Chapter 11

Orthogonal Polynomial Ensembles

11.1 Orthogonal Polynomials of Scalar Argument

Let $w(x)$ be a weight function on a real interval, or the unit circle, or generally on some curve in the complex plane. Without being too technical we think of $w(x)$ as piecewise continuous and we allow for delta functions. We will let I denote the support of w. We require that $w(x)$ have finite mass, i.e., $\int_I w(x)\mathrm{d}x$ is finite.

Every famous mathematician seems to have a set of orthogonal polynomials:

Example 1: *Legendre Polynomials*: Let $w(x) = 1$ on $[-1, 1]$. The polynomials are named for Legendre and denoted $P_n(x)$.

Example 2: *Chebyshev Polynomials*: $w(x) = (1-x^2)^{-1/2}$. The polynomials are denoted by $T_n(x)$ and satisfy $T_n(\cos\theta) = \cos(n\theta)$. For example $\cos 3\theta = 4\cos^3(\theta) - 3$ and $T_3(x) = 4x^3 - 3$.

Example 3: *Hermite Polynomials*: $w(x) = e^{-x^2/2}$ on the whole real line. The $H_n(x)$ have a special place in random eigenvalue history because of the Gaussian Ensembles.

Example 4: *Laguerre Polynomials*: $w(x) = x^\gamma e^{-x/2}$. For every parameter γ there is a set of Laguerre polynomials denoted $L_n^\gamma(x)$. These polynomials play a vital role in the Wishart matrices of multivariate statistics.

Example 5: *Jacobi Polynomials*: $w(x) = (1-x)^{\gamma_1}(1+x)^{\gamma_2}$. This is the granddaddy of all the above. When $\gamma_1 = \gamma_2 = 0$, we obtain the Legendre Polynomials. If $\gamma_1 = \gamma_2 = -1/2$, we obtain Chebyshev. Taking limits appropriately we can obtain the Hermite and Laguerre Polynomials.

There is a beautiful relationship among all of the following quantities

Quantity 1: *The moments* $s_k = \int_I x^k w(x)\mathrm{d}x$

Quantity 2: *The sequence of Orthogonal Polynomials* $p_0(x), p_1(x), \ldots$, where $\int_I p_j(x)p_k(x)w(x)\mathrm{d}x = \delta_{jk}$. It is assumed that $p_j(x)$ has degree j.

Quantity 3: *Infinite Symmetric Tridiagonal Matrices* (also known as Jacobi matrices). The three term recurrence for the orthogonal polynomials is often best seen as a tridiagonal matrix. The characteristic polynomial of the top j by j section is $p_j(x)$. If one has a finite symmetric tridiagonal matrix then one obtains a finite sequence of orthogonal polynomials corresponding to a finite measure.

Quantity 4: *The eigenvalues and the first row of the eigenvectors* of the above tridiagonal matrix as captured in $w_i(x) = \sum q_i^2 \delta(x - \lambda_i)$.

Quantity 5: *Gaussian Quadrature*: The abscissas are the eigenvalues of T_n, and the weights are the q_i^2.

Quantity 6: *Continued Fractions*: If we evaluate the continued fraction given by the diagonals and the offdiagonals squared, we obtain a sequence of numerators and denominators. The numerators are the orthogonal polynomials and the denominators are closely related.

Quantity 7: *Random Variables*: If we normalize $w(x)$ to have integral 1, then we have random variables with that density. For example, the Hermite density corresponds to the normal distribution, the Laguerre density corresponds to χ distributions, and the Jacobi density corresponds to the F (check this!!!!!!!!!!!!!!!) distribution.

Quantity 8: *Jump Condition*: One can seek function $Y(x)$ in the complex plane that are defined off of the support of $w(x)$ and satisfy a jump condition:

$$\lim_{\epsilon \to 0} \{Y(x + i\epsilon) - Y(x - i\epsilon)\} = \pi_n(x) w(x),$$

where all we specify about $\pi_n(x)$ is that it is a polynomial of degree n. We also require the growth condition $Y(x) = O(x^{-n})$ as $x \to \infty$. The solution is unique. π_n turns out to be the nth orthogonal polynomial, and $Y(x)$ is its moment generating function. For more on this see the section on Riemann Hilbert problem.

Quantity 9: *The Lanczos Tridiagonalization Algorithm*

11.2 Orthogonal Polynomial Ensembles and Matrix Models

Let $\omega(x)$ be a weight function. For parameters β we define the multivariate weight function: $\omega_n(x; \beta) = \Delta^\beta \prod_{i=1}^n (\omega(x_i))$, where

$$\Delta = \left| \prod (x_i - x_j) \right| .$$

Thus Δ is the absolute value of the Vandermonde determinant, i.e., the determinant of the matrix $(x_i^{j-1})_{i,j=1,\dots,n}$.

