of orthogonal polynomials.

Theorem

A sequence of orthogonal polynomials $\{P_n(x)\}$ satisfies a 3-term recurrence relation of the form.

$$P_{n+1}(x) = (A_n x + B_n) P_n(x) - C_n P_{n-1}(x) \quad \text{for } n = 0, 1, \dots \quad (4)$$

where we set $P_{-1}(x) \equiv 0$ and $P_0(x) \equiv 1$.

Here, A_n, B_n and C_n are real constants, $n = 0, 1, 2, \dots$

If the leading coefficient of $P_n(x)$ is $k_n > 0$, then

$$A_n = \frac{k_{n+1}}{k_n}, \quad C_{n+1} = \frac{A_{n+1} h_{n+1}}{A_n h_n}$$

Since $P_{n+1}(x)$ has degree exactly $(n+1)$ and so does $xP_n(x)$, we can determine A_n such that $P_{n+1}(x) - A_n x P_n(x)$ is a polynomial of degree at most n . Thus

$$P_{n+1}(x) - A_n x P_n(x) = \sum_{k=0}^n b_k P_k(x) \quad (5)$$

for some constants b_k . Now, if $Q(x)$ is **any** polynomial of degree $m < n$, we know from (3) that

$$\int_a^b P_n(x) Q(x) w(x) dx = 0.$$

If we multiply both sides of (5) by $w(x)P_m(x)$ where $m \in \{0, 1, \dots, n-2\}$, we obtain (upon integration)

$$\begin{aligned} & \int_a^b P_{n+1}(x) P_m(x) w(x) dx - A_n \int_a^b x P_n(x) P_m(x) w(x) dx \\ &= \sum_{k=0}^n \int_a^b b_k P_k(x) P_m(x) w(x) dx. \end{aligned}$$

$$\begin{aligned} \int_a^b P_{n+1}(x)P_m(x)w(x)dx - A_n \int_a^b xP_n(x)P_m(x)w(x)dx & \quad (6) \\ = \sum_{k=0}^n \int_a^b b_k P_k(x)P_m(x)w(x)dx. \end{aligned}$$

Now the left hand side of (6) is zero for each $m \in \{0, 1, \dots, n-2\}$ since then $xP_m(x)$ is a polynomial of degree $(m+1)$ which is less than or equal to $(n-1)$.

On the right hand side of (6), as k runs from 0 to n , the only integral in the sum that is not equal to zero is the one involving $k = m$.

Therefore $b_m h_m = 0$ for each $m \in \{0, 1, \dots, n-2\}$ and, since $h_m \neq 0$, we have $b_m = 0$, $m = 0, 1, \dots, n-2$.

Therefore, from

$$\begin{aligned} P_{n+1}(x) - A_n x P_n(x) &= \sum_{k=0}^n b_k P_k(x), \\ P_{n+1}(x) - A_n x P_n(x) &= b_{n-1} P_{n-1}(x) + b_n P_n(x), \end{aligned}$$

as required.

It is clear from the choice of A_n that $A_n = \frac{k_{n+1}}{k_n}$.

To prove the final part, multiply

$$P_{n+1}(x) = (A_n x + B_n) P_n(x) - C_n P_{n-1}(x)$$

by $P_{n-1}(x)w(x)$ and integrate, to obtain

$$0 = A_n \int_a^b x P_n(x) P_{n-1}(x) w(x) dx - C_n \int_a^b P_{n-1}^2(x) w(x) dx.$$

Now

$$P_{n-1}(x) = k_{n-1} x^{n-1} + (\text{poly of degree } \leq n-2) \quad (7)$$

and

$$P_n(x) = k_n(x)^n + (\text{poly of degree } \leq n-1)$$

Then

$$\begin{aligned} x P_{n-1}(x) &= k_{n-1}(x)^n + (\text{poly of degree } \leq n-1) \\ &= \frac{k_{n-1}}{k_n} k_n x^n + (\text{poly of degree } \leq n-1) \end{aligned}$$

