CAMBRIDGE

Cambridge University Press
978-1-107-07189-6 - Encyclopedia of Mathematics and its Applications: Orthogonal Polynomials of Several Variables: Second Edition
Charles F. Dunkl and Yuan Xu
Excerpt
More information

1

Background

The theory of orthogonal polynomials of several variables, especially those of classical type, uses a significant amount of analysis in one variable. In this chapter we give concise descriptions of the needed tools.

1.1 The Gamma and Beta Functions

It is our opinion that the most interesting and amenable objects of consideration have expressions which are rational functions of the underlying parameters. This leads us immediately to a discussion of the gamma function and its relatives.

Definition 1.1.1 The gamma function is defined for $\operatorname{Re} x > 0$ by the integral

$$\Gamma(x) = \int_0^\infty t^{x-1} \mathrm{e}^{-t}\, \mathrm{d}t.$$

It is directly related to the beta function:

Definition 1.1.2 The beta function is defined for $\operatorname{Re} x > 0$ and $\operatorname{Re} y > 0$ by

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t.$$

By making the change of variables $s = uv$ and $t = (1-u)v$ in the integral $\Gamma(x)\Gamma(y) = \int_0^\infty \int_0^\infty s^{x-1} t^{y-1} \mathrm{e}^{-(s+t)} \mathrm{d}s\, \mathrm{d}t$, one obtains

$$\Gamma(x)\Gamma(y) = \Gamma(x+y)B(x,y).$$

This leads to several useful definite integrals, valid for $\operatorname{Re} x > 0$ and $\operatorname{Re} y > 0$:

1. $\displaystyle\int_0^{\pi/2} \sin^{x-1}\theta \cos^{y-1}\theta\, \mathrm{d}\theta = \tfrac{1}{2} B\left(\tfrac{x}{2}, \tfrac{y}{2}\right) = \frac{\frac{1}{2}\Gamma\left(\frac{x}{2}\right)\Gamma\left(\frac{y}{2}\right)}{\Gamma\left(\frac{x+y}{2}\right)};$

2. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ (set $x = y = 1$ in the previous integral);
3. $\displaystyle\int_0^\infty t^{x-1}\exp(-at^2)\,dt = \tfrac{1}{2}a^{-x/2}\Gamma\left(\tfrac{x}{2}\right)$, for $a > 0$;
4. $\displaystyle\int_0^1 t^{x-1}(1-t^2)^{y-1}\,dt = \tfrac{1}{2}B\left(\tfrac{x}{2}, y\right) = \frac{\frac{1}{2}\Gamma\left(\frac{x}{2}\right)\Gamma(y)}{\Gamma\left(\frac{x}{2}+y\right)}$;
5. $\Gamma(x)\Gamma(1-x) = B(x, 1-x) = \dfrac{\pi}{\sin \pi x}$.

The last equation can be proven by restricting x to $0 < x < 1$, making the substitution $s = t/(1-t)$ in the beta integral $\int_0^1 [t/(1-t)]^{x-1}(1-t)^{-1}\,dt$ and computing the resulting integral by residues. Of course, one of the fundamental properties of the gamma function is the recurrence formula (obtained from integration by parts)

$$\Gamma(x+1) = x\Gamma(x),$$

which leads to the fact that Γ can be analytically continued to a meromorphic function on the complex plane; also, $1/\Gamma$ is entire, with (simple) zeros exactly at $\{0, -1, -2, \ldots\}$. Note that Γ interpolates the factorial; indeed, $\Gamma(n+1) = n!$ for $n = 0, 1, 2, \ldots$

Definition 1.1.3 The Pochhammer symbol, also called the shifted factorial, is defined for all x by

$$(x)_0 = 1, \quad (x)_n = \prod_{i=1}^{n}(x+i-1) \qquad \text{for } n = 1, 2, 3, \ldots$$

Alternatively, one can recursively define $(x)_n$ by

$$(x)_0 = 1 \quad \text{and} \quad (x)_{n+1} = (x)_n(x+n) \qquad \text{for } n = 1, 2, 3, \ldots$$

Here are some important consequences of Definition 1.1.3:

1. $(x)_{m+n} = (x)_m(x+m)_n$ for $m, n \in \mathbb{N}_0$;
2. $(x)_n = (-1)^n(1-n-x)_n$ (writing the product in reverse order);
3. $(x)_{n-i} = (x)_n(-1)^i/(1-n-x)_i$.

