LECTURES ON ORTHOGONAL POLYNOMIALS

JANG SOO KIM

Contents

1. Introduction to the lectures

Orthogonal polynomials are classical objects arising from the study of continued fractions. Due to the long history of orthogonal polynomials, they have now become important objects of study in many areas: classical analysis and PDE, mathematical physics, probability, random matrix theory, and combinatorics.

The combinatorial study of orthogonal polynomials was pioneered by Flajolet and Viennot in 1980s. In these lectures we will learn fascinating combinatorial properties of orthogonal polynomials.

We will first study basic properties of orthogonal polynomials based on Chihara's book, Chapter 1 [1]. We will then focus on the combinatorial approach of orthogonal polynomials, which will be based on Viennot's lecture notes [2]. We will also cover more recent developments in the combinatorics of orthogonal polynomials such as their connections with ASEP, staircase tableaux, lecture hall partitions, and orthogonal polynomials of type R_1.

The prerequisites of this course are Calculus 1, Linear Algebra, and Discrete Mathematics.

2. Elementary Theory of Orthogonal Polynomials

In this section we will cover the first chapter of Chihara's book [1].

2.1. **Introduction.** Since

$$2\cos m\theta \cos n\theta = \cos(m+n)\theta + \cos(m-n)\theta,$$

for nonnegative integers m and n, we have

$$(2.1) \qquad \int_0^\pi \cos m\theta \cos n\theta d\theta = 0, \qquad m \neq n.$$

In this situation we say that $\cos m\theta$ and $\cos n\theta$ are orthogonal over the interval $(0, \pi)$.

Note that $\cos n\theta$ is a polynomial in $\cos \theta$ of degree n. So we can write $\cos n\theta = T_n(\cos \theta)$ for a polynomial $T_n(x)$ of degree x.

By the change of variable $x = \cos \theta$, (2.2) can be rewritten as

$$(2.2) \qquad \int_{-1}^1 T_m(x)T_n(x)(1-x^2)^{-1/2}dx = 0, \qquad m \neq n.$$

The polynomials $T_n(x)$, $n \geq 0$, are called the *Tchebyshev polynomials of the first kind*. The first few polynomials are:

$$T_0(x) = 1,$$
$$T_1(x) = \cos\theta = x,$$
$$T_2(x) = \cos 2\theta = 2\cos^2\theta - 1 = 2x^2 - 1,$$
$$T_3(x) = 4x^3 - 3x.$$

Recall that in an inner product space V with inner product $\langle\cdot,\cdot\rangle$, a set of vectors $v_1,\ldots,v_n$ are said to be orthogonal if $\langle v_i, v_j\rangle = 0$ for all $i \neq j$. In this sense the Tchebyshev polynomials $T_n(x)$ are orthogonal, where $V = \mathbb{R}[x]$ is the space of polynomials with real coefficients with the inner product given by

$$\langle f(x), g(x)\rangle = \int_{-1}^{1} f(x)g(x)(1-x^2)^{-1/2}dx.$$

We say that $T_n(x)$ are *orthogonal polynomials* with respect to the *weight function* $(1-x^2)^{-1/2}$ on the interval $(-1,1)$.

Definition 2.1. Suppose that $w(x)$ is a nonnegative and integrable function on (a,b) with $\int_a^b w(x)dx > 0$ and $\int_a^b x^n dx < \infty$ for all $n \geq 0$. A sequence of polynomials $\{P_n(x)\}_{n\geq 0}$ is called an *orthogonal polynomial sequence (OPS)* with respect to the *weight function* $w(x)$ on (a,b) if the following conditions hold:

(1) $\deg P_n(x) = n$, $n \geq 0$,
(2) $\int_a^b P_m(x)P_n(x)w(x)dx = 0$ for $m \neq n$.

There is another way to define orthogonal polynomials without using the weight function. For a polynomial $f(x)$, if we define

$$\mathcal{L}(f(x)) = \int_a^b f(x)w(x)dx,$$

then $\mathcal{L}(f(x))$ is completely determined by the *moments* $\mu_n = \int_a^b x^n w(x)dx$. So, if we are only interested in polynomials, then we can define a linear functional $\mathcal{L}$ using a moment sequence $\mu_0, \mu_1, \ldots$. Not every sequence $\mu_0, \mu_1, \ldots$ gives rise to an OPS, though. We will see later a criterion for a sequence to be a moment sequence.

