General Mathematics Vol. 7 /1999

Orthogonal polynomials in nonlinear analysis of regression

Romeo Vomişescu

Abstract

In this note on proved the important role which have orthogonal polynomials of $L_2(\mu)$ in the problem of regression analysis of mathematical statistics.

1991 Mathematical Subject Classification ...

1. Let X be an arbitrary set, $\mathcal{B}$ a σ–algebra of subsets of the set X. Any element of $\mathcal{B}$ is called a measurable set. Denote by $(X, \mathcal{B}, \mu)$ a σ–finite measurable space and by $L_2(\mu)$–the space defined by the equality

$$L_2(\mu) = \{f : \int |f|^2 d\mu < \infty\}$$

where f is equivalence class of the functions with complex valued with respect to measure μ.

For any $f \in L_2(\mu)$ we define norm of f by the equality

$$\|f\|_2 = \left(\int |f|^2 d\mu\right)^{\frac{1}{2}}.$$

We shall assume that the functions $f_1, f_2, \ldots, f_n, f$ belong of the Hilbert Space $L_2(X, \mathcal{B}, \mu)$ and H is vectorial subspace of $L_2(X, \mathcal{B}, \mu)$ generate by $1, f_1, \ldots, f_n$. Because H is closed result from general theory of Hilbert space that the orthogonal projection $pr_H f$ of the function f on H have the property that

$$\inf_{g \in H} \|g - f\|_2 = pr_H f,$$

namely $pr_H f$ is the best linear approximation in $L_2(X, \mathcal{B}, \mu)$ of the function f by the functions $1, f_1, \ldots, f_n$.

In applications, $f_1, f_2, \ldots, f_n$ are in general random variables which belongs in times past therefore it is known, and f is a random variable which depend for the future and she must express by the function $f_1, \ldots, f_n$. We now formulate the converse statement

Theorem 1 *Let $f_k : (X, \mathcal{B}, \mu) \to \mathbb{R}$, $k = \overline{1, n}$ be real random variables ($RRV's$) square – integrable on $(X, \mathcal{B})$ with respect to measure μ and let H be vectorial subspace of $L_2(X, \mathcal{B}, \mu)$ generate by the functions $1, f_1, f_2, \ldots, f_n$. For any $f \in L_2$, the orthogonal projection $pr_H f$ which is the best linear approximation of f by $1, f_1, \ldots, f_n$ in the sens of norm L_2, have the expression*

$$pr_H f = Mf + \sum_{k=1}^{n} a_k f_k \tag{1}$$

where M is mean value, and the vector $a^\top = (a_1, a_2, \ldots, a_n)$ is a arbitrary solution of system

$$\sum_{k=1}^{n} a_k \cdot cov\,(f_k, f_m) = cov\,(f, f_m) \quad , \quad 1 \le m \le n.$$

The random variables $(f_k)_{1 \le k \le m}$ which belongs of the vectorial spaces L_2 is said to be linearly independent if and only if $C_{(f_1, \cdots, f_n)}$ is inversable. We have at the time

$$a_k = \sum_{m=1}^{n} cov\,(f, f_m) C^{-1}_{(f_1, \cdots, f_n)}(m, k) \quad , \quad 1 \le k \le n$$

and

$$M[(f - pr_H f)^2] =$$
$$= D^2(f) - \sum_{k,m=1}^{n} \cdot cov\,(f, f_k) C^{-1}_{(f_1, \cdots, f_n)}(m, k) \cdot cov\,(f_m, f).$$

The equality (1) is called linear regression of f by $f_1, \ldots, f_n$.

2. The orthogonal polynomials from $L_2(\mu)$ have an important role in the problems of nonlinear analysis of regression where the distribution μ on $\mathbb{R}$ fulfiled the condition that is regular, which is satisfied of the most distributions of statistics. It is well meaning following

Theorem 2 *If μ is a measure of probability on $\mathbb{R}$ so that $(\forall)a > 0$,*

(2) $$\int e^{a|x|} d\mu(x) < \infty$$

then all the moments of measure μ exist, and the sequence of orthogonal polynomials is complete in $L_2(\mu)$.

3. Let f, g be two real random variables with common repartition P on $\mathbb{R}^2$, so that $Mf^2 < \infty$, $M|f|^n < \infty$, $n = 0, 1, 2, \ldots$.

Let $\mu = Pf^{-1}$ be the distribution of the random variable f on $\mathbb{R}$, and let $\{p_n\}_{n \geq 0}$ be the sequence of the orthogonal polynomials in $L_2(\mu)$. We will have the best approximation of g by $\{1, f, f^2, \ldots, f^n\}$, by projection of g on the linear covering $\{1, f, f^2, \ldots, f^n\}$ in $L_2(P)$. The linear cover $\{1, f, f^2, \ldots, f^n\}$ is the same time the linear cover $p_0(f) = 1, p_1(f), \ldots, p_n(f)$.

