Chapter 6. Orthogonal Polynomials

In Hilbertian spaces $L^2_w(a,b)$, the introduction of a basis of orthogonal polynomials $\{P_n(x)\}$ allows one to obtain, in a constructive way, the so-called best approximation of the functions $f(x)$ of the space by a finite linear combination of polynomials of the basis. By minimizing the deviation $\max_{x\in(a,b)} ||f(x) - \sum_{k=0}^{N} a_k P_k(x)|| = min$, a remarkable extension of the classical least squares method, which goes back to Carl Friedrich Gauss, is obtained.

6.1 General Properties of Orthogonal Polynomials

In this section, we give the simplest properties of the orthogonal polynomials with respect to a weight w. For a more in-depth study of the subject one can consult the classic texts.

Denoting by k_n the leading coefficient of the nth polynomial $P_n(x)$ and putting, as usual:

$$\mathbf{h}_k = ||P_k(x)||^2 = \int_a^b P_k^2(x) w(x) dx \tag{6.1}$$

we begin by noting that given the weight w and the interval $[a,b]$ the orthogonal polynomials are each determined up to a multiplicative constant (which can be arranged to make the system orthonormal).

Recurrence relation

Three consecutive orthogonal polynomial system polynomials, associated with weight w on the interval (a,b), are related by the following recurrence relation:

$$P_n(x) = (A_n x + B_n) P_{n-1}(x) - C_n P_{n-2}(x) \quad (n = 2, 3, ...) \tag{6.2}$$

where A_n, B_n, C_n are constants such that $\forall n$, $A_n \neq 0$ and $C_n > 0$.

Furthermore, we have:

$$\begin{cases} A_n = \dfrac{k_n}{k_{n-1}}, \qquad C_n = \dfrac{A_n\ \mathbf{h}_{n-1}}{A_{n-1}\ \mathbf{h}_{n-2}} \\ B_n = A_n \left(\dfrac{k'_n}{k_n} - \dfrac{k'_{n-1}}{k_{n-1}} \right) \end{cases} \tag{6.3}$$

Christoffel-Darboux identity

For orthogonal polynomials, associated with the weight w on $[a, b]$, the following *Christoffel-Darboux identity* holds:

$$\begin{aligned} &\frac{1}{\mathbf{h}_0} P_0(y) P_0(x) + \frac{1}{\mathbf{h}_1} P_1(y) P_1(x) + \cdots + \frac{1}{\mathbf{h}_n} P_n(y) P_n(x) \\ &= \frac{1}{\mathbf{h}_n} \frac{k_n}{k_{n+1}} \frac{P_{n+1}(y) P_n(x) - P_{n+1}(x) P_n(y)}{y - x} \end{aligned} \tag{6.4}$$

Location of zeros

$\forall n$, the zeros $x_1, x_2, \dots, x_n$ of the polynomial P_n, belonging to the set of polynomials orthogonal in $[a, b]$ with respect to the weight w, are all *real*, *distinct* and *internal* to the interval $[a, b]$.

Separation of zeros

Two consecutive orthogonal polynomials $P_n(x)$ and $P_{n+1}(x)$ of the set of orthogonal polynomials in (a, b) with respect to an assigned weight w have no common zeros.

Moreover, there exists the so-called *theorem of separation of zeros.* Denoting by $x_1 < x_2 < \cdots < x_{n+1}$ the zeros of the polynomial $P_{n+1}(x)$, belonging to the set of polynomials orthogonal in (a, b) with respect to the weight w, in each of the open intervals (x_k, x_{k+1}) $(k = 1, 2, ..., n)$ exactly one zero of $P_n(x)$ falls.

