Jerzy Boryczka

DWUWYMIAROWY ROZKŁAD PRAWDOPODOBIEŃSTWA ELEMENTÓW I ZJAWISK METEOROLOGICZNYCH

WSTĘP

Współzależność dwóch parametrów meteorologicznych x, y określa się zwykle dwoma liniami regresji, które wyznacza się według metody najmniejszych kwadratów, traktując zawsze jedną ze zmiennych jako zależną.

Fakt istnienia dwóch prostych regresji, z których jedna określa zależność elementu y od x, a druga x od y jest bardzo niewygodny, zwłaszcza przy interpolacji i ekstrapolacji. Biorąc pod uwagę dużą dowolność przyporządkowania osi współrzędnych zmiennym x, y, graficzne ujęcie zależności między parametrami meteorologicznymi nie zawsze jest jednoznaczne.

Dlatego też, lepiej jest zamiast dwóch prostych regresji wyznaczać jedną *prostą główną*

$$ax+by+c=0$$

mającą tę własność, że suma kwadratów odległości punktów empirycznych (x_i, y_i) od niej jest najmniejsza. W odróżnieniu od prostych regresji, istnieje tylko jedna prosta główna, która spełnia ten warunek. Tutaj nie jest istotne, który z parametrów x, y będziemy uważać jako zmienną zależną. Każdej wartości zmiennej x odpowiada dokładnie jedna wartość y i odwrotnie: $y=-\frac{1}{b}(ax+c)$, $x=-\frac{1}{a}(by+c)$.

Miarą „dopasowania” tej prostej do zbioru punktów empirycznych (x_i, y_i) jest *wariancja odległości* punktów od prostej.

Trzeba zaznaczyć, że ta prosta główna najlepiej aproksymuje, w sensie metody najmniejszych kwadratów, dane ciągi obserwacji elementów meteorologicznych. Jeśli x, y mają normalny dwuwymiarowy rozkład gęstości prawdopodobieństwa, to prosta główna pokrywa się z dużą osią elipsy prawdopodobieństw. Długości półosi tej elipsy i współczynnik kierunkowy prostej głównej przyjęto jako parametry rozkładu prawdopodobieństwa. Miarą dokładności jest elipsoida błędów wyznaczona według zasady największej wiarygodności (prawdopodobieństwa). Jej objętość jest miernikiem „dobroci" ocen wymienionych parametrów rozkładu w zbiorowości generalnej.

Wyprowadzenia wzorów podstawowych znacznie upraszczają się przy zastosowaniu rachunku krakowianowego. Wtedy obliczenia sprowadzają się do elementarnych przekształceń algebraicznych.

Przypuszcza się, że sugerowana metoda znajdzie w najbliższej przyszłości szerokie zastosowanie w badaniach klimatologicznych.

1. WYZNACZANIE PROSTYCH GŁÓWNYCH

Dane są dwa ciągi obserwacji parametrów meteorologicznych, np. prędkość wiatru x i temperatura powietrza y o liczebności n

$$\mathbf{X} = \begin{Bmatrix} x_1 \, x_2 \ldots x_n \\ y_1 \, y_2 \ldots y_n \end{Bmatrix} \tag{1}$$

Mając zbiór punktów empirycznych $\mathbf{X}_i = \begin{Bmatrix} x_i \\ y_i \end{Bmatrix}$, określimy współzależność parametrów $\mathbf{X} = \begin{Bmatrix} x \\ y \end{Bmatrix}$

Jak wiadomo, w prostokątnym układzie współrzędnych, równaniem normalnym prostej prostopadłej do wektora $\mathbf{a} = \begin{Bmatrix} a_1 \\ a_2 \end{Bmatrix}$ jest

$$a_1 x + a_2 y + a_0 = 0 \tag{2}$$

gdzie $a_1^2 + a_2^2 = 1$.

Odległość względna δ_i (w stosunku do $\mathbf{a}$) punktu $\mathbf{X}_i$ od tej prostej równa jest

$$\delta_i = a_1 x_i + a_2 y_i + a_0 \tag{3}$$

przy czym a_0 jest odległością początku współrzędnych $\mathbf{O}$ od tej prostej.

Odległość bezwzględna (zwykła) punktu $\mathbf{X}_i$ od tej prostej równa jest wartości bezwzględnej $|\delta_i|$.