More formally,

$$xP_{n-1}(x) = \frac{k_{n-1}}{k_n} P_n(x) + \sum_{k=0}^{n-1} d_k P_k(x).$$

From (7), we see that

$$\begin{aligned} 0 &= A_n \frac{k_{n-1}}{k_n} h_n - C_n h_{n-1}, \text{ or} \\ C_n &= A_n \frac{k_{n-1}}{k_n} \frac{h_n}{h_{n-1}}, \text{ so} \\ C_{n+1} &= A_{n+1} \frac{k_n}{k_{n+1}} \frac{h_{n+1}}{h_n} \end{aligned}$$

and since $\frac{k_{n+1}}{k_n} = A_n$, we have

$$C_{n+1} = \frac{A_{n+1}}{A_n} \frac{h_{n+1}}{h_n}.$$

Three-term recurrence for monic polynomials

Theorem

Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of monic orthogonal polynomials with respect to a positive measure μ . Then,

$$p_{n+1}(x) = (x - \alpha_n)p_n(x) - \beta_n p_{n-1}(x), \quad n = 0, 1, 2, \dots,$$

with initial conditions $p_{-1} \equiv 0$ and $p_0 \equiv 1$.

Notice that the choice of p_{-1} makes the initial value of β_0 irrelevant.

The recurrence coefficient α_n is given as:

$$\begin{aligned} \alpha_n &= \frac{\langle xp_n, p_n \rangle}{\langle p_n, p_n \rangle} \\ &= \frac{1}{h_n} \int_{\mathbb{R}} x p_n^2(x) d\mu(x), \quad n = 0, 1, \dots, \end{aligned}$$

If $p_n(x) = x^n + \ell_n x^{n-1} + \dots$, then, for each $n \in \mathbb{N}$,

$$\alpha_n = \ell_n - \ell_{n+1}$$

Three-term recurrence for monic polynomials

Theorem

Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of monic orthogonal polynomials with respect to a positive measure μ . Then,

$$p_{n+1}(x) = (x - \alpha_n)p_n(x) - \beta_n p_{n-1}(x), \quad n = 0, 1, 2, \dots,$$

with initial conditions $p_{-1} \equiv 0$ and $p_0 \equiv 1$.

The recurrence coefficient β_n is given as:

$$\begin{aligned}\beta_n &= \frac{\langle xp_n, p_{n-1} \rangle}{\langle p_{n-1}, p_{n-1} \rangle} = \frac{1}{h_{n-1}} \int_{\mathbb{R}} x p_n(x) p_{n-1}(x) d\mu(x) \\ &= \frac{\langle p_n, p_n \rangle}{\langle p_{n-1}, p_{n-1} \rangle} = \frac{h_n}{h_{n-1}} \\ &> 0, \quad n = 1, 2, \dots\end{aligned}$$

It follows that $h_n = \beta_n \beta_{n-1} \dots \beta_1$.

The converse: spectral theorem for orthogonal polynomials

Theorem

If a family of monic polynomials satisfies a three term recurrence relation of the form

$$xp_n(x) = p_{n+1}(x) + \alpha_n p_n(x) + \beta_n p_{n-1}(x)$$

with initial conditions $p_0(x) = 1$ and $p_{-1}(x) = 0$ where $\alpha_{n-1} \in \mathbb{R}$ and $\beta_n > 0$ for all $n \in \mathbb{N}$, then there exists a positive Borel measure μ on the real line such that these polynomials are monic orthogonal polynomials satisfying

$$\int_{\mathbb{R}} p_n(x)p_m(x) d\mu(x) = h_n \delta_{mn}, \quad m, n = 0, 1, 2, \dots$$

- Proof does not give explicit information about measure or support.
- Measure need not be unique and depends on Hamburger moment problem
- Can be traced back to earlier work on continued fractions with a rudimentary form given by Stieltjes in 1894;
- Also appears in books by Wintner [1929] and Stone [1932].
- Often referred to as Favard's theorem but was in fact independently discovered by Favard, Shohat and Natanson around 1935.

Jacobi matrix

Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of monic orthogonal polynomials satisfying

$$p_{n+1}(x) = (x - \alpha_n)p_n(x) - \beta_n p_{n-1}(x), \quad n = 0, 1, 2, \dots,$$

with $p_{-1} = 0$ and $p_0 = 1$.

The recurrence coefficients can be collected in a tridiagonal matrix of the form

$$J = \begin{pmatrix} \alpha_0 & \sqrt{\beta_1} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & \\ & & \sqrt{\beta_3} & \alpha_3 & \ddots \\ & & & \ddots & \ddots \end{pmatrix}$$

known as the Jacobi matrix or Jacobi operator which acts as an operator (on a subset of) $\ell^2(\mathbb{N})$.