6.2 The Classical Orthogonal Polynomials

The orthogonal polynomials that are most frequently encountered in applications are those called *classical orthogonal polynomials*, which are solutions of a differential equation of the hypergeometric type (see [75]), that is, of the type:

$$\sigma(x)y'' + \tau(x)y' + \lambda_n y = 0 \tag{6.5}$$

where $\sigma(x)$ is a polynomial of degree not greater than 2, $\tau(x)$ is a polynomial of degree not greater than 1 and λ_n denotes a constant which is related to the other coefficients by the equation:

$$\lambda_n = -n\tau'(x) - \frac{n(n-1)}{2}\sigma''(x) \tag{6.6}$$

These polynomials, disregarding inessential linear changes in the independent variable, can be reduced to the following:

(I) Jacobi polynomials: $P_n^{(\alpha,\beta)}(x)$ $(\alpha > -1,\ \beta > -1)$ orthogonal in $(-1, 1)$ with respect to the weight:

$$w(x) = (1-x)^\alpha(1+x)^\beta$$

(II) Laguerre polynomials: $L_n^{(\alpha)}(x)$ $(\alpha > -1)$ orthogonal in $(0, +\infty)$ with respect to the weight:

$$w(x) = x^\alpha e^{-x}$$

(III) Hermite polynomials: $H_n(x)$ orthogonal in $(-\infty, +\infty)$ with respect to the weight:

$$w(x) = e^{-x^2}$$

All of the above systems of polynomials constitute *complete systems* in the respective spaces L_w^2. The main reason why the above orthogonal polynomials are frequently used in applications is the possibility to obtain a lot of information from them, as a consequence of the fact that they verify the properties listed below (we remind that these properties characterize the sets of classical orthogonal polynomials, see e.g. [31] pp. 150 and after):

(A) The weight $w(x)$ of the classical orthogonal polynomials satisfies the following Pearson differential equation:

$$\frac{w'(x)}{w(x)} = \frac{D + Ex}{A + Bx + Cx^2} \quad (A, B, C, D, E = \text{constants}) \tag{6.7}$$

(B) Classical orthogonal polynomials satisfy the following generalized Rodrigues' formula:

$$P_n(x) = \frac{1}{K_n w(x)} \frac{d^n}{dx^n} \left(\left[A + Bx + Cx^2\right]^n \; w(x) \right) \tag{6.8}$$

with K_n being a normalization constant which can be chosen arbitrarily and which in the following is chosen in order to respect the traditional standardization. In all cases (I), (II) and (III) it is verified that:

$$\frac{d^k}{dx^k} \left(\left[A + Bx + Cx^2\right]^n \; w(x) \right) \qquad \forall k = 0, 1, \dots, n \tag{6.9}$$

vanishes at the extremes of the interval under consideration. So it is possible to prove the orthogonality of $P_n(x)$ with respect to each power x^k (where $k = 0, 1, \dots, n-1$) representing $P_n(x)$ with the generalized Rodrigues' formula and performing successive integrations by parts. The classical orthogonal polynomials satisfy the hypergeometric differential equation (6.5) which can be rewritten in the form:

$$\begin{aligned} &(A + Bx + Cx^2)\, y'' + [B + D + (2C + E)x]y' \\ &\quad -n[(n+1)C + E]y = 0 \end{aligned} \tag{6.10}$$

In what follows, we limit ourselves to consider only the Chebyshev polynomials.

6.3 Chebyshev Polynomials

Starting from the identity $(e^{it})^n = e^{i\,nt}$, by using Euler's formula:

$$(\cos t + i \sin t)^n = \cos(nt) + i \sin(nt) \tag{6.11}$$

and expanding the first member with Newton's binomial formula, we get:

$$\begin{aligned}
&\sum_{k=0}^{n} i^k \binom{n}{k} \cos^{n-k} t \sin^k t \\
&\quad = \sum_{h=0}^{\left[\frac{n}{2}\right]} (-1)^h \binom{n}{2h} \cos^{n-2h} t \sin^{2h} t \\
&\qquad + i \sum_{h=0}^{\left[\frac{n-1}{2}\right]} (-1)^h \binom{n}{2h+1} \cos^{n-2h-1} t \sin^{2h+1} t \qquad (6.12)\\
&\quad = \sum_{h=0}^{\left[\frac{n}{2}\right]} (-1)^h \binom{n}{2h} \cos^{n-2h} t \left(1-\cos^2 t\right)^h \\
&\qquad + i \sin t \sum_{h=0}^{\left[\frac{n-1}{2}\right]} (-1)^h \binom{n}{2h+1} \cos^{n-2h-1} t \left(1-\cos^2 t\right)^h
\end{aligned}$$

