Regresja form kwadratowych i algebry Jordana

Jacek Wesołowski

Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska, Warszawa

Konferencja Statystyki Matematycznej
Wisła 2009

współautor: G. Letac (Tuluza, Francja)

Plan

Klasyczne twierdzenie Lukacsa (1955)

Rozkład gamma $\gamma_{s,b}$ dla $s, b > 0$ ma gęstość

$$f(x) = \frac{x^{s-1} e^{-x/b}}{b^s \Gamma(s)} I_{(0,\infty)}(x).$$

Twierdzenie. *Niech $X > 0$ i $Y > 0$ będą niezależne i niezdegenerowane. Jeśli*

$$U = \frac{X}{X+Y} \quad \textit{oraz} \quad V = X + Y$$

są niezależne, to istnieją stałe $p, q, a > 0$ takie że

$$X \sim \gamma_{p,a} \quad \textit{oraz} \quad Y \sim \gamma_{q,a}.$$

Macierzowy rozkład Wisharta

Niech $\bar{\Pi}_r(\mathbb{R})$ - stożek nieujemnie określonych macierzy rzeczywistych (symetrycznych) $r \times r$.

Rozkład Wisharta $\gamma_{p,\sigma}$ definiuje się przez transformatę Laplace'a

$$\int_{\bar{\Pi}_r(\mathbb{R})} e^{-\operatorname{tr}\theta x}\,\gamma_{p,\sigma}(dx) = \frac{1}{[\det(I_r + \theta\sigma)]^p}$$

jeśli $\sigma \in \Pi_r(\mathbb{R})$ (macierze dodatnio określone) oraz

$$p \in \Lambda = \left\{0, \frac{1}{2}, \frac{2}{2}, \ldots, \frac{r-1}{2}\right\} \cup \left(\frac{r-1}{2}, \infty\right).$$

Mówi o tym tw. Gindykina (1975).

Macierzowy rozkład Wisharta, cd

Jeśli $p > \frac{r-1}{2}$ to rozkład $\gamma_{p,\sigma}$ skupiony jest na $\Pi_r(\mathbb{R})$ i ma gęstość

$$f(x) = \frac{[\det(x)]^{p-\frac{r+1}{2}} e^{-\mathrm{tr}\,(x\sigma^{-1})}}{\Gamma_r(p)[\det(\sigma)]^p} I_{\Pi_r(\mathbb{R})}(x).$$

Jeśli $Z_1, \ldots, Z_N$ są niezależne i mają rozkład normalny w $\mathbb{R}^n$

$$Z_i \sim \mathcal{N}(0, \Sigma), \quad i = 1, \ldots, N,$$

to

$$Z_1 Z_1^T + \ldots + Z_N Z_N^T \sim \gamma_{\frac{N}{2}, 2\Sigma} \qquad \text{Wishart (1928).}$$

Twierdzenie Olkina-Rubina (1962)

Theorem

Niech X i Y będą niezależnymi macierzami losowymi o wartościach w $\bar{\Pi}_r(\mathbb{R})$ oraz $X+Y \in \Pi_r(\mathbb{R})$. Jeśli

$$U = (X+Y)^{-1} X (X+Y)^{-1} \quad i \quad V = X+Y \quad \textit{niezależne},$$

oraz

$$U \stackrel{d}{=} OUO^T \quad \forall\, O \ \textit{ ortogonalnej } r \times r$$

to istnieją $p, q \in \Lambda$ i $\sigma \in \Pi_r(\mathbb{R})$ takie, że

$$X \sim \gamma_{p,\sigma} \quad \textit{oraz} \quad Y \sim \gamma_{q,\sigma}.$$

Uogólnienia

1. Casalis, Letac (1996) - na stożki symetryczne w algebrach Jordana
2. Bobecka, Wesołowski (2002) - dla macierzy bez niezmienniczości rozkładu U, ale z gładkimi gęstościami
3. Hassairi, Lajmi, Zine (2008) - BW z algorytmem dzielenia definiowanym przez macierze trójkątne (rozkład Wisharta-Riesza)
4. Boutoria (2009) - BW na tzw. stożkach jednorodnych

Wersje regresyjne

Laha, Lukacs (1960):
Niech $X > 0$ i $Y > 0$ niezależne i niezdegenerowane. Jeśli istnieją stałe a i b takie, że

$$\mathbb{E}(X|X+Y) = a(X+Y) \quad \text{oraz} \quad \mathbb{E}(X^2|X+Y) = b(X+Y)^2$$

to $X \sim \gamma_{p,c}$ oraz $Y \sim \gamma_{q,c}$.

