Introduction to orthogonal polynomials

Michael Anshelevich

November 6, 2003

$\mu =$ probability measure on $\mathbb{R}$ with finite moments

$$m_n(\mu) = \int_{\mathbb{R}} x^n \, d\mu(x) < \infty.$$

Induces a functional on polynomials $\mathbb{C}[x]$,

$$\varphi\left[P(x)\right] = \int_{\mathbb{R}} P(x) \, d\mu(x).$$

On the polynomials $\mathbb{C}[x]$, define the sesquilinear inner product

$$\left\langle x^n, x^k \right\rangle_\mu = \varphi_\mu\left[x^{n+k}\right] = m_{n+k}(\mu).$$

The set $\{x^n\}_{n=0}^{\infty}$ is a basis for $\mathbb{C}[x]$. Gram-Schmidt with respect to the inner product $\langle \cdot, \cdot \rangle_\mu$, get a family of polynomials

$$\{P_n\}_{n=0}^{\infty}.$$

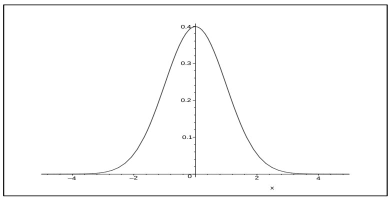
Note they have **real coefficients**. They are orthogonal with respect to μ:

$$\varphi_\mu\left[P_n P_k\right] = \int_{\mathbb{R}} P_n(x) P_k(x) \, d\mu(x) = 0$$

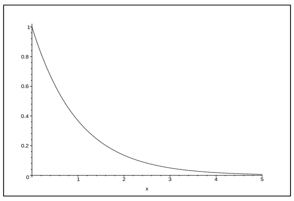
if $n \neq k$.

Hermite polynomials: normal (Gaussian) distribution

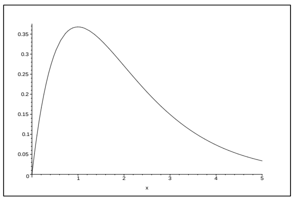
$$\frac{1}{\sqrt{2\pi t}}\, e^{-x^2/2t}\, dx.$$



Laguerre polynomials: Gamma distribution



$$\frac{1}{\Gamma(t)}\, e^{-x} x^{t-1}\, \mathbf{1}_{[0,\infty)}(x)\, dx.$$



Jacobi polynomials:

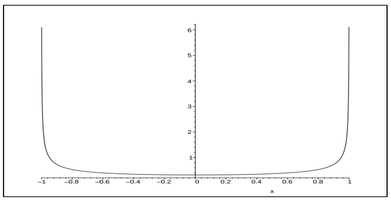
$$\frac{1}{Z(\alpha,\beta)}\,(1-x)^{\alpha}(1+x)^{\beta}\,\mathbf{1}_{[-1,1]}(x)\,dx.$$

In particular: **Ultraspherical (Gegenbauer)**,
$\alpha = \beta = \lambda - \frac{1}{2}$,

$$\frac{1}{Z(\lambda)}\,(1-x^2)^{\lambda-\frac{1}{2}}\,\mathbf{1}_{[-1,1]}(x)\,dx,$$

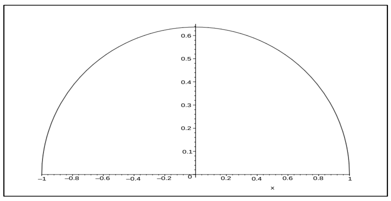
Chebyshev of the 1st kind, $\alpha = \beta = -\frac{1}{2}$, $\lambda = 0$,

$$\frac{1}{\pi} \frac{1}{\sqrt{1-x^2}} \mathbf{1}_{[-1,1]}(x)\, dx,$$



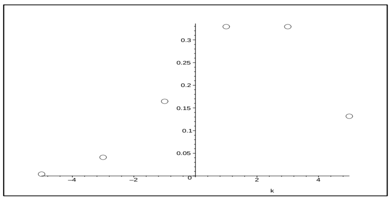
Chebyshev of the 2nd kind, $\alpha = \beta = \frac{1}{2}$, $\lambda = 1$,

$$\frac{2}{\pi} \sqrt{1-x^2}\, \mathbf{1}_{[-1,1]}(x)\, dx,$$



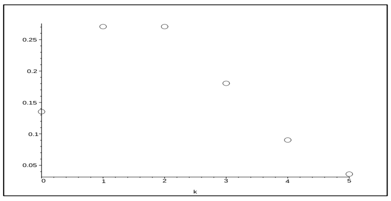
Kravchuk: binomial distribution, for $q = 1 - p$,