When β is understood from context, we may suppress it in the notation.

Definition: Random Matrix Models We say that we have a random matrix model for the multivariate weight function $\omega_n(x; \beta)$, if we have a random $n \times n$ matrix A_n whose eigenvalues have $\omega_n(x\beta)$ as a joint density.

Constructing nice random matrix models is a central problem in random matrix theory. It would be cheating to say, let x be a random variable with multivariate density $\omega_n(\beta)$ and take A_n to be the diagonal matrix with x on the diagonal.

11.2.1 The Cauchy-Binet Theorem and its Continuous Limit

Let $C = AB$ be a matrix product of any kind. Let $M\binom{i_1 \dots i_m}{j_1 \dots j_m}$ denote the $m \times m$ minor

$$\det(M_{i_k j_p})_{1 \le k \le m, 1 \le l \le m}.$$

In other words, it is the determinant of the submatrix of M formed from rows of $i_1, \ldots, i_m$ and columns $j_1, \ldots, j_m$.

The Cauchy-Binet Theorem states that

$$C\begin{pmatrix} i_1, \ldots, i_m \\ k_1, \ldots, k_m \end{pmatrix} = \sum_{j_1 < j_2 < \cdots < j_m} A\begin{pmatrix} i_1, \ldots, i_m \\ j_1, \ldots, j_m \end{pmatrix} B\begin{pmatrix} j_1, \ldots, j_m \\ k_1, \ldots, k_m \end{pmatrix}.$$

Notice that when $m = 1$ this is the familiar formula for matrix multiplication. When all matrices are $m \times m$, then the formula states that

$$\det C = \det A \det B\,.$$

Cauchy-Binet extends to matrices with infinitely many columns. If the columns are indexed by a continuous variable, we now have a vector of functions.

Replacing A_{ij} with $\varphi_i(x_j)$, B_{jk} with $\psi_k(x_j)$, we see that Cauchy-Binet becomes

$$\det C = \int \cdots \int \det(\varphi_i(x_j))_{i,j=1,\ldots,n} \det(\psi_k(x_j))_{k,j=1,\ldots,n} \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n\,.$$

where $C_{ik} = \int \varphi_i(x)\psi_k(x)\mathrm{d}x,\ i, k = 1, \ldots, n$.

Apparently, this continuous version of Cauchy-Binet may be traced back to an 1883 paper by Andréief [9].

11.2.2 Cauchy-Binet Examples

Corollary 11.1. *Let $A \in \mathbb{R}^{n,p}, D \in \mathbb{R}^{n,n}$. We have*

$$\begin{aligned} \det(A^T D A) &= \sum_{i_1 < \ldots < i_p} A\begin{pmatrix} i_1 & \ldots & i_p \\ 1 & \ldots & p \end{pmatrix}^2 \mathrm{d}_{i_1} \ldots \mathrm{d}_{i_p} \\ &= \mathrm{Pl}(A)^T \begin{pmatrix} \ddots & & \\ & \mathrm{d}_{i_1} \ldots \mathrm{d}_{i_p} & \\ & & \ddots \end{pmatrix}_{i_1 < \ldots < i_p} \mathrm{Pl}(A) \end{aligned}$$

Example 11.1. *Let* $\mathrm{d}_i = \begin{cases} 1 & \text{if } i \le l \\ 0 & \text{if } i > l \end{cases}$

We have $A_l = A(1:l,:)$ and then

$$\det(A_l^T A_l) = \sum_{\text{All } i \le l} A\begin{pmatrix} i_1 & \ldots & i_p \\ 1 & \ldots & p \end{pmatrix}^2$$

Example 11.2. *Let* $\mathrm{d}_i = \begin{cases} 1+z & \text{if } i = 5 \\ 1 & \text{if } i \ne 5 \end{cases}$

Also let $A^T A = I_p$.

$$\begin{aligned} \det(A^T \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1+z & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} A) &= 1 + z \sum_{\text{Some } i=5} A\begin{pmatrix} i_1 & \ldots & i_p \\ i & \ldots & p \end{pmatrix}^2 \\ &= 1 + z\|A(S,:)\|^2 \end{aligned}$$

Example 11.3. *Let* $\mathrm{d}_i = Z_i$ *symbolic*

$$\det(A^T DA) = \sum A \begin{pmatrix} i_1 & \ldots & i_p \\ 1 & \ldots & p \end{pmatrix}^2 Z_{i_1} \ldots Z_{i_p}$$