Wang (1981): Niech $X = (X_1, X_2)$ i $Y = (Y_1, Y_2)$ niezależne o wartościach w $(0, \infty)^2$. Jeśli

$$\mathbb{E}(X_i|X+Y) = a(X_i+Y_i) \quad \text{oraz} \quad \mathbb{E}(X_i^2|X+Y) = b(X_i+Y_i)^2$$

to X i Y mają dwuwymiarowe rozkłady gamma o transformacie Laplace'a postaci

$$\phi(s_1, s_2) = (1 - a_1 s_1 - a_2 s_2 + (a_1 a_2 - a_{12}) s_1 s_2)^{-p}.$$

Regresja form kwadratowych

Letac i Wesołowski (2008): Niech v - forma kwadratowa w $\mathbb{R}^n$ o macierzy M_v rzędu r. Niech Q_v przestrzeń form kwadratowych q na $\mathbb{R}^n$ ortogonalnych do v, tzn. $\langle q, v\rangle = \mathrm{tr}\, M_q M_v = 0$. Niech X i Y niezależne o wartościach w $\mathbb{R}^n$ i wykładniczych momentach. Jeśli $r \geq 2$,

$$\mathbb{E}(X|X+Y) = a(X+Y) \quad \text{i} \quad \forall\, q \in Q_v \;\; \mathbb{E}(q(X)|X+Y) = bq(X+Y).$$

dla $0 < a < 1$ i $b^2 < a < b$, to istnieją $c \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ takie, że

$$\mathbb{E}\, e^{\langle s,X\rangle} = (1 - \langle c, s\rangle + \lambda v(s))^{-p} = \left(\mathbb{E}\, e^{\langle s,Y\rangle}\right)^{\frac{a}{1-a}}$$

dla $p = \frac{a(a-b)}{b-a^2}$.

Algebra Jordana

Przestrzeń liniową V nad ciałem K ($\mathbb{R}$ lub $\mathbb{C}$) z dwuliniowym przekształceniem $V \times V \ni (x, y) \mapsto x \circ y \in V$ (produktem Jordana) jest algebrą Jordana jeśli

$$x \circ y = y \circ x \quad \text{oraz} \quad x \circ (x^{\circ 2} \circ y) = x^{\circ 2} \circ (x \circ y).$$

Przykłady:
(I) $V = S(m, \mathbb{R})$ - rzeczywiste macierze symetryczne $m \times m$

$$x \circ y = \frac{1}{2}(xy + yx).$$

(II) $V = \mathbb{R} \times E$, gdzie E przestrzeń euklidesowa,

$$(x_0, x) \circ (y_0, y) = (x_0 y_0 + \langle x, y \rangle, x_0 y + y_0 x).$$

Algebra Jordana, cd

Dla V skończenie wymiarowej definiuje się $\det(x)$ i $\mathrm{tr}\,(x)$ jako współczyniki wielomianu minimalnego elementu $x \in V$, oraz $\mathrm{rank}\, V$ jako maksymalny stopień wielomianów minimalnych.

Przykłady:

(I) $V = S(m, \mathbb{R}) \;\Rightarrow$

$$\det(x) = Det(x), \quad \mathrm{tr}\,(x) = Tr(x), \quad \mathrm{rank}\, V = m.$$

(II) $V = \mathbb{R} \times E \;\Rightarrow$

$$\det(x_0, x) = x_0^2 - ||x||^2, \quad \mathrm{tr}\,(x) = x_0^2 + ||x||^2 \quad \mathrm{rank}\, V = 1 + \dim E.$$

Stożek symetryczny algebry Jordana

Niech V będzie przestrzenią euklidesową z iloczynem skalarnym $\langle\cdot,\cdot\rangle$. Algebrę Jordana $(V,\circ)$ nazywamy euklidesową jeśli

$$\langle x\circ u, v\rangle = \langle u, x\circ v\rangle \quad \forall\, x,u,v\in V.$$