The Pochhammer symbol incorporates binomial-coefficient and factorial notation:

1. $(1)_n = n!$, $2^n\left(\frac{1}{2}\right)_n = 1 \times 3 \times 5 \times \cdots \times (2n-1)$;
2. $(n+m)! = n!(n+1)_m$;
3. $\dbinom{n}{i} = (-1)^i \frac{(-n)_i}{i!}$, where $\binom{n}{i}$ is the binomial coefficient;
4. $(x)_{2n} = 2^{2n}\left(\frac{x}{2}\right)_n\left(\frac{x+1}{2}\right)_n$, the duplication formula.

The last property includes the formula for the gamma function:

$$\Gamma(2x) = \frac{2^{2x-1}}{\sqrt{\pi}} \Gamma(x+\tfrac{1}{2})\Gamma(x).$$

For appropriate values of x and n, the formula $\Gamma(x+n)/\Gamma(x) = (x)_n$ holds, and this can be used to extend the definition of the Pochhammer symbol to values of $n \notin \mathbb{N}_0$.

1.2 Hypergeometric Series

The two most common types of hypergeometric series are (they are convergent for $|x| < 1$)

$$ {}_1F_0(a;x) = \sum_{n=0}^{\infty} \frac{(a)_n}{n!} x^n,$$

$$ {}_2F_1\left(\begin{matrix} a, b \\ c \end{matrix}; x\right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n n!} x^n,$$

where a and b are the "numerator" parameters and c is the "denominator" parameter. Later we will also use ${}_3F_2$ series (with a corresponding definition). The ${}_2F_1$ series is the unique solution analytic at $x = 0$ and satisfying $f(0) = 1$ of

$$x(1-x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) + [c - (a+b+1)x]\frac{\mathrm{d}}{\mathrm{d}x} f(x) - abf(x) = 0.$$

Generally, classical orthogonal polynomials can be expressed as hypergeometric polynomials, which are terminating hypergeometric series for which a numerator parameter has a value in $-\mathbb{N}_0$. The two series can be represented in closed form. Obviously ${}_1F_0(a;x) = (1-x)^{-a}$; this is the branch analytic in $\{x \in \mathbb{C} : |x| < 1\}$, which has the value 1 at $x = 0$. The Gauss integral formula for ${}_2F_1$ is as follows.

Proposition 1.2.1 *For* $\mathrm{Re}(c-b) > 0$, $\mathrm{Re}\, b > 0$ *and* $|x| < 1$,

$$ {}_2F_1\left(\begin{matrix} a, b \\ c \end{matrix}; x\right) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-xt)^{-a}\,\mathrm{d}t.$$

Proof Use the ${}_1F_0$ series in the integral and integrate term by term to obtain a multiple of

$$\sum_{n=0}^{\infty} \frac{(a)_n}{n!} x^n \int_0^1 t^{b+n-1}(1-t)^{c-b-1}\mathrm{d}t = \sum_{n=0}^{\infty} \frac{(a)_n \Gamma(b+n)\Gamma(c-b)}{n!\Gamma(c+n)} x^n,$$

from which the stated formula follows. □

Corollary 1.2.2 *For* $\operatorname{Re} c > \operatorname{Re}(a+b)$ *and* $\operatorname{Re} b > 0$ *the Gauss summation formula is*

$$ {}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix}; 1\right) = \frac{\Gamma(c)\Gamma(c-b-a)}{\Gamma(c-a)\Gamma(c-b)}. $$

The terminating case of this formula, known as the Chu–Vandermonde sum, is valid for a more general range of parameters.

