Multivariate Krawtchouk polynomials as Bernoulli systems

Philip Feinsilver

Southern Illinois University

Carbondale, Illinois USA 62901

Multivariate Krawtchouk polynomials are constructed.
The associated Bernoulli systems data are studied and the associated Lie algebra is found. The operator formulation of the corresponding observables provides recurrences satisfied by the multivariate Krawtchouk polynomials.

The 1st GSIS-RCPAM International Symposium
Tohoku University
Sendai, Japan
March 8, 2013

1 M.P. Krawtchouk

Krawtchouk polynomials are part of the legacy of Mikhail Kravchuk.

A symposium in honor of his work and memory was held in Kiev and an accompanying volume was produced that is most highly recommended:

N. Virchenko, et al., eds.

Development of the Mathematical Ideas of

Mykhailo Kravchuk (Krawtchouk),

Shevchenko Scientific Society, Kyiv-New York, 2004.

Krawtchouk polynomials appear in diverse areas of mathematics and science. Applications range from coding theory to image processing.

2 Krawtchouk polynomials in one variable and the binomial distribution

Krawtchouk polynomials may be defined via the generating function

$$(1+pv)^{N-x}(1-qv)^x = \sum_{0\le k\le N} v^k\, K_k(x,N)$$

The polynomials $K_k(x,N)$ are orthogonal with respect to the binomial distribution with parameters N, p. The associated probabilities have the form

$$\left\{ \binom{N}{0} q^N p^0, \ldots, \binom{N}{k} q^{N-k} p^k, \ldots, \binom{N}{N} q^0 p^N \right\}$$

2.1 Matrix formulation

Setting $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ we have

$$y_1^{N-x} y_2^x = \sum_k v_1^{N-k} v_2^k \Phi_{kx}$$

and the expression of orthogonality takes the form

$$\Phi B \mathrm{p} \Phi^\top = D$$

where B is the diagonal matrix with entries the binomial coefficients $\binom{N}{k}$, the matrix p is diagonal with entries the corresponding probabilities $q^{N-k} p^k$ and D is the diagonal matrix of squared norms, standardized by $D_{00} = 1$.

3 Symmetric tensor powers

➧ **Given** a $d \times d$ matrix A, the action on the symmetric tensor algebra of the underlying vector space defines its second quantization or "symmetric representation".

Introduce commuting variables $v_1, \ldots, v_d$. Map

$$y_i = \sum_j A_{ij} v_j$$

We will use multi-indices, $m = (m_1, \ldots, m_d)$, $m_i \geq 0$, similarly for n.

The induced matrix, $\bar{A}$, at level (homogeneous degree) $N = \sum n_i$ has entries $\bar{A}_{nm}$ determined by the expansion

$$y^n = y_1^{n_1} \cdots y_d^{n_d} = \sum_m \bar{A}_{nm} x^m \ .$$

➧ The map $A \to \bar{A}$ is at each level a **multiplicative homomorphism**,

$$\overline{A_1 A_2} = \bar{A}_1 \, \bar{A}_2 \ .$$

Multinomial matrix

We introduce the special matrix B which is a diagonal matrix with multinomial coefficients as entries

$$B_{nm} = \delta_{nm} \binom{N}{n} = \frac{N!}{n_1!\, n_2! \,\cdots n_d!} \; .$$

B is the diagonal of the induced matrix at level N of the all 1's matrix.

If p is a diagonal matrix with entries $p_i > 0$, $\sum_i p_i = 1$, then the diagonal matrix

$$B\overline{\mathrm{p}}$$

yields the probabilities for the corresponding multinomial distribution.

3.1 Transpose Lemma

The main lemma is the relation between the induced matrix of A with that of its transpose, $A^\top$.

Transpose Lemma.