$$\sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} \delta_{2k-N}(x).$$



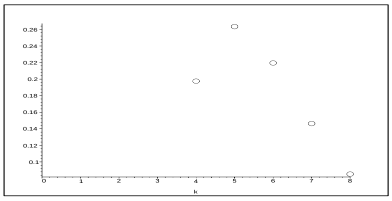
Charlier: Poisson distribution

$$\sum_{k=0}^{\infty} e^{-t} \frac{t^k}{k!} \delta_k(x).$$



Meixner: negative binomial (Pascal) distribution

$$\sum_{k \geq t} \binom{k-1}{t-1} p^t q^{k-t} \delta_k(x).$$



Can describe in various ways:

- Using the measure of orthogonality.

- 2nd order differential or difference equations.

- Explicit formulas using hypergeometric functions.

- Three-term recursion relations.

Favard's theorem. (Stone 1932) Let $\{P_n\}_{n=0}^{\infty}$ be a polynomial family, that is, P_n has degree n.

$\{P_n\}$ are **orthogonal** with respect to some measure μ with positive leading coefficients $\Longleftrightarrow$ they satisfy a **three-term recursion relation**

$$\begin{aligned}
P_0(x) &= 1, \\
xP_0(x) &= \alpha_0 P_1(x) + \beta_0 P_0(x), \\
xP_n(x) &= \alpha_n P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x),
\end{aligned}$$

where all β_n are real, $\alpha_n > 0$, $\gamma_n \geq 0$.

Moreover, $\{P_n\}$ are monic iff all $\alpha_n = 1$.
$\{P_n\}$ are ortho**normal** iff $\gamma_n = \alpha_{n-1}$.
μ has finite support iff $\gamma_n = 0$ for some n.

$\Rightarrow$) Suppose the polynomials are orthogonal with respect to μ and have positive leading coefficients. Then

$$P_n \perp \text{span}\left(P_0, P_1, \ldots, P_{n-1}\right).$$

So

$$\langle xP_n, P_k\rangle_\mu = 0$$

for $n + 1 < k$, and also

$$\langle xP_n, P_k\rangle_\mu = \varphi_\mu\left[xP_nP_k\right] = \langle P_n, xP_k\rangle = 0$$

for $n > k + 1$, $n - 1 > k$. So

$$xP_n(x) = \alpha_n P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x)$$

for some real $\alpha_n, \beta_n, \gamma_n$.

Leading coefficients of P_n, P_{n+1} are positive $\Rightarrow \alpha_n > 0$.
Also, (leading coefficients are 1) $\Leftrightarrow \alpha_n = 1$.

From the recursion relation,

$$\langle xP_n, P_{n-1}\rangle = \gamma_n \langle P_{n-1}, P_{n-1}\rangle .$$

Also, from the recursion

$$\begin{aligned} & xP_{n-1}(x) \\ & \quad = \alpha_{n-1}P_n(x) + \beta_{n-1}P_{n-1}(x) + \gamma_{n-1}P_{n-2}(x), \end{aligned}$$

it follows that

$$\langle xP_n, P_{n-1}\rangle = \langle P_n, xP_{n-1}\rangle = \alpha_{n-1} \langle P_n, P_n\rangle .$$

So

$$\gamma_n \langle P_{n-1}, P_{n-1}\rangle = \alpha_{n-1} \langle P_n, P_n\rangle .$$

So $\gamma_n \geq 0$, and $\{P_n\}$ are orthonormal iff $\gamma_n = \alpha_{n-1}$. Finally, $\gamma_n = 0 \Leftrightarrow \langle P_n, P_n\rangle = 0$.

$\Leftarrow$) Suppose the polynomials satisfy a three-term recursion relation

$$xP_n(x) = \gamma_{n+1} P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x).$$

On $\mathbb{C}[x]$, **define** a functional φ by

$$\varphi\left[P_n P_k\right] = \delta_{nk}$$

and extend linearly: for

$$A(x) = \sum_n a_n P_n(x), \qquad B(x) = \sum_k b_k P_k(x),$$

$$\varphi\left[AB\right] = \sum_n \sum_k a_n \bar{b}_k \varphi\left[P_n P_k\right] = \sum_n a_n \bar{b}_n.$$