Proposition 1.2.3 *For* $n \in \mathbb{N}_0$*, any* a, b*, and* $c \neq 0, 1, \ldots, n-1$ *the following hold:*

$$ \sum_{i=0}^{n} \frac{(a)_{n-i}(b)_i}{(n-i)!\,i!} = \frac{(a+b)_n}{n!} \quad \textit{and} \quad {}_2F_1\left(\begin{matrix} -n,b \\ c \end{matrix}; 1\right) = \frac{(c-b)_n}{(c)_n}. $$

Proof The first formula is deduced from the coefficient of x^n in the expression $(1-x)^{-a}(1-x)^{-b} = (1-x)^{-(a+b)}$. The left-hand side can be written as

$$ \frac{(a)_n}{n!} \sum_{i=0}^{n} \frac{(-n)_i(b)_i}{(1-n-a)_i\, i!}. $$

Now let $a = 1-n-c$; simple computations involving reversals such as $(1-n-c)_n = (-1)^n(c)_n$ finish the proof. □

The following transformation often occurs:

Proposition 1.2.4 *For* $|x| < 1$,

$$ {}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix}; x\right) = (1-x)^{-a}\, {}_2F_1\left(\begin{matrix} a, c-b \\ c \end{matrix}; \frac{x}{x-1}\right). $$

Proof Temporarily assume that $\operatorname{Re} c > \operatorname{Re} b > 0$; then from the Gauss integral we have

$$ \begin{aligned} {}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix}; x\right) &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-xt)^{-a}\,\mathrm{d}t \\ &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 s^{c-b-1}(1-s)^{b-1}(1-x)^{-a}\left(1-\frac{xs}{x-1}\right)^{-a}\mathrm{d}s, \end{aligned} $$

where one makes the change of variable $t = 1-s$. The formula follows from another application of the Gauss integral. Analytic continuation in the parameters extends the validity to all values of a, b, c excluding $c \in \mathbb{N}_0$. For this purpose we tacitly consider the modified series

$$ \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{\Gamma(c+n)\,n!} x^n, $$

which is an entire function in a, b, c. □

Corollary 1.2.5 *For* $|x| < 1$,

$$ {}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix}; x\right) = (1-x)^{c-a-b}\,{}_2F_1\left(\begin{matrix} c-a, c-b \\ c \end{matrix}; x\right). $$

Proof Using Proposition 1.2.4 twice,

$$ \begin{aligned} {}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix}; x\right) &= (1-x)^{-a}\,{}_2F_1\left(\begin{matrix} a, c-b \\ c \end{matrix}; \frac{x}{x-1}\right) \\ &= (1-x)^{-a}\left(1-\frac{x}{x-1}\right)^{b-c}{}_2F_1\left(\begin{matrix} c-a, c-b \\ c \end{matrix}; x\right) \end{aligned} $$

and $1 - x/(x-1) = (1-x)^{-1}$. □

Equating coefficients of x^n on the two sides of the formula in Corollary 1.2.5 proves the *Saalschütz summation formula* for a balanced terminating ${}_3F_2$ series.

Proposition 1.2.6 *For* $n = 0,1,2,\ldots$ *and* $c,d \neq 0,-1,-2,\ldots,-n$ *with* $-n+a+b+1 = c+d$ *(the "balanced" condition), we have*

$$ {}_3F_2\left(\begin{matrix} -n,a,b \\ c,d \end{matrix}; 1\right) = \frac{(c-a)_n\,(c-b)_n}{(c)_n\,(c-a-b)_n} = \frac{(c-a)_n\,(d-a)_n}{(c)_n\,(d)_n}. $$

Proof Considering the coefficient of x^n in the equation

$$ (1-x)^{a+b-c}\,{}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix}; x\right) = {}_2F_1\left(\begin{matrix} c-a, c-b \\ c \end{matrix}; x\right) $$

yields

$$ \sum_{j=0}^{n} \frac{(c-a-b)_{n-j}\,(a)_j\,(b)_j}{(n-j)!\,j!\,(c)_j} = \frac{(c-a)_n\,(c-b)_n}{n!\,(c)_n}, $$

but

$$ \frac{(c-a-b)_{n-j}}{(n-j)!} = \frac{(c-a-b)_n\,(-n)_j}{n!\,(1-n-c+a+b)_j}; $$

this proves the first formula with $d = -n+a+b-c+1$. Further, $(c-b)_n = (-1)^n(1-n-c+b)_n = (-1)^n(d-a)_n$ and $(-1)^n\,(c-a-b)_n = (1-n-c+a+b)_n = (d)_n$, which proves the second formula. □

