Orthogonality of free Sheffer systems

Michael Anshelevich

August 5, 2004

$\mu =$ infinitely divisible probability measure, $\{\mu_t : t \in [0, \infty)\}$ the corresponding convolution semigroup,

$$\mu_t * \mu_s = \mu_{t+s}.$$

Assume all their moments are finite, μ centered, variance 1.

$$k(z) = \log \int_{\mathbb{R}} e^{xz} \, d\mu(x) = \log \sum_{n=0}^{\infty} \frac{1}{n!} m_n(\mu) z^n$$

is the cumulant-generating function, $k(z) = \frac{1}{2} z^2 + \ldots$.

Let $U = z + \ldots$ be an analytic function / formal power series. The Sheffer polynomials (Sheffer 1937) are defined by

$$\sum_{n=0}^{\infty} \frac{1}{n!} P_n(x,t) z^n = \exp\Big(xU(z) - tk(U(z)) \Big).$$

Note $P_n(x,t) = x^n + \ldots$.

Martingale polynomials for the corresponding Lévy processes; main objects in umbral calculus; exponential families in statistics.

Relation to the measure:

$$\int_{\mathbb{R}} 1\, d\mu_t(x) = 1,$$
$$\int_{\mathbb{R}} P_n(x,t)\, d\mu_t(x) = 0.$$

When have moreover

$$\int_{\mathbb{R}} P_n(x,t) P_k(x,t)\, d\mu_t(x) = \langle P_n, P_k \rangle = 0$$

for $n \neq k$? An important subclass: Meixner family. Many properties and characterizations.

1. ORTHOGONALITY.

Meixner polynomials are the orthogonal Sheffer polynomials.

2. EXPLICIT FORM. (Meixner 1934)

Three-term recursion relations

$$x P_n = P_{n+1} + \alpha n P_n + n(t + \beta(n-1)) P_{n-1}.$$

Two degenerate cases:

$\alpha = \beta = 0$: Hermite, orthogonal with respect to the Gaussian distribution.

$\beta = 0$, $\alpha = 1$: Charlier, orthogonal with respect to the centered Poisson distribution.

Otherwise, fix $\beta = 1$.

$\alpha = 2$: Laguerre, orthogonal with respect to the centered Gamma distribution.

$\alpha > 2$: Meixner, orthogonal with respect to the centered negative binomial distribution.

$0 \leq \alpha < 2$: Meixner-Pollaczek, orthogonal with respect to the centered continuous binomial (hyperbolic cosine) distribution.

If not infinitely divisible, an extra case with $t = -N$ integer: Krawtchouk, orthogonal with respect to the binomial distribution.

3. QUADRATIC VARIANCE FUNCTION. (Morris 1983)

Sheffer polynomials are Meixner if

$$\sum_{n=0}^{\infty} \frac{1}{n!} P_n(x,t) z^n = \exp\Big(xU(z) - tk(U(z))\Big)$$

for

a) $U(z) = (k'(z))^{<-1>}$ (inverse under composition), and

b) $k''(z) = a(k'(z))^2 + bk'(z) + 1$.

4. QUADRATIC TWO-SIDED CONDITIONAL VARIANCE. (Wesołowski 1993)

Let $\{X_t : t \geq 0\}$ be a separable square-integrable stochastic process, normalized so that for all $t, s > 0$

$$\mathrm{E}[X_t] = 0, \mathrm{E}[X_t X_s] = \min(t, s),$$

and for all $0 \leq s < t < u$,

$$\mathrm{E}\left[X_t | \mathcal{F}_{\leq s} \vee \mathcal{F}_{\geq u}\right] = aX_s + bX_u,$$

$$\mathrm{Var}\left[X_t | \mathcal{F}_{\leq s} \vee \mathcal{F}_{\geq u}\right] = C_2(X_u - X_s)^2 + C_1(X_u - X_s) + C_0,$$

where C_0, C_1, C_2 are deterministic functions of $s < t < u$, and $C_2 \neq ab$. If for every $t > 0$ the distribution of X_t has at least 3 point support, then the process has independent increments, with marginal and conditional distributions belonging to the Meixner family.