The induced matrices at each level satisfy

$$\overline{A^\top} = B^{-1} \bar{A}^\top B \; .$$

Proof: is accomplished by expanding the bilinear form $(\sum\limits_{i,j} u_i A_{ij} v_j)^N$ two ways, in terms of the induced matrix for A, then that of $A^\top$, and comparing.

4 Construction of Krawtchouk polynomial systems

Start with U, an orthogonal (unitary) matrix.

Make all entries of first column positive by taking out phases and form the probability matrix thus

$$\mathrm{p} = \begin{pmatrix} U_{00}^2 & & \\ & \ddots & \\ & & U_{d0}^2 \end{pmatrix} = \begin{pmatrix} p_0 & & \\ & \ddots & \\ & & p_d \end{pmatrix}$$

row and column indices running from 0 to d.

Define

$$A = \frac{1}{\sqrt{\mathrm{p}}}\, U \sqrt{D}$$

where D is diagonal with all positive entries on the diagonal. The **essential property** satisfied by A is

$$A^\top \mathrm{p} A = D\,.$$

4.1 Krawtchouk systems

In any degree N, the induced matrix $\bar{A}$ satisfies

$$\overline{A^\top}\bar{\mathrm{p}}\bar{A} = \bar{D}\,.$$

Using the Transpose Lemma

$$B\overline{A^\top} = \bar{A}^\top B$$

with B the special multinomial diagonal matrix yields

$$\Phi\, B\bar{\mathrm{p}}\,\Phi^\top = B\bar{D}$$

the Krawtchouk matrix Φ being thus defined as $\bar{A}^\top$.

The entries of Φ are the values of the **multivariate Krawtchouk polynomials** thus determined.

$B\bar{D}$ is the diagonal matrix of squared norms according to the orthogonality of the Krawtchouk polynomial system.

♫ Example

Start with the orthogonal matrix $U = \begin{pmatrix} \sqrt{q} & \sqrt{p} \\ \sqrt{p} & -\sqrt{q} \end{pmatrix}$.

Factoring out the squares from the first column we have

$$\mathrm{p} = \begin{pmatrix} q & 0 \\ 0 & p \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix} .$$

We have

$$A^\top \mathrm{p} A = \begin{pmatrix} 1 & 0 \\ 0 & pq \end{pmatrix} = D .$$

Take $N = 4$.

We have the Kravchuk matrix

$$\Phi = \bar{A}^{\top} =$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 4p & -q+3p & -2q+2p & -3q+p & -4q \\ 6p^2 & -3pq+3p^2 & q^2-4pq+p^2 & 3q^2-3pq & 6q^2 \\ 4p^3 & -3p^2q+p^3 & 2pq^2-2p^2q & -q^3+3pq^2 & -4q^3 \\ p^4 & -p^3q & p^2q^2 & -pq^3 & q^4 \end{pmatrix}.$$

p becomes the induced matrix

$$\bar{\mathrm{p}} = \begin{pmatrix} q^4 & 0 & 0 & 0 & 0 \\ 0 & q^3p & 0 & 0 & 0 \\ 0 & 0 & q^2p^2 & 0 & 0 \\ 0 & 0 & 0 & qp^3 & 0 \\ 0 & 0 & 0 & 0 & p^4 \end{pmatrix}.$$

and the binomial coefficient matrix

$$B = \mathrm{diag}(1, 4, 6, 4, 1).$$

5 Appell and Bernoulli systems

An **Appell system** of polynomials is a sequence $\{\phi_n(x)\}$ such that

1. $\deg \phi_n = n$
2. $\partial_x \phi_n = n\phi_{n-1}$ where $\partial_x = \dfrac{d}{dx}$.

Introduce the `raising operator`

$$R\phi_n = \phi_{n+1}$$

The pair ∂_x, R satisfy the commutation relations

$$[\partial_x, R] = I$$

of the Heisenberg-Weyl algebra or boson commutation relations.