1.2.1 Lauricella series

There are many ways to define multivariable analogues of the hypergeometric series. One straightforward and useful approach consists of the Lauricella generalizations of the ${}_2F_1$ series; see Exton [1976]. Fix $d = 1,2,3,\ldots$ vector parameters $\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbb{R}^d$, scalar parameters α,γ and the variable $x \in \mathbb{R}^d$. For concise formulation we use the following: let $\mathbf{m} \in \mathbb{N}_0^d$, $\mathbf{m}! = \prod_{j=1}^d (m_j)!$, $|\mathbf{m}| = \sum_{j=1}^d m_j$, $(\mathbf{a})_\mathbf{m} = \prod_{j=1}^d (a_j)_{m_j}$ and $x^\mathbf{m} = \prod_{j=1}^d x_j^{m_j}$.

The four types of Lauricella functions are (with summations over $\mathbf{m} \in \mathbb{N}_0^d$):

1. $F_A(\alpha, \mathbf{b}; \mathbf{c}; x) = \sum_\mathbf{m} \frac{(\alpha)_{|\mathbf{m}|} (\mathbf{b})_\mathbf{m}}{(\mathbf{c})_\mathbf{m}\, \mathbf{m}!} x^\mathbf{m}$, convergent for $\sum_{j=1}^d |x_j| < 1$;
2. $F_B(\mathbf{a}, \mathbf{b}; \gamma; x) = \sum_\mathbf{m} \frac{(\mathbf{a})_\mathbf{m} (\mathbf{b})_\mathbf{m}}{(\gamma)_{|\mathbf{m}|}\, \mathbf{m}!} x^\mathbf{m}$, convergent for $\max_j |x_j| < 1$;
3. $F_C(\alpha, \beta; \mathbf{c}; x) = \sum_\mathbf{m} \frac{(\alpha)_{|\mathbf{m}|} (\beta)_{|\mathbf{m}|}}{(\mathbf{c})_\mathbf{m} \mathbf{m}!} x^\mathbf{m}$, convergent for $\sum_{j=1}^d |x_j|^{1/2} < 1$;
4. $F_D(\alpha, \mathbf{b}; \gamma; x) = \sum_\mathbf{m} \frac{(\alpha)_{|\mathbf{m}|} (\mathbf{b})_\mathbf{m}}{(\gamma)_{|\mathbf{m}|}\, \mathbf{m}!} x^\mathbf{m}$, convergent for $\max_j |x_j| < 1$.

There are integral representations of Euler type (the following are subject to obvious convergence conditions; further, any argument of a gamma function must have a positive real part):

1. $$F_A(\alpha, \mathbf{b}; \mathbf{c}; x) = \prod_{j=1}^d \frac{\Gamma(c_j)}{\Gamma(b_j)\Gamma(c_j - b_j)} \times \int_{[0,1]^d} \prod_{j=1}^d \left(u_j^{b_j - 1} (1-u_j)^{c_j - b_j - 1}\right) \left(1 - \sum_{j=1}^d u_j x_j\right)^{-\alpha} \mathrm{d}u;$$
2. $$F_B(\mathbf{a}, \mathbf{b}; \gamma; x) = \prod_{j=1}^d \Gamma(a_j)^{-1} \frac{\Gamma(\gamma)}{\Gamma(\delta)} \times \int_{T^d} \prod_{j=1}^d \left(u_j^{a_j - 1} (1 - u_j x_j)^{-b_j}\right) \left(1 - \sum_{j=1}^d u_j\right)^{\delta - 1} \mathrm{d}u,$$
   where $\delta = \gamma - \sum_{j=1}^d a_j$ and T^d is the simplex $\{u \in \mathbb{R}^d : u_j \geq 0$ for all j, and $\sum_{j=1}^d u_j \leq 1\}$;
3. $$F_D(\alpha, \mathbf{b}; \gamma; x) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma - \alpha)} \times \int_0^1 u^{\alpha - 1} (1-u)^{\gamma - \alpha - 1} \prod_{j=1}^d (1 - u x_j)^{-b_j} \mathrm{d}u,$$
   a single integral.