5. BASIS FOR REPRESENTATIONS OF sl_2.

Relations

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

can be realized as

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

with

$$J_- + J_+ + aJ_0 = X.$$

6. LÉVY MEASURE BELONGS TO ITS INFINITELY DIVISIBLE FAMILY.

$$\log \int_{\mathbb{R}} e^{i\theta x}\, d\mu_t(x)$$

$$= t\left[i\gamma\theta + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2} \right) \frac{1+x^2}{x^2}\, d\nu(x) \right].$$

With some fudging, $\mu = \nu$.

MULTIVARIATE FAMILIES.

Let μ be an infinitely divisible measure on $\mathbb{R}^n$, centered.

$$k(\mathbf{z}) = \log \int_{\mathbb{R}^n} e^{\mathbf{x}\cdot\mathbf{z}}\, d\mu(\mathbf{x}) = \log\left(1+\sum_{\vec{u}} \frac{1}{|\vec{u}|!} m_{\vec{u}}(\mu) z_{\vec{u}}\right)$$

is the cumulant-generating function,

$$k(\mathbf{z}) = \sum_{i,j=1}^{n} v_{ij} z_i z_j + \ldots,$$

$V = [v_{ij}] =$ covariance matrix.

Let $\mathbf{U} = (U_1, U_2, \ldots, U_n)$ be an n-tuple of formal power series in $(z_1, \ldots, z_n)$, $U_i = z_i + \ldots$. The Sheffer polynomials are defined by

$$1 + \sum_{\vec{u}} \frac{1}{|\vec{u}|!} P_{\vec{u}}(\mathbf{x}, t) z_{\vec{u}} = \exp\Big(\mathbf{x}\cdot\mathbf{U}(\mathbf{z}) - tk(\mathbf{U}(\mathbf{z}))\Big).$$

Note $P_{\vec{u}}(\mathbf{x}, t) = x_{\vec{u}} + \ldots$.

Meixner: orthogonal among these.

Say $\{P_{\vec{u}}\}$ are pseudo-orthogonal with respect to μ if

$$\int_{\mathbb{R}^n} P_{\vec{u}}(\mathbf{x})^* P_{\vec{v}}(\mathbf{x})\, d\mu(\mathbf{x}) = 0$$

whenever $|\vec{u}| \neq |\vec{v}|$. They are orthogonal if this is the case whenever $\vec{u} \neq \vec{v}$.

Again, Sheffer polynomials are Meixner if

a) $\mathbf{U}(\mathbf{z}) = (\nabla k(z))^{<-1>}\Big((V\mathbf{z})_1, (V\mathbf{z})_2, \ldots, (V\mathbf{z})_n\Big)$, and

b) For each i, j, $A_{i,j} =$ matrix, $B_{ij} =$ vector:

$$\partial_i \partial_j k(\mathbf{z}) = \Big(A_{ij} \nabla k(z)\Big) \cdot \nabla k(z) + B_{ij} \cdot \nabla k(\mathbf{z}) + v_{ij}.$$

(Feinsilver 1986, Pommeret 1996)

Again simple recursion relations. Complete set of measures of orthogonality: unknown!

Known cases: (Letac 1988, Casalis 1991, 1996)

Linear: $\partial_i \partial_j k(\mathbf{z}) = B_{ij} \cdot \nabla k(\mathbf{z}) + C_{ij}$.

Products of Gaussian and Poisson distributions.

Simple quadratic:

$$\partial_i \partial_j k(\mathbf{z}) = A\, \partial_i k(\mathbf{z})\, \partial_j k(\mathbf{z}) + B_{ij} \cdot \nabla k(\mathbf{z}) + C_{ij}.$$

$2n + 4$ families (including non-infinitely-divisible), marginal and conditional distributions belong to the Meixner class.

Homogeneous quadratic:

$$\partial_i \partial_j k(\mathbf{z}) = \Big(A_{ij} \nabla k(\mathbf{z})\Big) \cdot \nabla k(\mathbf{z}).$$

Wishart distributions on symmetric cones.