Comparing Equations (6.11) and (6.12) we find:

$$\cos(nt) = \sum_{h=0}^{\left[\frac{n}{2}\right]} (-1)^h \binom{n}{2h} \cos^{n-2h} t \ \left(1-\cos^2 t\right)^h \qquad (6.13)$$

$$\frac{\sin(nt)}{\sin t} = \sum_{h=0}^{\left[\frac{n-1}{2}\right]} (-1)^h \binom{n}{2h+1} \cos^{n-2h-1} t \ \left(1-\cos^2 t\right)^h \qquad (6.14)$$

Putting $x = \cos t$ in (6.13) and (6.14), we obtain two polynomials in x of degrees n and $n-1$ which are called, respectively, Chebyshev polynomials (CP in short) of the first and second kind:

$$T_n(x) := \cos(n \ \arccos x) = \sum_{h=0}^{\left[\frac{n}{2}\right]} (-1)^h \binom{n}{2h} x^{n-2h}(1-x^2)^h$$

$$U_{n-1}(x) := \frac{\sin(n \ \arccos x)}{\sin(\arccos x)} = \sum_{h=0}^{\left[\frac{n-1}{2}\right]} (-1)^h \binom{n}{2h+1} x^{n-2h-1}(1-x^2)^h$$

Such polynomials enjoy many important properties [71, 84] of which we recall the most simple ones.

6.3.1 First Kind and Second Kind Chebyshev Polynomials

Main properties of the first kind CP

The trigonometric identity:

$$\cos((n+1)t) + \cos((n-1)t) = 2\cos t \cos(nt)$$

gives the recursion:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

from which, by using the initial values $T_0(x) = 1$ and $T_1(x) = x$, the subsequent polynomials easily follow:

$$\begin{aligned}
&T_2(x) = 2x^2 - 1 \\
&T_3(x) = 4x^3 - 3x \\
&T_4(x) = 8x^4 - 8x^2 + 1 \\
&T_5(x) = 16x^5 - 20x^3 + 5x \\
&T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1 \\
&T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x \\
&T_8(x) = 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1 \\
&\dots
\end{aligned}$$

Note that:

- The leading coefficient of $T_n(x)$ is 2^{n-1}.
- When $n = 2m$ ($m \in \mathbb{N}$), $T_{2m}(x)$ is an even function of x, while $T_{2m+1}(x)$ is an odd function of x.
- $\forall\ n \in \mathbb{N}, \quad T_n(1) = 1$ and $T_n(-1) = (-1)^n$.

From the equation:

$$\int_0^{\pi} \cos(nt)\cos(mt)\ dt = 0 \qquad (\text{if} \quad m \neq n)$$

by the change of variable $t = \arccos x$, we find the orthogonality property in $[-1, 1]$ with respect to the weight function $(1 - x^2)^{-1/2}$:

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1 - x^2}} dx = 0 \qquad (\text{if} \quad m \neq n)$$

Furthermore, it follows that:

$$\int_{-1}^{1} \frac{1}{\sqrt{1 - x^2}} dx = \pi$$

$$\int_{-1}^{1} \frac{T_n^2(x)}{\sqrt{1 - x^2}} dx = \int_0^{\pi} \cos^2(nt)\ dt = \frac{\pi}{2} \qquad (n \in \mathbf{N})$$

All the n zeros of $T_n(x)$ are real, simple and internal to $[-1, 1]$. More precisely, they are given by:

$$x_k = \cos\left(\frac{(2k+1)}{n}\ \frac{\pi}{2}\right) \qquad (k = 0, 1, \ldots, n-1)$$