Why consistent? Suppose

$$A_1 B_1 = A_2 B_2,$$

to show

$$\varphi\left[A_1 B_1\right] = \varphi\left[A_2 B_2\right].$$

Factoring over $\mathbb{C}$, enough to show

$$\varphi\left[(A(x)(x-a))B(x)\right] = \varphi\left[A(x)((x-a)B(x))\right]$$

for $a \in \mathbb{C}$. By linearity, enough to show

$$\varphi\left[(P_k(x)x)P_n(x)\right] = \varphi\left[P_k(x)(xP_n(x))\right].$$

But from the recursion relation, these are

$$\begin{aligned}\gamma_{k+1}\delta_{k+1,n} + \beta_k\delta_{k,n} + \gamma_k\delta_{k-1,n}&\\ = \gamma_{n+1}\delta_{k,n+1} + \beta_n\delta_{k,n} + \gamma_n\delta_{k,n-1}&.\end{aligned}$$

Remains to show $\varphi = \varphi_\mu$ for some measure μ. Note: μ not unique.

$$\left\langle \sum_{i=0}^{n} a_i P_i(x), \sum_{i=0}^{n} a_i P_i(x) \right\rangle_{\varphi} = \sum_{i=0}^{n} |a_i|^2 > 0$$

so the induced inner product is positive definite (assume non-degenerate) Let $\mathcal{H}$ be the Hilbert space completion of $\mathbb{C}[x]$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\varphi}$, and X be the operator with dense domain $\mathbb{C}[x]$ defined by

$$XA(x) = xA(x).$$

Then X is an unbounded symmetric operator. Has a self-adjoint extension? von Neumann: look at the deficiency subspaces

$$\mathcal{K}_\pm = \ker(X^* \mp i).$$

A self-adjoint extension exists iff the deficiency indices $d_\pm = \dim(\mathcal{K}_\pm)$ are equal.

Let

$$C\Big(\sum_{j=0}^{n} a_j x^j\Big) = \sum_{j=0}^{n} \bar{a}_j x^j.$$

Since

$$\langle C(P), C(P)\rangle_\varphi = \overline{\langle P, P\rangle_\varphi} = \langle P, P\rangle_\varphi\,,$$

C can extend to an operator on $\mathcal{H}$. It commutes with X, so

$$C(\mathcal{K}_\pm) = \mathcal{K}_\mp.$$

Indeed,

$$\begin{aligned}\mathcal{K}_+ &= \{\eta \in \mathcal{H} : X^*\eta = i\eta\} \\ &= \{\eta \in \mathcal{H} : \forall P \in \mathbb{R}[x], \langle \eta, XP\rangle = \langle i\eta, P\rangle\}\,.\end{aligned}$$

But

$$\begin{aligned}\langle C(\eta), X(P)\rangle &= \langle \eta, CX(P)\rangle = \langle \eta, XC(P)\rangle \\ &= \langle i\eta, C(P)\rangle = \langle -i\eta, P\rangle\,.\end{aligned}$$

We conclude that

$$d_+ = d_-.$$

Let $\tilde{X}$ be a self-adjoint extension of X. It has a spectral measure: for any subset $S \subset \mathbb{R}$, have a projection $E(S)$. Define

$$\mu(S) = \langle 1, E(S)1 \rangle .$$

This is a probability measure. In fact,

$$\begin{aligned} m_n(\mu) &= \int_{\mathbb{R}} x^n \, d\mu(x) = \left\langle 1, \int_{\mathbb{R}} x^n \, dE(x) \, 1 \right\rangle \\ &= \left\langle 1, \tilde{X}^n 1 \right\rangle = \langle 1, x^n \rangle = \varphi \left[x^n \right] . \end{aligned}$$

$\tilde{X}$ not unique $\Rightarrow$ μ not unique.

Hermite:

$$xP_n = P_{n+1} + tnP_{n-1}.$$

Laguerre:

$$xL_n = L_{n+1} + (t+2n)L_n + n(t+n-1)L_{n-1}.$$

Jacobi:

$$xP_n = P_{n+1} + \frac{\beta^2-\alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)}P_n + \frac{4n(n+\alpha)(n+\beta)(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)^2(2n+\alpha+\beta+1)}P_{n-1}.$$

Ultraspherical:

$$xP_n = P_{n+1} + \frac{n(n+2\lambda-1)}{4(n+\lambda-1)(n+\lambda)}P_{n-1}.$$

Chebyshev:

$$xP_n = P_{n+1} + \frac{1}{4}P_{n-1}.$$

Kravchuk:

for $a = (1 - 2p)$, $b = 4pq$, $a^2 - b = 1$,

$$xK_n = K_{n+1} + \mathbf{a}(2n - N)K_n + \mathbf{b}n(N - n + 1)K_{n-1}.$$

Charlier:

$$xP_n = P_{n+1} + (t + n)P_n + tnP_{n-1}.$$

Meixner:

for $a = p^{-1} - \frac{1}{2}$, $b = qp^{-2}$, $a^2 - b = \frac{1}{4} > 0$,

$$xM_n = M_{n+1} + \mathbf{a}(t + 2n)M_n + \mathbf{b}n(t + n - 1)M_{n-1}.$$

Meixner-Pollaczek: same for $a^2 - b < 0$. Note for Laguerre, $a^2 - b = 0$.