Definition 2.2. Let $\mathcal{L}$ be the linear functional defined on the space of polynomials in x. A sequence of polynomials $\{P_n(x)\}_{n\geq 0}$ is called an *orthogonal polynomial sequence (OPS)* with respect to $\mathcal{L}$ if the following conditions hold:

(1) $\deg P_n(x) = n$, $n \geq 0$,
(2) $\mathcal{L}(P_m(x)^2) \neq 0$ for $m \geq 0$,
(3) $\mathcal{L}(P_m(x)P_n(x)) = 0$ for $m \neq n$.

Note that the second condition above was not necessary in Definition 2.1 because it follows from the facts that $w(x)$ is nonnegative and $\int_a^b w(x)dx > 0$.

Remark 2.3. The moments of the Tchebyshev polynomials are

$$\mu_{2n} = \int_{-1}^{1} x^{2n}(1-x^2)^{-1/2}dx = \frac{\pi}{2^{2n}}\binom{2n}{n}, \qquad \mu_{2n+1} = 0.$$

This suggests that there could be some interesting combinatorics behind the scene. We will later find a combinatorial way to understand this situation.

Example 2.4 (Charlier polynomials)**.** The *Charlier polynomials* $P_n(x)$ are defined by

$$P_n(x) = \sum_{k=0}^{n} \binom{x}{k}\frac{(-a)^{n-k}}{(n-k)!},$$

where $\binom{x}{k} = x(x-1)\cdots(x-k+1)/k!$. We will find a different type of orthogonality for $P_n(x)$.

The generating function for $P_n(x)$ is

$$G(x,w) = \sum_{n\geq 0} P_n(x)w^n = \sum_{n\geq 0}\left(\sum_{k=0}^{n}\binom{x}{k}\frac{(-a)^{n-k}}{(n-k)!}\right)w^n = \sum_{n\geq 0}\binom{x}{n}w^n\sum_{n\geq 0}\frac{(-a)^m}{m!}w^m,$$

which means

$$G(x,w) = e^{-aw}(1+w)^x.$$

Thus

$$a^x G(x,v)G(x,w) = e^{-a(v+w)}\left(a(1+v)(1+w)\right)^x.$$

We have

$$\sum_{k\geq 0}\frac{a^k G(k,v)G(k,w)}{k!} = \sum_{k\geq 0}\frac{e^{-a(v+w)}\left(a(1+v)(1+w)\right)^k}{k!} = e^{-a(v+w)}e^{a(1+v)(1+w)} = e^a e^{avw}.$$

Thus

(2.3)
$$\sum_{k\geq 0}\frac{a^k G(k,v)G(k,w)}{k!} = \sum_{n\geq 0}\frac{e^a(avw)^n}{n!}.$$

On the other hand

$$\sum_{k\geq 0}\frac{a^k G(k,v)G(k,w)}{k!} = \sum_{k\geq 0}\frac{a^k}{k!}\sum_{m,n\geq 0}P_m(k)P_n(k)v^m w^n$$

(2.4)
$$= \sum_{m,n\geq 0}\left(\sum_{k\geq 0}P_m(k)P_n(k)\frac{a^k}{k!}\right)v^m w^n.$$

Comparing the coefficients of $v^m w^n$ in (2.3) and (2.4) we obtain

(2.5)
$$\sum_{k\geq 0}P_m(k)P_n(k)\frac{a^k}{k!} = \frac{e^a a^n}{n!}\delta_{n,m}.$$

Therefore, if we define a linear functional $\mathcal{L}$ by

$$\mathcal{L}(x^n) = \sum_{k\geq 0}k^n\frac{a^k}{k!},$$

then $P_n(x)$ are orthogonal polynomials with respect to $\mathcal{L}$.

Note that we describe the orthogonality of $P_n(x)$ using only the linear functional $\mathcal{L}$ without referring to any weight function. However, we can also find a weight function in this case. Let $\psi(x)$ be the step function with a jump at $k = 0, 1, 2, \ldots$ of magnitude $a^k/k!$. Then the linear functional $\mathcal{L}$ can be written as the following Riemann–Stieltjes integral
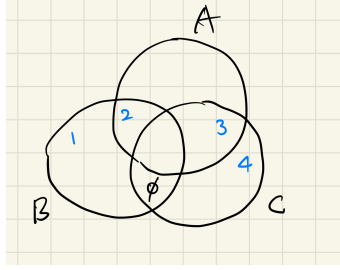
$$\mathcal{L}(f(x)) = \int_{-\infty}^{\infty}f(x)d\psi(x).$$