In fact, it follows that:

$$|T_n(x_k)| = |\cos(n\ \arccos x_k)| = \left|\cos\left((2k+1)\frac{\pi}{2}\right)\right| = 0$$

Main properties of the second kind CP

Similarly, the properties of the second kind Chebyshev polynomials can be obtained, but we limit ourselves to list them here. They verify the same recursion as the $T_n(x)$:

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x)$$

with the initial conditions $U_0(x) = 1$ and $U_1(x) = 2x$. Therefore, the first few of them are:

$$
\begin{aligned}
&U_2(x) = 4x^2 - 1 \\
&U_3(x) = 8x^3 - 4x \\
&U_4(x) = 16x^4 - 12x^2 + 1 \\
&U_5(x) = 32x^5 - 32x^3 + 6x \\
&U_6(x) = 64x^6 - 80x^4 - 24x^2 - 1 \\
&U_7(x) = 128x^7 - 192x^5 + 80x^3 - 8x \\
&U_8(x) = 256x^8 - 448x^6 + 240x^4 - 40x^2 + 1 \\
&\ldots
\end{aligned}
$$

The second kind Chebyshev polynomials play an important role in representing the powers of a 2×2 non-singular matrix [76, 81]. Extension of this polynomial family to the multivariate case has been considered for representing the powers of an $r \times r$ ($r \geq 3$) non-singular matrix (see [80, 81]).

Remark 1. Chebyshev polynomials are a particular case of the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$, which are orthogonal in the interval $[-1, 1]$ with respect to the weight $(1-x)^\alpha(1+x)^\beta$, since:

$$
T_n(x) = P_n^{(-1/2,-1/2)}(x)\,, \qquad U_n(x) = P_n^{(1/2,1/2)}(x)
$$

Therefore, properties of the Chebyshev polynomials could be deduced in a more general framework of the hypergeometric functions.

6.3.2 Third Kind and Fourth Kind Chebyshev Polynomials

In connection with interpolation and quadrature problems, another couple of Chebyshev polynomials have been considered. They correspond to different choices of weights:

$$
V_n(x) = P_n^{(1/2,-1/2)}(x), \qquad W_n(x) = P_n^{(-1/2,1/2)}(x)
$$

These were called the third and fourth kind Chebyshev polynomials by Walter Gautschi [38].

The third and fourth kind Chebyshev polynomials are defined in $[-1,1]$ as follows:

$$V_n(x) = \frac{\cos[(n+1/2)\arccos x]}{\cos[(\arccos x)/2]}$$

$$W_n(x) = \frac{\sin[(n+1/2)\arccos x]}{\sin[(\arccos x)/2]}$$

Since $W_n(x) = (-1)^n V_n(-x)$, the third kind Chebyshev polynomials transform into those of the fourth kind by interchanging the ends of the interval $[-1,1]$ and so they are not essentially different from each other.

6.4 Non-Trigonometric Fourier Series

Taking up the case of the expansions of a function $f(x)$ in (a,b) by means of the uniformly convergent series:

$$f(x) = f_0 u_0(x) + f_1 u_1(x) + \cdots + f_n u_n(x) + \cdots$$

whose Fourier coefficients are:

$$f_k = (f, u_k)_w = \int_a^b f(x) u_k(x) w(x) dx$$

it is worth to note that the Bessel inequality always holds, that is:

$$f_0^2 + f_1^2 + \cdots + f_n^2 + \cdots = \sum_{k=0}^{\infty} f_k^2 \leq \|f\|_w = \int_a^b f(x) w(x) dx$$

Moreover, if the system of the $u_k(x)$ functions is complete, then the extension to the L_w^2 space of the Pythagorean theorem, which is known as the Parseval equality, holds:

$$\sum_{k=0}^{\infty} f_k^2 = \|f\|_w = \int_a^b f(x) w(x) dx$$

If the system used is not complete, the convergence in quadratic mean does not occur towards the function $f(x)$, but towards the projection of the function $f(x)$ on the linear manifold generated by the linear combinations of the functions of the system being used.