Can study zeros, asymptotics, continued fractions. Instead:

Denote $\{e_n\}$ the basis of $\mathcal{H}$ coming from $\{P_n\}$

$$Xe_n = \gamma_{n+1} e_{n+1} + \beta_n e_n + \gamma_n e_{n-1}.$$

X has a tri-diagonal matrix

$$\begin{pmatrix} \beta_0 & \gamma_1 & & & \\ \gamma_1 & \beta_1 & \gamma_2 & & \\ & \gamma_2 & \beta_2 & \gamma_3 & \\ & & \gamma_3 & \ddots & \ddots \\ & & & \ddots & \end{pmatrix}.$$

Thus X is a sum of three operators: creation

$$A^* e_n = \gamma_{n+1} e_{n+1},$$

annihilation

$$A e_n = \gamma_n e_{n-1},$$

and preservation

$$N e_n = \beta_n e_n.$$

Since β_n is real, N is symmetric, and

$$\begin{aligned} \langle A^* e_n, e_k \rangle &= \gamma_{n+1} \left\langle e_{n+1}, e_k \right\rangle = \gamma_{n+1} \delta_{n+1,k} \\ &= \gamma_k \langle e_n, e_{k-1} \rangle = \langle e_n, A e_{n-1} \rangle . \end{aligned}$$

A, A^* are (formal) adjoints of each other.

For the Hermite polynomials, $N = 0$, A, A^* are the usual annihilation and creation operators.

For the Charlier polynomials, A, A^* are as before, and N is (almost) the number operator.

Another way to look at this. Heisenberg Lie algebra H_t has the basis $\{A, A^*, t\}$ subject to the relation

$$[A, A^*] = t, \quad t \text{ scalar.}$$

Each H_t has a representation on some Hilbert space $\mathcal{H}$ with basis $\{e_n\}_{n=0}^{\infty}$:

$$\begin{aligned} A\, e_n &= \sqrt{tn}\, e_{n-1}, \\ A^* e_n &= \sqrt{t(n+1)}\, e_{n+1}. \end{aligned}$$

$$(AA^* - A^*A)e_n = [t(n+1) - tn]e_n = te_n.$$

If we want

$$A + A^* = X,$$

where

$$XP(x) = xP(x),$$

can take

$$\mathcal{H} = L^2(\mathbb{R}, \mu_t)$$

for

$$\mu_t = \text{ Gaussian measure},$$

and

$$e_n = P_n(x,t) = \text{Hermite polynomials.}$$

In this representation,

$$A = t\partial_x, \qquad A^* = x - t\partial_x.$$

Oscillator Lie algebra Os_t has the basis $\{A, A^*, N, t\}$ subject to the relations

$$[A, A^*] = t, \quad [A, N] = A, \quad [N, A^*] = A^*, \quad t \text{ scalar.}$$

Each Os_t has a representation on some $\mathcal{H}$:

$$\begin{aligned} A\, e_n &= \sqrt{tn}\, e_{n-1}, \\ A^* e_n &= \sqrt{t(n+1)}\, e_{n+1}, \\ N\, e_n &= (t+n) e_n. \end{aligned}$$

Note e_0 is a cyclic vector. In fact, the representation is irreducible. If we want

$$A + A^* + N = X,$$

can take $\mathcal{H} = L^2(\mathbb{R}, \mu_t)$ for

$$\mu_t = \text{Poisson measure},$$

and

$$e_n = P_n(x, t) = \text{Charlier polynomials.}$$

$$A = t\Delta, \quad \Delta(P) = P(x) - P(x-1),$$
$$A^*(P) = xP(x-1) - tP, N(P) = x\Delta P + tP(x-1).$$