Znajdziemy taką prostą /2/, aby suma odległości względnych δ_i danych punktów empirycznych od niej równała się zeru, a suma odległości $|\delta_i|$ była najmniejsza

$$\sum_{i=1}^{n} \delta_i = 0 \tag{4}$$

$$\sum_{i=1}^{n} |\delta_i| = \min . \tag{5}$$

Prostą tę wyznaczamy, przyjmując warunek, aby suma kwadratów odległości tych punktów od niej osiągnęła minimum

$$\sum_{i=1}^{n} |\delta_i|^2 = \sum_{i=1}^{n} \delta_i^2 = \min . \tag{6}$$

Trzeba więc znaleźć takie rozwiązanie układu równań liniowych

$$\begin{gathered} a_1 x_1 + a_2 y_1 + a_0 = \delta_1 \\ a_1 x_2 + a_2 y_2 + a_0 = \delta_2 \\ \dots\dots\dots\dots \\ a_1 x_n + a_2 x_n + a_0 = \delta_n \end{gathered} \tag{7}$$

żeby suma

$$\sigma = \frac{1}{n} \sum_{i=1}^{n} \delta_i^2 = \frac{1}{n} \sum_{i=1}^{n} (a_1 x_i + a_2 y_i + a_0)^2 \tag{8}$$

była minimalna.

Sumując stronami te równania i dzieląc przez n otrzymuje się

$$\frac{1}{n} \sum_{i=1}^{n} \delta_i = a_1 \bar{x} + a_2 \bar{y} + a_0 \tag{9}$$

gdzie $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$ są to średnie arytmetyczne danych ciągów obserwacji.

Jak widać, warunek /4/ może być spełniony tylko wtedy, gdy punkt o współrzędnych $\bar{\mathbf{X}} = \begin{Bmatrix} \bar{x} \\ \bar{y} \end{Bmatrix}$ leży na szukanej prostej.

Równanie i-te układu /7/ przekształćmy i zapiszmy w symbolice krakowianowej.

$$a\Delta\mathbf{X}_i + a_1 \bar{x} + a_2 \bar{y} + a_0 = \delta_i \tag{10}$$

gdzie

$$\Delta\mathbf{X}_i = \mathbf{X}_i - \bar{\mathbf{X}}, \ \Delta\mathbf{X} = \begin{Bmatrix} \Delta x_1 \, \Delta x_2 \ldots \Delta x_n \\ \Delta y_1 \, \Delta y_2 \ldots \Delta y_n \end{Bmatrix}$$

Można sprawdzić, że

$$(\mathbf{a}\,\Delta\mathbf{X}_i)^2 = \mathbf{a} \cdot \begin{Bmatrix} \Delta^2 x_i & \Delta y_i \, \Delta x_i \\ \Delta x_i \Delta y_i & \Delta^2 y_i \end{Bmatrix} \cdot \mathbf{a} \tag{11}$$

A więc

$$\sigma = \mathbf{a} \cdot \mathbf{S} \cdot \mathbf{a} + (a_1 \bar{x} + a_2 \bar{y} + a_0)^2 \tag{12}$$

przy czym

$$\mathbf{S} = \frac{1}{n} (\tau\Delta\mathbf{X})^2 = \begin{Bmatrix} S_x^2 & S_{yx} \\ S_{xy} & S_y^2 \end{Bmatrix} \tag{13}$$

jest krakowianem kowariancji

$S_x^2 = \frac{1}{n} \sum_{i=1}^{n} \Delta^2 x_i$, $S_y^2 = \frac{1}{n} \sum_{i=1}^{n} \Delta^2 y_i$ – wariancje zmiennych x, y.

$S_{xy} = S_{yx} = \frac{1}{n} \sum_{i=1}^{n} \Delta x_i \Delta y_i$ – kowariancje zmiennych x, y.

Suma kwadratów /12/ jest najmniejsza, jeśli

$$a_1 \bar{x} + a_2 \bar{y} + a_0 = 0 \tag{14}$$

czyli

$$\mathbf{a} \cdot \mathbf{S} \cdot \mathbf{a} = \sigma \tag{15}$$

Oznacza to, że punkt o współrzędnych $\bar{\mathbf{X}}$ leży na tej prostej, bowiem spełniony jest wtedy warunek /4/.