1.3 Orthogonal Polynomials of One Variable

1.3.1 General properties

We start with a determinant approach to the Gram–Schmidt process, a method for producing orthogonal bases of functions given a linearly (totally) ordered basis. Suppose that X is a region in $\mathbb{R}^d$ (for $d \geq 1$), μ is a probability measure on X

and $\{f_i(x) : i = 1,2,3\ldots\}$ is a set of functions linearly independent in $L^2(X,\mu)$. Denote the inner product $\int_X fg\,\mathrm{d}\mu$ as $\langle f,g\rangle$ and the elements of the Gram matrix $\langle f_i, f_j\rangle$ as g_{ij}, $i, j \in \mathbb{N}$.

Definition 1.3.1 For $n \in \mathbb{N}$ let $d_n = \det(g_{ij})_{i,j=1}^n$ and

$$D_n(x) = \det \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ g_{n-1,1} & g_{n-1,2} & \cdots & g_{n-1,n} \\ f_1(x) & f_2(x) & \ldots & f_n(x) \end{bmatrix}.$$

Proposition 1.3.2 *The functions* $\{D_n(x) : n \ge 1\}$ *are orthogonal in* $L^2(X,\mu)$, $\operatorname{span}\{D_j(x) : 1 \le j \le n\} = \operatorname{span}\{f_j(x) : 1 \le j \le n\}$ *and* $\langle D_n, D_n\rangle = d_{n-1}d_n$.

Proof By linear independence, $d_n > 0$ for all n; thus $D_n(x) = d_{n-1}f_n(x) + \sum_{j<n} c_j f_j(x)$ for some coefficients c_j (where $d_0 = 1$) and $\operatorname{span}\{D_j : j \le n\} = \operatorname{span}\{f_j : j \le n\}$. The inner product $\langle f_j, D_n\rangle$ is the determinant of the matrix in the definition of D_n with the last row replaced by $(g_{j1}, g_{j2}, \ldots, g_{jn})$ and hence is zero for $j < n$. Thus $\langle D_j, D_n\rangle = 0$ for $j < n$ and $\langle D_n, D_n\rangle = d_{n-1}\langle f_n, D_n\rangle = d_{n-1}d_n$. □

There are integral formulae for d_n and $D_n(x)$ which are interesting foreshadowings of multivariable weight functions P_n involving the discriminant, as follows.

Definition 1.3.3 For $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n \in X$ let

$$P_n(x_1, x_2, \ldots, x_n) = \det(f_j(x_i))_{i,j=1}^n.$$

Proposition 1.3.4 *For* $n \in \mathbb{N}$ *and* $x_1, x_2, \ldots, x_n \in X$,

$$\int_{X^n} P_n(x_1, x_2, \ldots, x_n)^2\,\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_n) = n!d_n,$$

and

$$\int_{X^n} P_n(x_1, x_2, \ldots, x_n)P_{n+1}(x_1, x_2, \ldots, x_n, x)\,\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_n) = n!D_{n+1}(x).$$

Proof In the first integral, make the expansion

$$P_n(x_1, x_2, \ldots, x_n)^2 = \sum_\sigma\sum_\tau \varepsilon_\sigma\varepsilon_\tau \prod_{i=1}^n f_{\sigma i}(x_i) f_{\tau i}(x_i),$$

where the summations are over the symmetric group S_n (on n objects); $\varepsilon_\sigma, \sigma i$ denote the sign of the permutation σ and the action of σ on i, respectively. Integrating over X^n gives the sum $\sum_\sigma\sum_\tau \varepsilon_\sigma\varepsilon_\tau \prod_{i=1}^n g_{\sigma i,\tau i} = n!\sum_\tau \varepsilon_\tau \prod_{i=1}^n g_{\tau i,i} = n!d_n$.