Free analogs?

$$\sum_{n=0}^{\infty} P_n(x,t) z^n = \frac{1}{1 - xU(z) + tR(U(z))}.$$

Here $R(z)$ = free cumulant generating function of the measure μ, $z \times$the R-transform. Why are these free Sheffer polynomials?

a) Give martingale polynomials for free Lévy processes. (Biane 1998, Voiculescu 2000)

b) In the particular case $U(z) = z$ (free Appell), many combinatorial formulas, identical to the formulas for the Appell polynomials, with the lattice of all partitions replaces by non-crossing partitions. (M.A. 2004)

Free Meixner class (orthogonal free Sheffer):

a) $U(z) = (R(z)/z)^{<-1>}$, and

b) $R(z)/z^2 = a(R(z)/z)^2 + bR(z)/z + 1$.

Recursion relations: almost constant,

$$\begin{aligned} xP_0 &= P_1, \\ xP_1 &= P_2 + \alpha P_1 + tP_0, \\ xP_n &= P_{n+1} + \alpha P_n + (t+\beta)P_{n-1}. \end{aligned}$$

Measures: semicircular, free Poisson, and new families similar to free Poisson. (M.A. 2003)

Free binomial distribution: sum of N freely independent projections. Polynomials of this form.

MULTI-VARIATE SITUATION.

$\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{z} = (z_1, \ldots, z_n)$ n-tuples of non-commuting indeterminates.

Define

$$1 + \sum_{\vec{u}} P_{\vec{u}}(\mathbf{x}) z_{\vec{u}} = F(\mathbf{z})\Big(1 - \mathbf{x} \cdot \mathbf{V}(\mathbf{z})\Big)^{-1}$$

(inverse under multiplication), where $\mathbf{V}(\mathbf{z}) =$ n-tuple of formal power series,

Then in fact, $F(z) = \Big(1 + R(\mathbf{U}(\mathbf{z}))\Big)^{-1}$,

$$V_i(z) = U_i(z) F(z),$$

where $R(\mathbf{z}) =$ free cumulant generating function of φ, a state on $\mathbb{R}\langle x_1, x_2, \ldots, x_n \rangle$.

φ is determined by: $\varphi[1] = 1$, $\varphi[P_{\vec{u}}] = 0$.

$$1 + \sum_{\vec{u}} P_{\vec{u}}(\mathbf{x}, t) z_{\vec{u}} = \Big(1 - \mathbf{x} \cdot \mathbf{U}(\mathbf{z}) + tR(\mathbf{U}(\mathbf{z}))\Big)^{-1}$$

$\{P_{\vec{u}}\}$ pseudo-orthogonal with respect to a functional φ if

$$\varphi\left[P_{\vec{u}}^{*} P_{\vec{v}}\right] = 0$$

whenever $|\vec{u}| \neq |\vec{v}|$. Orthogonal if this is the case whenever $\vec{u} \neq \vec{v}$.

Define a linear operator D_j (right partial derivative) on non-commutative power series via

$$D_j w_{u(1),\ldots,u(n)} = \begin{cases} 0, & u(n) \neq j, \\ w_{u(1),\ldots,u(n-1)}, & u(n) = j. \end{cases}$$

Proposition. If the polynomials are pseudo-orthogonal, then

$$(D_i R)(\mathbf{U}(\mathbf{z})) = \sum_{j=1}^{n} v_{ij} z_j.$$

Proposition. If the polynomials are pseudo-orthogonal, then

$$z_i = U_i + \sum_{\alpha,j} b_{i,\alpha,j} z_\alpha U_j + \sum_{\alpha,\beta,j} b_{i,(\alpha,\beta),j} z_\alpha z_\beta U_j.$$

No higher-order terms.

$$\begin{aligned} P_{\vec{u},\alpha,\beta} x_j = P_{\vec{u},\alpha,\beta,j} &+ \sum_i b_{i,\beta,j} P_{\vec{u},\alpha,i} \\ &+ \sum_i (\delta_{i\alpha} v_{j\beta} + b_{i,(\alpha,\beta),j}) P_{\vec{u},i}. \end{aligned}$$