Therefore the linear regression of g by $1, f, f^2, \ldots, f^n$ is equal with

$$\sum_{k=0}^{n} [Mgp_k(f)]p_k(f)$$

and is called the polynomial regression of g by f of degree n.

4. Let μ be the standard normal distribution in $\mathbb{R}$, namely

$$\mu(E) = \frac{1}{\sqrt{2\pi}} \int_E e^{-\frac{1}{2}x^2} dx \quad , \quad E \subset \mathbb{R}.$$

Let $\{H_n(x)\}_{n \geq 0}$ be the sequence of the polynomials defined by the equality

(3) $$K_N(t,x) = e^{-\frac{1}{2}(t^2-2tx)} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x) \quad , \quad t \in \mathbb{R}.$$

We will show that for any $t, s \in \mathbb{R}$ we have

$$\int K_N(t,x) K_N(s,x) d\mu(x) = e^{ts}.$$

Indeed,

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{t^n}{n!} \cdot \frac{s^k}{k!} H_n(x) H_k(x) =$$

$$e^{-\frac{1}{2}(t^2-2xt+s^2-2sx)} = e^{-\frac{1}{2}(t^2+s^2)} e^{(t+s)x}$$

and

$$\int K_N(t,x) K_N(s,x) d\mu(x) =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{t^n}{n!} \cdot \frac{s^k}{k!} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} H_n(x) H_k(x) dx =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t^2+s^2)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2+(t+s)} dx.$$

We make the substitute $x = \sqrt{2} \cdot z$ and we take into account that

$$\int_{-\infty}^{\infty} e^{-z^2+2bz} dz = \sqrt{\pi} \cdot e^{b^2}.$$

Therefore

$$\int K_N(t,x) K_N(s,x) d\mu(x) =$$

$$\frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}(t^2+s^2)} \sqrt{\pi} e^{\frac{1}{2}(t^2+s^2+2st)} = e^{st}$$

Identifyed the coefficients of $t^k s^\ell$, we obtain

$$\int \frac{H_k(x)}{\sqrt{k!}} \cdot \frac{H_\ell(x)}{\sqrt{\ell!}} d\mu(x) = \delta_{k\ell} \quad , \quad k, \ell = 0, 1, 2, \ldots$$

where $\delta_{k\ell}$ is Kronecker's $\delta-$ symbol.

Since H_n is polynomials of degree n result from Theorem 2, that the sequence $\left\{\dfrac{H_n(x)}{\sqrt{n!}}\right\}_{n\geq 0}$ is a sequence complete orthonormal of orthogonal polynomials for the standard nornal distribution.

H_n are called the Hermite polynomials of degee n. We next apply the equality (3) to obtain

$$e^{-\frac{1}{2}(x-t)^2}=\sum_{n=0}^{\infty}\frac{t^n}{n!}H_n(x)e^{-\frac{1}{2}x^2}.$$

Using Taylor$'$s series, we have

$$e^{-\frac{1}{2}(x-t)^2}=\sum_{n=0}^{\infty}(-1)^n\frac{t^n}{n!}\cdot\frac{d^n}{dx^n}(e^{-\frac{1}{2}x^2}).$$

By comparing the coefficient of t^n on both sides, we obtain

$$H_n(x)=(-1)^n e^{\frac{1}{2}x^2}\frac{d^n}{dx^n}(e^{-\frac{1}{2}x^2}).$$

Using the orthonormal sequence above, without proof, we state the following theorem

Theorem 3 *Let f,g be the random variables which have the bidimensional normal distributions with the mean values the null vector and the covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Let $\xi(f)$ and $\nu(g)$ two random variables which are functions only of f respectively of g. We suppose that*

$$M\xi(f)=M\nu(g)=0 \quad , \quad M\xi^2(f)=M\nu^2(g)=1.$$

Then $|cov\,(\xi(f),\nu(g))| \leq |\rho|$. The equality take place if and only if $\xi(f)=\pm f$ and $\nu(g)=\pm g$.
If f,g are real random variables and f have normal distribution with the mean value m and the variance σ^2, then the

polynomial regression of g by f of degree n is

$$\sum_{k=0}^{n} \frac{1}{k!} \left\{ M \left[g H_k \left(\frac{f-m}{\sigma} \right) \right] \right\} H_k \left(\frac{f-m}{\sigma} \right)$$

where H_k is Hermite's polynomial of degree k.