➧ **We take** as generating function for the sequence $\{\phi_n\}$

$$e^{xz-tH(z)} = \sum_{n=0}^{\infty} \frac{z^n}{n!}\,\phi_n(x,t)$$

introducing an additional "time" variable, with

$$\int_{-\infty}^{\infty} e^{zx}\,p_t(dx) = e^{tH(z)}$$

for a convolution family of probability measures, p_t.

Note that t may be running only through discrete values.

In the infinitely divisible case, we have the exponential martingale for the corresponding process with independent increments.

➧ We extend to the multivariate case, taking $\partial_j = d/dx_j$, with R_i raising the index n_i to $n_i + 1$. In the exponent, $xz = x \cdot z = \sum x_i z_i$.

5.1 Canonical Appell system

Consider canonical raising and lowering (*velocity*) operators defined by

$$\mathcal{V}_j \phi_n = n_j \phi_{n-e_j} \qquad \text{and} \qquad \mathcal{R}_i \phi_n = \phi_{n+e_i}$$

satisfying

$$[\mathcal{V}_j, \mathcal{R}_i] = \delta_{ij}\, I$$

where $\mathcal{V} = (\mathcal{V}_1, \ldots, \mathcal{V}_d)$ is given by a function V of $\partial = (\partial_1, \ldots, \partial_d)$, analytic in a neighborhood of 0 in $\mathbb{C}^n$, with a locally analytic inverse.

The generating function takes the form

$$e^{xz - tH(z)} = \sum_{n \geq 0} \frac{V(z)^n}{n!}\, \phi_n(x, t)$$

with multi-index $n! = n_1! \cdots n_d!$.

5.2 Bernoulli systems

A **Bernoulli system** is a canonical Appell system such that, for each t, the polynomials $\{\phi_n(x,t)\}$ form an orthogonal system with respect to the measure p_t.

To indicate this, write J_n for the corresponding canonical Appell sequence.

Operator formulation

➨ The operator form of the generating function is the exponential of the raising operators acting on the vacuum state:

$$e^{V(z)\mathcal{R}}\Omega = e^{xz-tH(z)} = \sum_{n\geq 0} \frac{V(z)^n}{n!} J_n(x,t)$$

where the `vacuum state` Ω is here the constant function equal to 1.

➨ Introducing the inverse function $U(v)$, $z = U(V(z))$, the generating function takes the form

$$e^{v\mathcal{R}}\Omega = e^{xU(v)-tH(U(v))} = \sum_{n\geq 0} \frac{v^n}{n!} J_n(x,t) \ .$$

6 Quantization

➧ Introduce the operators X_j, multiplication by x_j.

Start with

$$e^{z \cdot X}\Omega = e^{tH(z)}\, e^{V(z)\mathcal{R}}\,\Omega\,.$$

Differentiating with respect to z_j we have

$$X_j = t\,\frac{\partial H}{\partial z_j} + \sum_i \mathcal{R}_i\,\frac{\partial V_i}{\partial z_j}$$

These yield a family of commuting self-adjoint operators.

These are the **observables** of the system.

➧ As operators the variables z_j act as the partial differentiation operators, ∂_j.

6.1 Specification of the system

- Express all operators in terms of the canonical raising and velocity operators $\mathcal{R}_i, \mathcal{V}_j$.

- Introduce the lowering operators $\{\mathcal{L}_1, \ldots, \mathcal{L}_d\}$, where $\mathcal{L}_i$ is adjoint to $\mathcal{R}_i$.

- Study the Lie algebra generated by the raising and lowering operators.

- Express X_j in manifestly self-adjoint form.

7 Multinomial distribution. Notations.

Consider a process that at each time step does one of $d+1$ possibilities:

1. With probability p_0, none of the levels 1 through d increase.

2. With probability p_i, $1 \leq i \leq d$, level i increases by 1.

The corresponding moment generating function for one time step is

$$
\begin{aligned}
p_0 + \sum_i p_i e^{z_i} &= 1 + \sum_i p_i (e^{z_i} - 1) \\
&= p_\mu e^{z_\mu}
\end{aligned}
$$

where we set $z_0 = 1$.