The summation over σ is done by first fixing σ and then replacing i by $\sigma^{-1}i$ and τ by $\tau\sigma^{-1}$. Similarly,

$$P_n(x_1,x_2,\ldots,x_n)P_{n+1}(x_1,x_2,\ldots,x_n,x) = \sum_\sigma\sum_\tau \varepsilon_\sigma\varepsilon_\tau \prod_{i=1}^n f_{\sigma i}(x_i)f_{\tau i}(x_i)f_{\tau(n+1)}(x),$$

and the τ-sum is over S_{n+1}. As before, the integral has the value

$$\sum_\sigma\sum_\tau \varepsilon_\sigma\varepsilon_\tau \prod_{i=1}^n g_{\sigma i,\tau i} f_{\tau(n+1)}(x),$$

which reduces to the expression $n!\sum_\tau \varepsilon_\tau \prod_{i=1}^n g_{\tau i,i} f_{\tau(n+1)}(x) = n! D_{n+1}(x)$. □

We now specialize to orthogonal polynomials; let μ be a probability measure supported on a (possibly infinite) interval $[a,b]$ such that $\int_a^b |x|^n \,\mathrm{d}\mu < \infty$ for all n. We may as well assume that μ is not a finite discrete measure, so that $\{1,x,x^2,x^3,\ldots\}$ is linearly independent in $L^2(\mu)$; it is not difficult to modify the results to the situation where $L^2(\mu)$ is of finite dimension. We apply Proposition 1.3.2 to the basis $f_j(x) = x^{j-1}$; the Gram matrix has the form of a Hankel matrix $g_{ij} = c_{i+j-2}$, where the nth moment of μ is

$$c_n = \int_a^b x^n \,\mathrm{d}\mu(x)$$

and the orthonormal polynomials $\{p_n(x) : n \geq 0\}$ are defined by

$$p_n(x) = (d_{n+1}d_n)^{-1/2} D_{n+1}(x);$$

they satisfy $\int_a^b p_m(x)p_n(x)\,\mathrm{d}\mu(x) = \delta_{mn}$, and the leading coefficient of p_n is $(d_n/d_{n+1})^{1/2} > 0$. Of course this implies that $\int_a^b p_n(x)q(x)\,\mathrm{d}\mu(x) = 0$ for any polynomial $q(x)$ of degree $\leq n-1$. The determinant P_n in Definition 1.3.3 is exactly the Vandermonde determinant $\det(x_i^{j-1})_{i,j=1}^n = \prod_{1\leq i<j\leq n}(x_j - x_i)$.

Proposition 1.3.5 *For $n \geq 0$,*

$$\int_{[a,b]^n} \prod_{1\leq i<j\leq n} (x_j - x_i)^2 \,\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_n) = n!d_n,$$

$$\int_{[a,b]^n} \prod_{i=1}^n (x - x_i) \prod_{1\leq i<j\leq n} (x_j - x_i)^2 \,\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_n) = n!(d_n d_{n+1})^{1/2} p_n(x).$$

It is a basic fact that $p_n(x)$ has n distinct (and simple) zeros in $[a,b]$.

Proposition 1.3.6 *For $n \geq 1$, the polynomial $p_n(x)$ has n distinct zeros in the open interval (a,b).*

Proof Suppose that $p_n(x)$ changes sign at $t_1,\ldots,t_m$ in (a,b). Then it follows that $\varepsilon p_n(x)\prod_{i=1}^m(x-t_i) \geq 0$ on $[a,b]$ for $\varepsilon = 1$ or -1. If $m < n$ then $\int_a^b p_n(x) \prod_{i=1}^m(x-t_i)\,\mathrm{d}\mu(x) = 0$, which implies that the integrand is zero on the support of μ, a contradiction. □

In many applications one uses orthogonal, rather than orthonormal, polynomials (by reason of their neater notation, generating function and normalized values at an end point, for example). This means that we have a family of nonzero polynomials $\{P_n(x) : n \geq 0\}$ with $P_n(x)$ of exact degree n and for which $\int_a^b P_n(x)x^j\,\mathrm{d}\mu(x) = 0$ for $j < n$.

We say that the squared norm $\int_a^b P_n(x)^2\,\mathrm{d}\mu(x) = h_n$ is a *structural constant*. Further, $p_n(x) = \pm h_n^{-1/2} P_n(x)$; the sign depends on the leading coefficient of $P_n(x)$.