Displays squares of elements of $\mathrm{Pl}(A)$ *with symbolic labels.*

Example 11.4. *Let* $\mathrm{d}_i = 1 + Z_i$

$$\det(A^T DA) = \sum A \begin{pmatrix} i_1 & \ldots & i_p \\ 1 & \ldots & p \end{pmatrix}^2 + \sum_{\substack{\text{Some} \\ Z_k = i}} Z_i A \begin{pmatrix} i_1 & \ldots & i_p \\ 1 & \ldots & p \end{pmatrix}^2$$

$$+ \sum_{\substack{\text{Some} \\ Z_k = i \\ i_k = j}} Z_i Z_j A \begin{pmatrix} i_1 & \ldots & i_p \\ 1 & \ldots & p \end{pmatrix}^2 + \cdots$$

11.3 Orthogonal Polynomial Ensembles when $\beta = 2$

In this section we assume that $\beta = 2$ so that $\omega_n(x) = \Delta^2 \prod_{i=1}^n \omega(x_i)$. For classical weight function $\omega(x)$, Hermitian matrix models have been constructed. We have already seen the GUE corresponding to Hermite matrix models, and complex Wishart matrices for Laguerre. We also get the complex MANOVA matrices corresponding to Jacobi.

Notation: We define $\phi_n(x) = p_n(x)\omega(x)^{1/2}$. Thus the $\phi_i(x)$ are not generally polynomials, but they do form an orthonormal set of functions on the support of ω.

It is a general fact that the level density is

$$f\omega(x) = \sum_{i=1}^{n-1} \phi_i(x)^2 \,.$$

Given any function $f(x)$ one can ask for

$$E(f) \equiv E_{\omega_n}(\prod (f(x_i)) \,.$$

When we have a matrix model, this is $E(\det(f(X))$.

It is a simple result that $E(f) = \int (\det(\phi_i(x)\phi_j(x)f(x))_{i,j=0,\ldots,n-1}\mathrm{d}x$. This implies by the continuous version of the Cauchy Binet theorem that

$$E(f) = \det C_n,$$

where $(C_n)_{ij} = \int \phi_i(x)\phi_j(x)f(x)\mathrm{d}x$.

Some important functions to use are $f(x) = 1 + \sum z_i(\delta(x - y_i))$. The coefficients of the resulting polynomial then is the marginal density of k eigenvalues. [say slightly better]

Another important one is $f(x) = 1 - \chi_{[a,b]}$, where $\chi_{[a,b]}$ the indicator function on $[a, b]$. Then we obtain the probability that no eigenvalue is in the interval $[a, b]$. If b is infinite, we obtain the probability that all eigenvalues are less than a, that is the distribution function for the largest eigenvalue.

11.4 Orthogonal Polynomials of Matrix Argument

Let $x = (x_1, \dots, x_n)$, for any β we can define the multivariate orthogonal polynomial $p_\kappa(x)$ by the condition that

$$\int p_\kappa(x) p\nu(x) \mathrm{d}\omega_n(x;\beta) = \delta_\kappa \nu \,.$$

Here κ is a finite non-increasing, sequence and the leading term in $p_\kappa(x) = x^\kappa$.

These orthogonal polynomials are symmetric. Indeed the set of polynomials of degree $\leq k$ form a basis for the symmetric polynomials of degree $\leq k$.

Conjecture: Let $\omega(x) = 1$ on the unit circle in the complex plane. Then the orthogonal polynomials are the Jack polynomials. These polynomials are homogeneous.

Given any $n \times n$ matrix, we can define $p_\kappa(X)$ as $p_\kappa(\lambda_1, \dots, \lambda_n)$. Since p_κ is symmetric, we see that $p_\kappa(X)$ is a polynomial in the elements of X. (For example the determinant, the trace, or any coefficient in the characteristic polynomial can be expressed as a polynomial in the entries of the matrix.)

When $\beta = 1, 2$, and 4 we can define a measure $\omega(X)$ on real symmetric, complex Hermitian, or symplectic self-dual matrix such that $\int p_\kappa(X) p_\nu(X) \omega(X) dX = \delta_{\kappa\nu}$. Here $\omega(X) = \prod \omega(\lambda_i(X)) = \det \omega(X)$.

Alternatively there is a measure on real tridiagonal matrices for any such β. Similarly we can construct a random tridiagonal matrix whose eigenvalue probability density is $\omega_n(x;\beta)$.