Hankel determinants

The coefficients in the three-term recurrence relation can also be expressed in terms of determinants whose entries are moments associated with measure μ .

$$\alpha_n = \frac{\tilde{\Delta}_{n+1}}{\Delta_{n+1}} - \frac{\tilde{\Delta}_n}{\Delta_n}, \quad \beta_n = \frac{\Delta_{n+1}\Delta_{n-1}}{\Delta_n^2},$$

where Δ_n is the Hankel determinant

$$\Delta_n = \det \left[\mu_{j+k} \right]_{j,k=0}^{n-1} = \begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-1} \\ \mu_1 & \mu_2 & \dots & \mu_n \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \dots & \mu_{2n-2} \end{vmatrix}, \quad n \geq 1,$$

with $\Delta_0 = 1$, $\Delta_{-1} = 0$, and $\tilde{\Delta}_n$ is the determinant

$$\tilde{\Delta}_n = \begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-2} & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_{n-1} & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \dots & \mu_{2n-3} & \mu_{2n-1} \end{vmatrix}, \quad n \geq 1,$$

with $\tilde{\Delta}_0 = 0$ and μ_k is the k th moment.

The monic polynomial $p_n(x)$ can be uniquely expressed as the determinant

$$p_n(x) = \frac{1}{\Delta_n} \begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \dots & \mu_{2n-1} \\ 1 & x & \dots & x^n \end{vmatrix},$$

The normalisation constants are given by

$$h_n = \frac{\Delta_{n+1}}{\Delta_n}, \quad h_0 = \Delta_1 = \mu_0.$$

Remark

$\Delta_n > 0$ ($h_n > 0$), $n \geq 1$ corresponds to a positive definite moment functional and orthogonal polynomials in the usual sense.

A more general notion of orthogonality can be defined for quasi-definite moment functionals when $\Delta_n \neq 0$.

Note that when the moments are non-real, the definition bears no relation to the standard concept of orthogonality of polynomials in a complex variable.

Hermite polynomials

The polynomials orthogonal with respect to the normal distribution e^{-x^2} are the Hermite polynomials, named for the French mathematician Charles Hermite (1822 – 1901).

Definition

The Hermite polynomials are denoted $H_n(x)$ and are defined by the generating function

$$e^{2xt-t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)t^n}{n!}$$

valid for all finite x and t .

Theorem

The Hermite polynomials can be represented explicitly by

$$H_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k n!}{k!(n-2k)!} (2x)^{n-2k}.$$

Theorem

The orthogonality property of $H_n(x)$ is

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2} dx = 2^n n! \sqrt{\pi} \delta_{nm},$$

i.e. the Hermite polynomials are orthogonal on the real line with respect to the normal distribution.

Theorem

The three-term recurrence relation for the Hermite polynomials is given by

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x) \quad n \geq 1.$$

Laguerre Polynomials

Laguerre polynomials, named for the French mathematician Edmond Nicolas Laguerre (1834 – 1886).

Definition

Laguerre polynomials are denoted $L_n^\alpha(x)$ and are defined by the generating function

$$(1-t)^{-\alpha-1} \exp\left(\frac{-xt}{1-t}\right) = \sum_{n=0}^{\infty} L_n^\alpha(x) t^n.$$

Theorem

The Laguerre polynomials can be represented explicitly by

$$L_n^\alpha(x) = \frac{(\alpha+1)_n}{n!} \sum_{k=0}^n \frac{(-n)_k x^k}{(\alpha+1)_k k!}$$

where $(a)_t$ is Pochhammer's symbol $(a)_t = a(a+1)\dots(a+t-1)$.

Theorem

The Laguerre polynomials are orthogonal on the positive real line with respect to the gamma distribution, i.e. the orthogonality relation for the Laguerre polynomials is contained in

$$\int_0^{\infty} L_n^{\alpha}(x)L_m^{\alpha}(x)x^{\alpha}e^{-x}dx = \frac{\Gamma(\alpha+n+1)}{n!}\delta_{mn}$$

for $\alpha > -1$.

Theorem

The Laguerre polynomials satisfy the three term recurrence relation given by

$$(n+1)L_{n+1}^{\alpha}(x) = (1+2n+\alpha-x)L_n^{\alpha}(x) - (n+\alpha)L_{n-1}^{\alpha}(x).$$

Remark

A Laguerre polynomial involves a parameter α . The Hermite polynomials did not rely on any parameters.