Coefficients do not depend on $\vec{u}$.

$$\begin{aligned} D_i D_j R(\mathbf{z}) = \sum_{\alpha,\beta} & A_{ij}^{\alpha\beta} D_\alpha R(\mathbf{z})\ D_\beta R(\mathbf{z}) \\ &+ \sum_\alpha B_{ij}^{\alpha} D_\alpha R(\mathbf{z}) + v_{ij}. \end{aligned}$$

If $V = \mathrm{Id}$, then

$$\begin{aligned} D_i D_j R(\mathbf{z}) = \sum_{\alpha,\beta} & b_{i,(\alpha,\beta),j} D_\alpha R(\mathbf{z})\ D_\beta R(\mathbf{z}) \\ & + \sum_{\alpha} b_{i,\alpha,j} D_\alpha R(\mathbf{z}) + \delta_{ij}. \end{aligned}$$

In the proof: if polynomials are pseudo-orthogonal, they satisfy a recursion

$$P_{\vec{u}} x_j = P_{\vec{u},j} + \sum_{|\vec{w}|=|\vec{u}|} B_{\vec{w},\vec{u},j} P_{\vec{w}} + \sum_{|\vec{v}|=|\vec{u}|-1} C_{\vec{v},\vec{u},j} P_{\vec{v}},$$

Also, use the right-acting

$$D_{z_i} : z_{\vec{u},j} = \delta_{ij} z_{\vec{u}},$$

left-acting

$$G_{x_i} : P_{j,\vec{u}} = \delta_{ij} P_{\vec{u}},$$

and the main identity: for

$$M(\mathbf{w}) = \sum_{\vec{u}} \varphi\left[x_{\vec{u}}\right] w_{\vec{u}}$$

$$R(w_1(1 + M(\mathbf{w})), \ldots, w_n(1 + M(\mathbf{w}))) = M(\mathbf{w}).$$

Added after the talk: more properties of the single-variable Meixner class.

7. QUADRATIC REGRESSION. (Laha, Lukacs 1960)

$X = (X_1, X_2, \cdots, X_n) = n$-tuple of i.i.d. random variables. Let $Q = X^t A X$ and $\Lambda = a^t X$. If the regression $\mathsf{E}[Q|\Lambda]$ is quadratic in Λ, then the distribution of X_i belongs to the Meixner class.

8. EXPLICIT LINEARIZATION COEFFICIENTS. (Kim, Zeng 2001)

Changing notation, assume the recursion

$$xP_n = P_{n+1} + (\alpha+\beta)nP_n + n(t+\alpha\beta(n-1))P_{n-1}.$$

Linearization coefficients are integrals of products of polynomials with respect to the orthogonality measure:

$$\int_{\mathbb{R}} P_{n_1}(x)P_{n_2}(x)\dots P_{n_k}(x)\,d\mu(x)$$

For the Meixner class, they are equal to

$$\sum_{\sigma\in\mathsf{Sym}(n_1,n_2,\dots,n_k)} \alpha^{a(\sigma)-\mathsf{cyc}(\sigma)}\beta^{d(\sigma)-\mathsf{cyc}(\sigma)}t^{\mathsf{cyc}(\sigma)}.$$

σ is a permutation, $a(\sigma) =$ number of ascents of σ (i with $i < \sigma(i)$), $d(\sigma) =$ number of descents of σ (i with $i > \sigma(i)$), $\mathsf{cyc}(\sigma) =$ number of cycles of σ.

$\pi_{n_1,n_2,\dots,n_k} =$ set partition whose k classes are intervals of lengths $n_1, n_2, \dots, n_k$. $\mathsf{Sym}(n_1, n_2, \dots, n_k) =$ generalized derangements $=$ inhomogeneous permutations $= \sigma$ such that for all i, $\sigma(i)$ and i lie in different classes of $\pi_{n_1,n_2,\dots,n_k}$.