Wtedy

$$\Omega = \mathrm{Int}\{x\circ x:\ x\in V\}$$

jest stożkiem symetrycznym, tzn.
(1)

$$\Omega = \Omega^* = \{y\in V:\ \langle x,y\rangle > 0\,\forall\, x\in\bar{\Omega}\setminus\{0\}\}$$

(2) grupa $G(\Omega)$ automorfizmów Ω działa tranzytywnie, tzn.

$$\forall\, x,y\in\Omega\ \ \exists\, g\in G(\Omega):\ g(x)=y.$$

Algebry Jordana i ich stożki symetryczne

V	Ω	$\dim V$	$\operatorname{rank} V$	d
$S(m, \mathbb{R})$	$\Pi_m(\mathbb{R})$	$\frac{m(m+1)}{2}$	m	1
$H(m, \mathbb{C})$	$\Pi_m(\mathbb{C})$	m^2	m	2
$H(m, \mathbb{H})$	$\Pi_m(\mathbb{H})$	$m(2m-1)$	m	4
$\mathbb{R} \times \mathbb{R}^{m-1}$	Λ_n	n	2	$n-1$
$H(3, \mathbb{O})$	$\Pi_3(\mathbb{O})$	27	3	8

Rozkład Wisharta na stożku symetrycznym

Rozkład Wisharta $\gamma_{p,\sigma}$ na stożku symetrycznym $\bar{\Omega}$ w algebrze Jordana rzędu r definiuje się przez transformatę Laplace'a

$$\int_{\bar{\Omega}} e^{-\mathrm{tr}\,\theta x}\,\gamma_{p,\sigma}(dx) = [\det(e+\theta\sigma)]^{-p}$$

dla $\sigma \in \Omega$ oraz p w zbiorze Gindykina

$$p \in \Lambda_V = \left\{\frac{d}{2}j,\ j = 0, 1, \ldots, r-1\right\} \cup \left(\frac{d}{2}(r-1), \infty\right).$$

Jeśli $p > \frac{d}{2}(r-1)$ to $\gamma_{p,\sigma}$ jest skupiona na stożku otwartym Ω i ma gęstość względem miary Lebesgue'a na V postaci

$$f(x) = \frac{[\det x]^{p-\frac{d}{2}(r+1)} e^{-\mathrm{tr}\,x\sigma^{-1}}}{\Gamma_\Omega(p)[\det \sigma]^p} I_\Omega(x).$$

Rozkład Wisharta na stożkach symetrycznych, cd

Rozkład Wisharta $\gamma_{p,\sigma}$ na stożku rzeczywistych macierzy symetrycznych, $S(m,\mathbb{R})$, to standardowy obiekt klasycznej statystycznej analizy wielowymiarowej dla $p = N/2$ znany od pracy Wisharta (1928). Dla pozostałych p od pracy Olkin, Rubin (1962).

Rozkład Wisharta na zespolonych macierzach hermitowskich, $H(m,\mathbb{C})$, badany jest od pracy Goodmana (1963) o zespolonych wektorach gaussowskich. W szczególności fizycy, Mehta (2004), w tym asymptotyka miar empirycznych zadanych przez wartości własne - rozkład Marchenko-Pastura.

Rozkład Wisharta na stożkach symetrycznych, cd

Rozkład Wisharta na kwaternionowych macierzach hermitowskich $H(m, \mathbb{H})$ badany jest od pracy Anderssona (1975).

Rozkład Wisharta na stożku Lorentza badany jest od pracy Jensena (1988).

Unifikacja w ramach teorii algebr Jordana i stożków symetrycznych: Letac (1989), Massam (1994), Faraut, Koranyi (1994).

Uogólnienie na tzw. stożki jednorodne: Andersson, Wojnar (2006) - rozkład hiper-Wisharta (Lauritzen), ważny w sieciach bayesowskich (*graphical models*).