Remark 2.5. We can also prove (2.5) in a combinatorial way. To do this, note that

$$\sum_{k\geq 0}P_m(k)P_n(k)\frac{a^k}{k!} = \sum_{k\geq 0}\sum_{i=0}^{m}\binom{k}{i}\frac{(-a)^{m-i}}{(m-i)!}\sum_{j=0}^{n}\binom{k}{j}\frac{(-a)^{n-j}}{(n-j)!}\frac{a^k}{k!}$$

$$= \sum_{k\geq 0}\sum_{i=0}^{m}\sum_{j=0}^{n}\binom{k}{m-i}\frac{(-a)^i}{i!}\binom{k}{n-j}\frac{(-a)^j}{j!}\frac{a^k}{k!}$$

$$= \sum_{N\geq 0}\frac{a^N}{N!}\sum_{i+j+k=N}(-1)^{i+j}\frac{N!}{i!j!k!}\binom{k}{m-i}\binom{k}{n-j}.$$

For a fixed N,

$$\sum_{i+j+k=N}(-1)^{i+j}\binom{N}{i,j,k}\binom{k}{m-i}\binom{k}{n-j} = \sum_{(A,B,C)\in X}(-1)^{|B\setminus A|+|C\setminus A|},$$

FIGURE 1. The triple (A, B, C).

where X is the set of triples (A, B, C) such that $A \cup B \cup C = \{1, \ldots, N\}$, $|A| = k$, $|B| = m$, $|C| = n$, $(B \cap C) \setminus A = \emptyset$. If $B \neq C$, we can find a sign-reversing involution, namely toggling the smallest integer in regions 1 and 2 or in regions 3 and 4 in Figure 1. Thus we have

$$\sum_{(A,B,C) \in X} (-1)^{|B \setminus A| + |C \setminus A|} = \sum_{(A,B,C) \in X : B = C} (-1)^{|B \setminus A| + |C \setminus A|} = \delta_{m,n} \binom{N}{n}.$$

This implies

$$\sum_{k \geq 0} P_m(k) P_n(k) \frac{a^k}{k!} = \delta_{m,n} \sum_{N \geq 0} \frac{a^N}{N!} \binom{N}{n} = \frac{e^a a^n}{n!} \delta_{n,m}.$$

Remark 2.6. In the theory of orthogonal polynomials, finding an explicit weight function is an important problem. However, in these lectures, we will not pursue in this direction and we will be mostly satisfied with Definition 2.2.

2.2. **The moment functional and orthogonality.** We will consider the space $\mathbb{C}[x]$ polynomials with complex coefficients. A *linear function* on $\mathbb{C}[x]$ is a map $\mathcal{L} : \mathbb{C}[x] \to \mathbb{C}$ such that $\mathcal{L}(af(x) + bg(x)) = a\mathcal{L}(f(x)) + b\mathcal{L}(g(x))$ for all $f(x), g(x) \in \mathbb{C}[x]$ and $a, b \in \mathbb{C}$.

Definition 2.7. Let $\{\mu_n\}_{n \geq 0}$ be a sequence of complex numbers. Let $\mathcal{L}$ be the linear functional on the space of polynomials defined by $\mathcal{L}(x^n) = \mu_n$, $n \geq 0$. In this case we say that $\mathcal{L}$ is the *moment functional* determined by the *moment sequence* $\{\mu_n\}$, and μ_n is called the *nth moment*.

We recall the definition of orthogonal polynomials.

Definition 2.8. Let $\mathcal{L}$ be the linear functional defined on the space of polynomials in x. A sequence of polynomials $\{P_n(x)\}_{n \geq 0}$ is called an *orthogonal polynomial sequence (OPS)* with respect to $\mathcal{L}$ if the following conditions hold:

(1) $\deg P_n(x) = n$, $n \geq 0$,
(2) $\mathcal{L}(P_m(x) P_n(x)) = K_n \delta_{m,n}$, for some $K_n \neq 0$.

We say that $P_n(x)$ are *orthonormal* if $\mathcal{L}(P_m(x) P_n(x)) = \delta_{m,n}$.

Theorem 2.9. *Let $\{P_n(x)\}$ be a sequence of polynomials and let $\mathcal{L}$ be a moment sequence. The following are equivalent:*

(1) *$\{P_n(x)\}$ is an OPS with respect to $\mathcal{L}$;*
(2) *$\mathcal{L}(\pi(x) P_n(x)) = 0$ if $\deg \pi(x) < n$ and $\mathcal{L}(\pi(x) P_n(x)) \neq 0$ if $\deg \pi(x) = n$;*
(3) *$\mathcal{L}(x^m P_n(x)) = K_n \delta_{m,n}$, $0 \leq m \leq n$, for some $K_n \neq 0$.*

Proof. (1) $\Rightarrow$ (2): Suppose that $\deg \pi(x) \leq n$. Since $\{P_n(x)\}$ is a basis of $\mathbb{C}[x]$, we can write

$$\pi(x) = c_0 + c_1 P_1(x) + \cdots + c_n P_n(x).$$

Then

$$\mathcal{L}(\pi(x) P_n(x)) = \sum_{k=0}^{n} \mathcal{L}(c_k P_k(x) P_n(x)) = c_n \mathcal{L}(P_n(x)^2),$$

which is zero if $\deg \pi(x) < n$ and nonzero if $\deg \pi(x) = n$.