The Lie algebra $sl(2,\mathbb{R})$ of real traceless 2×2 matrices has the basis $\left\{J_-, J_+, J_0\right\}$

$$J_- = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad J_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

subject to the relations

$$[J_-, J_+] = J_0, \quad [J_-, J_0] = 2J_-, \quad [J_0, J_+] = 2J_+.$$

Abstract representations, parameterized by $t > 0$:

$$\begin{aligned} J_- \, e_n &= \sqrt{n(t+n-1)}\, e_{n-1}, \\ J_+ \, e_n &= \sqrt{(n+1)(t+n)}\, e_{n+1}, \\ J_0 \, e_n &= (t+2n) e_n. \end{aligned}$$

Want

$$J_- + J_+ + cJ_0 = X.$$

For each t, c, $sl(2, \mathbb{R})$ has a representation on $L^2(\mathbb{R}, \mu_{t,c})$, for

$$\mu_{t,c} = \begin{cases} \text{negative binomial (Meixner)}, & c > 1, \\ \text{Gamma (Laguerre)}, & c = 1, \\ \ldots \text{ (Meixner-Pollaczek)}, & 0 < c < 1. \end{cases}$$

Where are the finite-dimensional representations?

Take $t = -N$, $N > 0$ integer.

$$\begin{aligned} J_- \, e_n &= -\sqrt{n(N-n+1)}\, e_{n-1}, \\ J_+ \, e_n &= \sqrt{(n+1)(N-n)}\, e_{n+1}, \\ J_0 \, e_n &= (2n - N) e_n. \end{aligned}$$

Finite-dimensional since $J_+ e_N = 0$. Can take

$$\mu_N = \text{binomial measure},$$
$$e_n = P_n(t, N) = \text{Kravchuk polynomials.}$$

For example, for $N = 1$,

$$d\mu(x) = q\delta_{-1}(x) + p\delta_1(x)$$

is the Bernoulli distribution,

$$\begin{aligned} K_0 &= 1, \\ K_1 &= x + (1 - 2p). \end{aligned}$$

In the basis (K_1, K_0) for

$$L^2(\mathbb{R}, \mu) = \mathbb{C}^2,$$

J_-, J_+, J_0 have matrices as above, and

$$X = J_+ - 4pqJ_- + (q - p)J_0 = \begin{pmatrix} q-p & 1 \\ 4pq & p-q \end{pmatrix}$$

sends

$$X : (K_0 = 1) \mapsto \Big(K_1 - (1 - 2p)K_0 = x\Big)$$

and

$$\begin{aligned} X : K_1 \mapsto (q-p)K_1 &+ 4pqK_0 \\ &= (q-p)x + (1-2p)^2 + 4p(1-p) \\ &= (q-p)x + 1. \end{aligned}$$

Note that for $x = -1$,

$$xK_1(x) = 2p = (q - p)x + 1,$$

and for $x = 1$,

$$xK_1(x) = 2q = (q - p)x + 1,$$

so

$$xK_1(x) = (q - p)x + 1$$

and

$$J_+ - 4pqJ_- + (q - p)J_0 = \text{ multiplication by } x.$$

Additional material:

1) The Hilbert space $\mathcal{H}$ (the closure of $\mathbb{C}[x]$ with respect to $\langle \cdot, \cdot \rangle_\mu$) is not always equal to $L^2(\mathbb{R}, \mu)$. In fact, $\mathbb{C}[x]$ is dense in $L^2(\mathbb{R}, \mu)$ if and only if μ comes from a self-adjoint extension of a symmetric operator as in slide 12; there may be measures that do not.

2) For the uses of representation theory to obtain results about orthogonal polynomials, see reference 4.

3) Different representations of the Lie algebra $sl(2, \mathbb{R})$ on slide 21, distinguished by the parameter c, correspond to representations of different Lie groups all of which have the same Lie algebra. See reference 3.

REFERENCES

1. N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Hafner Publishing Co., New York, 1965.

2. Barry Simon, *The classical moment problem as a self-adjoint finite difference operator*, Adv. Math. **137** (1998), no. 1, 82–203.

3. Philip Feinsilver and René Schott, *Algebraic structures and operator calculus. Vol. I*, Kluwer Academic Publishers Group, Dordrecht, 1993, Representations and probability theory.

4. H. T. Koelink and J. Van Der Jeugt, *Convolutions for orthogonal polynomials from Lie and quantum algebra representations*, SIAM J. Math. Anal. **29** (1998), no. 3, 794–822 (electronic).

5. Erik Koelink, *Spectral theory and special functions*, Lecture notes for the SIAM Activity Group OP-SF summer school 2000, Laredo, Spain. arXiv:math.CA/0107036.