➧ N-step process

The moment generating function for N steps is thus

$$(p_\mu e^{z_\mu})^N$$

Notations

We use the following *summation convention*

- repeated Greek indices λ, μ are summed from 0 to d.
- Latin indices i, j, k, run from 1 to d and are not summed, unless explicitly indicated.

8 Multivariate Krawtchouk polynomials as Bernoulli systems

In degree N, we have

$$(Av)^x = \sum_n v^n \Phi_{nx} = \sum_n v^n K_n(x, N) \,.$$

More explicitly,

$$(A_{0\mu} v_\mu)^{N - \sum x_i} (A_{1\mu} v_\mu)^{x_1} \cdots (A_{d\mu} v_\mu)^{x_d} \\ = \sum v^n K_n(x, N) \,.$$

Recall that the first column of A consists of all 1's, and set $\alpha_0 = A_{00} = 1$, $\alpha_i = A_{0i}$, $1 \le i \le d$.

We get

$$(\alpha_\mu v_\mu)^N \prod_i \left(\frac{A_{i\mu} v_\mu}{\alpha_\mu v_\mu} \right)^{x_i} = e^{x \cdot U(v) - N H(U(v))}$$

as the generating function for a Bernoulli system.

8.1 Identification of Bernoulli constituents

We have $t = N$ and

➨

$$H(z) = \log p_\mu e^{z_\mu}$$

with the identification

$$p_\mu e^{z_\mu} = \frac{1}{\alpha_\mu V_\mu} \,.$$

And U, the inverse to V, is given by

➨

$$U_k(v) = \log \frac{A_{k\mu} v_\mu}{\alpha_\mu v_\mu} \,.$$

8.2 Canonical velocity operators

Let

$$C = A^{-1} = D^{-1} A^{\top} \mathrm{p} \, .$$

Then we have

$$V_k = \frac{1}{p_\mu e^{z_\mu}} \, C_{k\lambda} e^{z_\lambda} \; .$$

We find the Riccati partial differential equations

$$\frac{\partial V_i}{\partial z_j} = \left(C_{ij} - p_j V_i\right) A_{j\mu} V_\mu$$

for $1 \leq i, j \leq d$.

➨ It is convenient to assign/adjoin *projective coordinates*, $v_0 = V_0 = 1$, and we have previously set $z_0 = 0$.

8.3 Observables

The relations

$$X_j = t\,\frac{\partial H}{\partial z_j} + \sum_i \mathcal{R}_i\,\frac{\partial V_i}{\partial z_j}$$

give the observables

$$X_j = \big(t\,p_j + \sum_i \mathcal{R}_i(C_{ij} - p_j\mathcal{V}_i)\big)\,A_{j\mu}\mathcal{V}_\mu$$

where we have identified the V's with the canonical velocity operators, $\mathcal{V}$'s.

9 Coherent states. Berezin transform. Lie algebra

The generating function $e^{V\mathcal{R}}\Omega$ is a type of *coherent state*. The inner product of coherent states has the form

$$\Upsilon = \langle e^{B\mathcal{R}}\Omega, e^{V\mathcal{R}}\Omega\rangle = \phi(B_1V_1, \ldots, B_dV_d)$$

by orthogonality.

With $\mathcal{L}_i$ denoting the adjoint of $\mathcal{R}_i$, we have

$\langle\Omega, e^{B\mathcal{L}}e^{V\mathcal{R}}\Omega\rangle$ equal to the

vacuum expectation value of the group element $e^{B\mathcal{L}}e^{V\mathcal{R}}$.

Comparing with the Heisenberg-Weyl group

$$e^{BD}e^{VX} = e^{VX}e^{BV}e^{BD}$$

we call Υ the **Leibniz function** of the system.