5. Let μ be the gamma distribution on $(0,\infty)$, i.e.,

$$\mu(E) = \frac{1}{\Gamma(\alpha)} \int_E x^{\alpha-1} e^{-x} dx \quad , \quad E \subset \mathbb{R}_+$$

where $\alpha > 0$. Let $\{L_n(x)\}_{n\geq 0}$ be the sequence of polynomials define by

$$\begin{aligned} K_G(t,x) &= \frac{1}{(1-t)^\alpha} e^{-\frac{tx}{1-t}} = \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} L_n(x) \quad , \quad t \in (-1,1). \end{aligned} \tag{4}$$

Afterwards,

$$\int K_G(t,x) K_G(s,x) d\mu(x) =$$

$$= \int_0^\infty \frac{1}{(1-t)^\alpha} e^{-\frac{tx}{1-t}} \frac{1}{(1-s)^\alpha} e^{-\frac{sx}{1-s}} \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} dx =$$

$$= \frac{1}{\Gamma(\alpha)} \cdot \frac{1}{(1-t)^\alpha} \cdot \frac{1}{(1-s)^\alpha} \int_0^\infty e^{-\left(\frac{t}{1-t}+\frac{s}{1-s}+1\right)x} x^{\alpha-1} dx =$$

$$= \frac{1}{\Gamma(\alpha)} \cdot \frac{1}{(1-t)^\alpha} \cdot \frac{1}{(1-s)^\alpha} \int_0^\infty x^{\alpha-1} e^{-\frac{1-ts}{(1-t)(1-s)}x} dx$$

Substituting

$$\frac{1-ts}{(1-t)(1-s)} x = u \quad , \quad dx = \frac{(1-t)(1-s)}{1-ts} du,$$

we have

$$\int K_G(t,x) K_G(s,x) d\mu(x) =$$

$$= \frac{\Gamma(\alpha)}{\Gamma(\alpha)(1-ts)^{\alpha}} = \frac{1}{(1-ts)^{\alpha}} = (1-ts)^{-\alpha},$$

therefore

$$\int K_G(t,x)K_G(s,x)d\mu(x) =$$

$$= \frac{1}{(1-ts)^{\alpha}} = \sum_{n=0}^{\infty} \frac{\alpha(\alpha+1)\dots(\alpha+n-1)}{n!}t^n s^n.$$

By comparing the coefficients of $t^k s^\ell$ on both sides, we obtain

$$\int p_k(x)p_\ell(x)d\mu(x) = \delta_{k\ell} \quad , \quad k,\ell = 0,1,2,\dots$$

where

$$(5)\, p_k(x) = \left[\frac{1}{k!\alpha(\alpha+1)\dots(\alpha+k-1)}\right]^{\frac{1}{2}} L_k(x) \ , \ k = 0,1,2,\dots$$

It is verified that

$$p_k(x) = c_k \sum_{r=0}^{k} \binom{k}{r}(-1)^r \frac{x^r}{\Gamma(r+\alpha)}$$

where c_k is the normed constant.

The polynomials L_n are called the Laguerre polynomials of degree n. Since $\int e^{tx}d\mu < \infty$ $(\forall)t \in (-1,1)$, from the Theorem 2 result that the sequence (p_k) defined of (5) is a complete orthonormal sequence in $L_2(\mu)$.

6. Let μ be the Poisson distribution with the parameter $\lambda > 0$, that is

$$\mu(\{x\}) = e^{-\lambda}\frac{\lambda^x}{x!} \quad , \quad x = 0,1,2,\dots \ .$$

Let $k_P(t,x) = e^{-t}(1+\lambda^{-1}t)^x$ be, where $t \in \mathbb{R}$. Then

$$\int k_P(t,x)k_P(s,x)d\mu(x) =$$

$$= \sum_{x=0}^{\infty} e^{-t}(1+\lambda^{-1}t)^x e^{-s}(1+\lambda^{-1}s)^x e^{-\lambda}\frac{\lambda^x}{x!} =$$
$$= e^{-t}e^{-s}e^{-\lambda}\sum_{x=0}^{\infty}[(1+\lambda^{-1}t)(1+\lambda^{-1}s)]^x\frac{\lambda^x}{x!} =$$
$$e^{-t}e^{-s}e^{-\lambda}\sum_{x=0}^{\infty}\frac{[(1+\lambda^{-1}t)(1+\lambda^{-1}s)]^x}{x!} =$$
$$e^{-t}e^{-s}e^{-\lambda}e^{(1+\lambda^{-1}t)(1+\lambda^{-1}s)\lambda} =$$
$$e^{-t}e^{-s}e^{-\lambda}e^{\lambda(1+\lambda^{-1}t+\lambda^{-1}s+\lambda^{-2}ts)} =$$
$$e^{-t}e^{-s}e^{-\lambda}e^{\lambda}e^{t}e^{s}e^{\frac{ts}{\lambda}} = e^{\frac{ts}{\lambda}},$$

that is

$$\int k_P(t,x)k_P(s,x)d\mu(x) = e^{\frac{ts}{\lambda}}.$$

Let $(P_n(x))_{n\geq 0}$ be the polynomials sequence defined by the formula

$$k_P(t,x) = e^{-t}(1+\lambda^{-1}t)^x = \sum_{n=0}^{\infty}\frac{t^n}{n!}P_n(x). \tag{6}$$