Reprezentacja kwadratowa

Dla $x \in V$ definiuje się operacje liniową $\mathbb{P}(x)$na V

$$\mathbb{P}(x)y = 2x \circ (x \circ y) - (x \circ x) \circ y.$$

$\mathbb{P}$ jest tzw. reprezentacją kwadratową w algebrze Jordana V.
Np. dla $V = S(m, \mathbb{R})$ mamy

$$\mathbb{P}(x)y = xyx.$$

Jeśli κ jest transformatą kumulantową dla rozkładu Wisharta na $\bar{\Omega}$, czyli $\kappa = \log L$ gdzie L jest transformatą Laplace'a, to

$$\kappa''(\theta) = \mathbb{P}(\kappa'(\theta)).$$

Równość ta charakteryzuje rozkład Wisharta na stożkach symetrycznych - Letac(1989).

Charakteryzacja regresyjna rozkładu Wisharta w $\bar{\Omega}$

Przypomnienie: dla $x \in V$ (przestrzeń euklidesowa) endomorfizm $x \otimes x : V \to V$ definiuje się wzorem

$$(x \otimes x)y = \langle x, y \rangle x.$$

Letac, Massam (1998): Dla dowolnych $x \in V$ niech Q_1 i Q_2 będą symetrycznymi endomorfizmami na V postaci

$$Q_1(x) = \mathbb{P}(x) + \frac{d}{2} x \otimes x \quad \text{i} \quad Q_2(x) = \mathbb{P}(x) - x \otimes x.$$

Niech X i Y niezależne o wartościach w $\bar{\Omega}$. Niech istnieją $a, b_1, b_2 \in \mathbb{R}$ takie, że

$$\mathbb{E}(X|X+Y) = a(X+Y) \quad \text{i} \quad \mathbb{E}(Q_i(X)|X+Y) = b_i Q_i(X+Y), \; i = 1, 2.$$

Wtedy X i Y mają rozkład Wisharta na Ω (o wspólnej skali σ).

Regresje form kwadratowych w algebrach Jordana

Niech $L_s(V)$ oznacza przestrzeń symetrycznych endomorfizmów przestrzeni V. Wtedy dla dowolnego $S \in L_s(V)$ definiujemy formę kwadratową $q_i(S)$ na V

$$q_i(S)(x) = \mathrm{trace}\, SQ_i(x), \quad i = 1, 2.$$

Niech

$$\mathcal{Q}_i = \{q_i(S) : \ S \in L_s(V)\}, \quad i = 1, 2.$$

Wtedy warunki regresyjne z tw. Letac-Massam można przepisać w postaci

$$\mathbb{E}(q(X)|X + Y) = b_i q(X + Y), \quad \forall\, q \in \mathcal{Q}_i, \ i = 1, 2.$$

Regresje form kwadratowych w przestrzeni liniowej

Letac, Wesołowski (2009), Bull. Soc. Math. France:

Tw. *Przestrzeń form kwadratowych $\mathcal{Q}$ na skończenie wymiarowej przestrzeni liniowej V ma dekompozycję na sumę prostą*

$$\mathcal{Q} = \mathcal{Q}_1 \oplus \ldots \oplus \mathcal{Q}_k.$$

Niech X i Y niezależne o wartościach w V i o rozkładach nieredukowalnych. Istnieją $a \in \mathbb{R}$ taka, że

$$\mathbb{E}(X|X+Y) = a(X+Y)$$

oraz różne liczby $b_1, \ldots, b_k$ takie, że

$$\mathbb{E}(q(X)|X+Y) = b_i q(X+Y), \quad \forall\, q \in \mathcal{Q}_i,\ i = 1, \ldots, k.$$

Regresje form kwadratowych w przestrzeni liniowej, cd

Tw. cd *Wtedy* $0 < a < 1$, $k = 2$ *oraz na* V *istnieje struktura euklidesowej algebry Jordana taka, że* X *i* Y *mają rozkłady Wisharta* $\gamma_{p,\sigma}$ *i* $\gamma_{p',\sigma}$ *na stożku symetrycznym* $\bar{\Omega}$ *w tej algebrze. Przestrzenie* $\mathcal{Q}_1$ *i* $\mathcal{Q}_2$ *są rozpinane odpowiednio przez*

$$q_1^{(s)}(x) = \mathrm{tr}\,(\mathbb{P}(x)(s)s) + \frac{d}{2}\mathrm{tr}^{\,2}(xs), \quad s \in V,$$

$$q_2^{(s)}(x) = \mathrm{tr}\,(\mathbb{P}(x)(s)s) - \mathrm{tr}^{\,2}(xs), \quad s \in V.$$