1.3.2 Three-term recurrence

Besides the Gram matrix of moments there is another important matrix associated with a family of orthogonal polynomials, the Jacobi matrix. The principal minors of this tridiagonal matrix provide an interpretation of the three-term recurrence relations. For $n \geq 0$ the polynomial $xP_n(x)$ is of degree $n+1$ and can be expressed in terms of $\{P_j : j \leq n+1\}$, but more is true.

Proposition 1.3.7 *There exist sequences $\{A_n\}_{n\geq 0}, \{B_n\}_{n\geq 0}, \{C_n\}_{n\geq 1}$ such that*

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

where

$$A_n = \frac{k_{n+1}}{k_n}, \qquad C_n = \frac{k_{n+1}k_{n-1}h_n}{k_n^2 h_{n-1}}, \qquad B_n = -\frac{k_{n+1}}{k_n h_n}\int_a^b xP_n(x)^2\,\mathrm{d}\mu(x),$$

and k_n is the leading coefficient of $P_n(x)$.

Proof Expanding $xP_n(x)$ in terms of polynomials P_j gives $\sum_{j=0}^{n+1} a_j P_j(x)$ with $a_j = h_j^{-1}\int_a^b xP_n(x)P_j(x)\,\mathrm{d}\mu(x)$. By the orthogonality property, $a_j = 0$ unless $|n-j| \leq 1$. The value of $a_{n+1} = A_n^{-1}$ is obtained by matching the coefficients of x^{n+1}. Shifting the label gives the value of C_n. □

Corollary 1.3.8 *For the special case of monic orthogonal polynomials, the three-term recurrence is*

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

where

$$C_n = \frac{d_{n+1}d_{n-1}}{d_n^2} \quad \text{and} \quad B_n = -\frac{d_n}{d_{n+1}} \int_a^b x P_n(x)^2 \, d\mu(x).$$

Proof In the notation from the end of the last subsection, the structure constant for the monic case is $h_n = d_{n+1}/d_n$. □

It is convenient to restate the recurrence, and some other, relations for orthogonal polynomials with arbitrary leading coefficients in terms of the moment determinants d_n (see Definition 1.3.1).

Proposition 1.3.9 *Suppose that the leading coefficient of $P_n(x)$ is k_n, and let $b_n = \int_a^b x p_n(x)^2 \, d\mu(x)$; then*

$$h_n = k_n^2 \frac{d_{n+1}}{d_n},$$

$$xP_n(x) = \frac{k_n}{k_{n+1}} P_{n+1}(x) + b_n P_n(x) + \frac{k_{n-1}h_n}{k_n h_{n-1}} P_{n-1}(x).$$

Corollary 1.3.10 *For the case of orthonormal polynomial p_n,*

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

where $a_n = k_n/k_{n+1} = (d_n d_{n+2}/d_{n+1}^2)^{1/2}$.

With these formulae one can easily find the reproducing kernel for polynomials of degree $\le n$, the Christoffel–Darboux formula:

Proposition 1.3.11 *For $n \ge 1$, if k_n is the leading coefficient of p_n then we have*

$$\sum_{j=0}^{n} p_j(x)p_j(y) = \frac{k_n}{k_{n+1}} \frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x - y},$$

$$\sum_{j=0}^{n} p_j(x)^2 = \frac{k_n}{k_{n+1}} \left[p'_{n+1}(x)p_n(x) - p'_n(x)p_{n+1}(x) \right].$$

Proof By the recurrence in the previous proposition, for $j \ge 0$,

$$(x-y)p_j(x)p_j(y) = \frac{k_j}{k_{j+1}} \left[p_{j+1}(x)p_j(y) - p_j(x)p_{j+1}(y) \right] + \frac{k_{j-1}}{k_j} \left[p_{j-1}(x)p_j(y) - p_j(x)p_{j-1}(y) \right].$$

The terms involving b_j arising from setting $n = j$ in Corollary 1.3.10 cancel out. Now sum these equations over $0 \le j \le n$; the terms telescope (and note that the case $j = 0$ is special, with $p_{-1} = 0$). This proves the first formula in the proposition; the second follows from L'Hospital's rule. □