But

$$e^{-t}(1+\lambda^{-1}t)^x = \sum_{k=0}^{\infty}\sum_{r=0}^{\infty}\binom{x}{r}\lambda^{-r}t^r\frac{(-1)^k t^k}{k!} =$$
$$= \sum_{k=0}^{\infty}\sum_{r=0}^{\infty}\binom{x}{r}(-1)^{n-r}\lambda^{-r}\frac{t^r t^{n-r}}{(n-r)!} =$$
$$= \sum_{k=0}^{\infty}\sum_{r=0}^{\infty}\binom{n}{r}(-1)^{n-r}\lambda^{-r}x(x-1)\ldots(x-r+1)\frac{t^n}{n!}.$$

Supposing

$$p_n(x) = \frac{\lambda^{\frac{n}{2}}}{\sqrt{n!}}P_n(x) \quad , \quad n = 0,1,2,\ldots$$

then $(p_n)_{n\geq 0}$ is a complete orthonormal sequence of polynomials in $L_2(\mu)$. The polynomial P_n is called the Poisson$'$ s poynomial of degree n.

7. Let μ be the binomial distribution defined on $\{0,1,2,\ldots,n\}$ set by the equality

$$\mu(\{x\}) = \binom{n}{x} p^x q^{n-x} \quad , \quad x = 0,1,2,\ldots,n.$$

We note $K_B(t,x) = (1+tq)^x(1-tp)^{n-x}$. Then

$$\int K_B(t,x)K_B(s,x)d\mu(x) =$$

$$= \sum_{x=0}^{n} (1+tq)^x(1-tp)^{n-x}(1+sq)^x(1-sp)^{n-x}\binom{n}{x}p^x q^{n-x} =$$

$$= \sum_{x=0}^{n} \binom{n}{x}(p+spq+tpq+tspq^2)^x(q-spq-tpq+tsp^2q)^{n-x} =$$

$$= (p+spq+tpq+tspq^2+q-spq-tpq+tsp^2q)^n = (1+tspq)^n$$

Therefore

$$\int K_B(t,x)K_B(s,x)d\mu(x) = (1+pqts)^n.$$

Moreover

$$K_B(t,x) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\left[\sum_{j=0}^{k}(-1)^j\binom{k}{j}p^j q^{k-j} x(x-1)\ldots\right.$$

$$\left.\ldots(x-k+j-1)(n-x)\ldots(n-x-j+1)\right] =$$

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

The sequence (p_k) defined by the equality

$$p_k(x) = \frac{(pq)^{-\frac{k}{2}}\binom{n}{k}^{-\frac{1}{2}}}{k!}K_k(x) \quad , \quad k = 0,1,2,\ldots,n$$

determine an orthonormal base in $L_2(\mu)$. The polynomials K_k are called the Kravciuc′s polynomials of degree k. Maybe show that

$$\lim_{n\to\infty} K_B\left(\frac{t}{\sqrt{npq}}, np + x\sqrt{npq}\right) = K_N(t,x)$$

$$\lim_{\substack{n\to\infty\\ np\to\lambda}} K_B(t\lambda^{-1}, x) = K_P(t,x).$$

Bibliografie

[1] Ciucu G., Tudor C., *Probabilităţi şi procese stocastice*, Ed. Academiei, Bucureşti, 1978.

[2] Mihoc Gh., Ciucu G., Craiu V., *Teoria probabilităţilor şi statistică matematică*, E.D.P., Bucureşti, 1970.

[3] Larry C. Andrews, *Special functions*, Macmillan Publ. Company, New York, 1985.

[4] Rainville E.D., *Special Functions*, New York: Chelsea, 1960.

[5] Şiriaev A.N., *Veroiatnosti*, Moscova, 1989.

[6] Vomişescu R., *B–Stieltjes Integrals with Probability Interpretations*, Acta Technica Napocensis, nr. 37, I, pg. 107.

Department of Mathematics
Faculty of Sciences
University "Lucian Blaga" of Sibiu
2400 – Sibiu Romania