9.1 Recovering the lowering operators

Start with the observations

$$\frac{\partial \Upsilon}{\partial V_i} = \langle e^{B\mathcal{R}}\Omega, \mathcal{R}_i e^{V\mathcal{R}}\Omega\rangle$$

$$\frac{\partial \Upsilon}{\partial B_i} = \langle e^{B\mathcal{R}}\Omega, \mathcal{L}_i e^{V\mathcal{R}}\Omega\rangle \ .$$

Thus we wish to express the partial derivatives $\frac{\partial \Upsilon}{\partial B_i}$ in terms of V_i and $\frac{\partial \Upsilon}{\partial V_i}$.

With the correspondence

$$\frac{\partial \Upsilon}{\partial V_i} \longleftrightarrow \mathcal{R}_i$$

we will have found the lowering operators in terms of the canonical raising and velocity operators.

9.2 Leibniz function for the Krawtchouk system

- We have the coherent state

$$e^{V\mathcal{R}}\Omega = e^{xU(V)-tH(U(V))}$$

Multiplying by $e^{B\mathcal{R}}\Omega$ and taking the expected value produces the exponent t times

$$H(U(B)+U(V))-H(U(B))-H(U(V)) = \psi(BV)$$

say. From the identifications of the Bernoulli constituents found above, via the fundamental relation $A^\top \mathrm{p} A = D$, we find

$$\psi(BV) = \log B_\mu D_\mu V_\mu$$

where $B_0 = V_0 = 1$ and $D_i = D_{ii}$ are the diagonal entries of D.

- Hence, the Leibniz function

$$\Upsilon = (B_\mu D_\mu V_\mu)^t \,.$$

9.3 Lowering operators for the Krawtchouk system. Lie algebra.

- We find the system of partial differential equations

$$\frac{1}{D_i}\frac{\partial \Upsilon}{\partial B_i} = tV_i\Upsilon - V_i\sum V_j\frac{\partial \Upsilon}{\partial V_j}$$

Hence the lowering operators

$$\mathcal{L}_i = D_i\Big(t - \sum_{j=1}^{d}\mathcal{R}_j\mathcal{V}_j\Big)\,\mathcal{V}_i$$

- Introducing the d^2 operators

$$\rho_{ij} = [\mathcal{L}_i, \mathcal{R}_j]$$

plus the $2d$ raising and lowering operators yields a Lie algebra of dimension $d^2 + 2d = (d+1)^2 - 1$. Thus, we have a copy of $s\ell(d+1)$.

10 Observables. Recurrence formulas.

Going back to the observables, we find $X_j =$

$$\sum_{1\le i\le d} (\mathcal{R}_i + \mathcal{L}_i) C_{ij} - p_j \sum_{1\le i\le d} \mathcal{R}_i \mathcal{V}_i + \sum_{\substack{1\le i\le d\\ 1\le k\le d}} C_{ij} A_{jk} \mathcal{R}_i \mathcal{V}_k$$

the first and second terms are manifestly self-adjoint. The last term can be expressed using the operators ρ_{ij}. The relations $\rho^*_{ij} = \rho_{ji}$ provide self-adjointness.

The form of the X_j in terms of the canonical raising and velocity operators

$$X_j = \bigl(t\, p_j + \sum_i \mathcal{R}_i (C_{ij} - p_j \mathcal{V}_i)\bigr)\, A_{j\mu} \mathcal{V}_\mu$$

are, in fact, recurrence formulas for the multivariate Krawtchouk polynomial system.

11 Acknowledgments

➨ We acknowledge the seminal paper of R. C. Griffiths

Orthogonal polynomials and multinomial distributions,

Australian J. Stat. **13**(1971) 27–35.

==

THANKS

Prof. Obata and the Center at Tohoku University

for the invitation.

I am very happy to have the opportunity to support your program as a participant in this

GSIS-RCPAM symposium.

and I am looking forward to visiting with all of you.

==