Co więcej

$$a = \frac{p}{p+p'}, \quad b_1 = \frac{p(p+1)}{(p+p')(p+p'+1)}, \quad b_2 = \frac{p(p-\frac{d}{2})}{(p+p')(p+p'-\frac{d}{2})}.$$

Uwaga o dowodzie

Ważnym elementem dowodu jest wynik Casalis (1991), który mówi, że jeśli funkcja wariancji V naturalnej rodziny wykładniczej określonej na skończenie wymiarowej przestrzeni liniowej jest jednorodna funkcją kwadratową, to na tej przestrzeni istnieje struktura algebry Jordana i rodzina wykładnicza generowana jest przez rozkład Wisharta na stożku symetrycznym w tej algebrze.

Nasz dowód polega na sprowadzeniu warunków regresji dla form kwadratowych do warunku rozważanego przez Casalis, a następnie na wykazaniu, że to już implikuje $k = 2$, a następnie na identyfikacji przestrzeni $\mathcal{Q}_1$ i $\mathcal{Q}_2$,

Regresje form kwadratowych a struktura algebry Jordana

Warunki regresyjne wprowadzają w sposób jednoznaczny strukturę Jordana w przestrzeń liniową V o wymiarze n. Liczby a, b_1 i b_2 spełniają $b_2 < a^2 < b_1 < a$ i wyznaczają stałą Pierce'a

$$d = 2\frac{(a - b_1)(a^2 - b_2)}{(b_1 - a^2)(a - b_2)}.$$

Rząd algebry Jordana wyznaczamy z identyczności $n = r + \frac{d}{2}r(r-1)$.

Konieczne jest (naturalne) założenie nieredukowalności rozkładu X (lub Y). Miara μ na V jest redukowalna jeśli istnieje dekompozycja $V = V_1 \oplus V_2$ oraz miary μ_i na V_i, $i = 1, 2$ takie, że $\mu = \mu_1 \otimes \mu_2$. Miara jest nieredukowalna jeśli nie jest redukowalna.

Struktura przestrzeni $\mathcal{Q}_1$ i $\mathcal{Q}_2$

Wygodniej badać podprzestrzenie $\mathcal{F}_1$ i $\mathcal{F}_2$ przestrzeni endomorfizmów symetrycznych $\mathcal{F} = L_s(V)$. Izomorfizm $\mathcal{Q} \ni q \leftrightarrow f \in \mathcal{F}$ definiuje się przez $q \mapsto \langle fx, x \rangle$.

Tw.

1. *Istnieje jedyny endomorfizm symetryczny* Ψ *na* $\mathcal{F}$*, tzn.* $\Psi \in L_s(\mathcal{F})$*, taki, że*

$$\Psi(y \otimes y) = \mathbb{P}(y), \quad \forall\, y \in V.$$

2. Ψ *spełnia równanie*

$$\Psi(\mathbb{P}(y)) = \frac{d}{2}\, y \otimes y + \left(1 - \frac{d}{2}\right) \mathbb{P}(y), \quad \forall\, y \in V.$$

Struktura przestrzeni $\mathcal{Q}_1$ i $\mathcal{Q}_2$, cd

3. *Przestrzenie $\mathcal{F}_1$ i $\mathcal{F}_2$ są ortogonalne oraz $\mathcal{F} = \mathcal{F}_1 \oplus \mathcal{F}_2$.*
4. *Przestrzenie $\mathcal{F}_1$ i $\mathcal{F}_2$ są przestrzeniami własnymi endomorfizmu Ψ związanymi z wartościami własnymi $\lambda_1 = 1$ oraz $\lambda_2 = \frac{d}{2}$, odpowiednio.*
5. *Wymiary przestrzeni $\mathcal{F}_1$ i $\mathcal{F}_2$ wynoszą*

$$\dim \mathcal{F}_1 + \dim \mathcal{F}_2 = \frac{n(n+1)}{2}$$

$$\dim \mathcal{F}_2 = \frac{r(r-1)}{2} \frac{4 + 2d(2r-3) + d^2(r-1)(r-2)}{2(2+d)}.$$