(2) $\Rightarrow$ (3): Trivial. (2) $\Rightarrow$ (3): Trivial. $\square$

Theorem 2.10. *Suppose that $\{P_n(x)\}_{n \geq 0}$ be an OPS with respect to $\mathcal{L}$. Then for any polynomial $\pi(x)$ of degree n,*

$$\pi(x) = \sum_{k=0}^{n} c_k P_k(x), \qquad c_k = \frac{\mathcal{L}(\pi(x)P_k(x))}{\mathcal{L}(P_k(x)^2)}.$$

Proof. Clearly, we can write

$$\pi(x) = \sum_{k=0}^{n} c_k P_k(x),$$

for some c_k. Multiplying $P_j(x)$ both sides and taking $\mathcal{L}$, we get

$$\mathcal{L}(\pi(x)P_j(x)) = \sum_{k=0}^{n} \mathcal{L}\left(c_k P_k(x)P_j(x)\right) = c_j \mathcal{L}(P_j(x)^2).$$

Dividing both sides by $\mathcal{L}(P_j(x)^2)$, we obtain the theorem. $\square$

Theorem 2.11. *Suppose that $\{P_n(x)\}_{n \geq 0}$ be an OPS with respect to $\mathcal{L}$. Then $P_n(x)$ is uniquely determined by $\mathcal{L}$ up to a nonzero factor. More precisely, if $\{Q_n(x)\}_{n \geq 0}$ is an OPS with respect to $\mathcal{L}$, then there are constants $c_n \neq 0$ such that $Q_n(x) = c_n P_n(x)$ for all $n \geq 0$.*

Proof. Let us write $Q_n(x) = \sum_{k=0}^{n} c_k P_k(x)$. Then by Theorem 2.10, $c_k = \mathcal{L}(Q_n(x)P_k(x))/\mathcal{L}(P_k(x)^2)$. But by Theorem 2.9, $\mathcal{L}(Q_n(x)P_k(x)) = 0$ unless $k = n$. Thus $Q_n(x) = c_n P_n(x)$. $\square$

Note that if $\{P_n(x)\}_{n \geq 0}$ is an OPS for $\mathcal{L}$, then so is $\{c_n P_n(x)\}_{n \geq 0}$ for any $c_n \neq 0$. Therefore there is a unique monic OPS, which is obtained by dividing each $P_n(x)$ by its leading coefficient. Note also that there is a unique orthonormal OPS as well given by $p_n(x) = P_n(x)/\mathcal{L}(P_n(x)^2)$. In summary we have the following corollary.

Corollary 2.12. *Suppose that $\mathcal{L}$ is a moment sequence such that there is an OPS for $\mathcal{L}$. Let K_n, $n \geq 0$, be a sequence of nonzero numbers.*

(1) *There is a unique monic OPS $\{P_n(x)\}_{n \geq 0}$ for $\mathcal{L}$.*
(2) *There is a unique orthonormal OPS $\{P_n(x)\}_{n \geq 0}$ for $\mathcal{L}$.*
(3) *There is a unique OPS $\{P_n(x)\}_{n \geq 0}$ for $\mathcal{L}$ such that the leading coefficient of $P_n(x)$ is K_n.*
(4) *There is a unique OPS $\{P_n(x)\}_{n \geq 0}$ for $\mathcal{L}$ such that $\mathcal{L}(x_n P_n(x)) = K_n$.*

Clearly, if $\{P_n(x)\}_{n \geq 0}$ is an OPS for $\mathcal{L}$, then it is also an OPS for $\mathcal{L}'$ given by $\mathcal{L}'(f(x)) = c\mathcal{L}(f(x))$ for some $c \neq 0$. Therefore, by dividing the linear functional by the value $\mathcal{L}(1)$, we may assume that $\mathcal{L}(1) = 1$.

REFERENCES

[1] T. S. Chihara. *An introduction to orthogonal polynomials.* Gordon and Breach Science Publishers, New York, 1978. Mathematics and its Applications, Vol. 13.
[2] G. Viennot. Une théorie combinatoire des polynômes orthogonaux généraux. Lecture Notes, UQAM, 1983.