Wystarczy teraz znaleźć wartości ekstremalne formy kwadratowej /15/. Korzystając, że $\mathbf{a}^2 = 1$ z równania /15/ mamy

$$\mathbf{a} \cdot \mathbf{S} = \sigma \mathbf{a} \tag{16}$$

a po przekształceniach

$$\mathbf{a} \cdot (\mathbf{S} - \sigma \boldsymbol{\tau}) = \mathbf{0}$$

czyli

$$\begin{Bmatrix} a_1 \\ a_2 \end{Bmatrix} \cdot \begin{Bmatrix} S_x^2 - \sigma & S_{yx} \\ S_{xy} & S_y^2 - \sigma \end{Bmatrix} = 0 \tag{17}$$

Są to równania jednorodne

$$\begin{aligned} a_1 (S_x^2 - \sigma) + a_2 S_{yx} &= 0 \\ a_1 S_{xy} + a_2 (S_y^2 - \sigma) &= 0 \end{aligned} \tag{18}$$

o niewiadomych a_1, a_2

We współrzędnych $\Delta \mathbf{X} = \begin{Bmatrix} \Delta x \\ \Delta y \end{Bmatrix}$ o początku $\bar{\mathbf{X}}$, szukana prosta ma równanie

$$a_1 \Delta x + a_2 \Delta y = 0 \tag{19}$$

Układ równań /18/ ma rozwiązanie niezerowe wtedy, gdy

$$\det (S - \sigma \tau) = \begin{vmatrix} S_x^2 - \sigma & S_{xy} \\ S_{xy} & S_y - \sigma \end{vmatrix} = 0 \tag{20}$$

czyli jeśli σ jest pierwiastkiem trójmianu

$$\sigma^2 - (S_x^2 + S_y^2) \sigma + \det S = 0 \tag{21}$$

jego pierwiastki

$$\sigma = \tfrac{1}{2} [(S_x^2 + S_y^2) \pm \sqrt{(S_x - S_y^2)^2 + 4 S_{xy}^2}] \tag{22}$$

są ekstremalnymi wartościami sumy /8/. Tylko jeden z nich, $\sigma_{\min}$, spełnia warunek najmniejszych odległości /6/.

Istnieje zatem tylko jedna prosta, dla której suma kwadratów odległości punktów od niej jest minimalna.

Ponieważ suma pierwiastków $\sigma_{\min}$, $\sigma_{\max}$ równania kwadratowego /21/ równa jest

$$\sigma_{\min} + \sigma_{\max} = S_x^2 + S_y^2 \tag{23}$$

a iloczyn

$$\sigma_{\min}\sigma_{\max}=\det S>0 \tag{23}$$

więc są one zawsze rzeczywiste i dodatnie

Proste odpowiadające pierwiastkom $\sigma_{\min}$ i $\sigma_{\max}$ będziemy nazywać głównymi

Równaniem kierunkowym prostej głównej $\sigma=\sigma_{\min}$ jest

$$\Delta y=\frac{S_{xy}}{S_x^2-\sigma_{\min}}\Delta x$$

lub

$$\Delta y=\frac{S_y^2-\sigma_{\min}}{S_{xy}}\Delta x \tag{24}$$

Można je również napisać w postaci

$$\Delta y=\tfrac{1}{2}(s-\sqrt{s^2+4})\Delta x$$

przyjmując

$$s=\frac{s_x^2-s_y^2}{S_{xy}} \tag{25}$$

Eliminując parametr σ z równań /18/ i korzystając, że $\mathbf{a}^2=1$ otrzymuje się

$$\begin{aligned} a_1^2&=\tfrac{1}{2}\left(1\pm\frac{s}{\sqrt{s^2+4}}\right)\\ a_2^2&=\tfrac{1}{2}\left(1\mp\frac{s}{\sqrt{s^2+4}}\right)\end{aligned} \tag{26}$$

Proste te przechodzą przez środek zbioru **X** i są prostopadłe

Niech krakowian

$$\mathbf{A}=\begin{Bmatrix} a_{11} & a_{21}\\ a_{12} & a_{22}\end{Bmatrix} \tag{27}$$

wyraża wektory prostopadłe do tych prostych uporządkowane od $\sigma_{\max}$ do $\sigma_{\min}$. Są to kierunki główne zbioru **X**.

Wartości $\sigma_{\min}$, $\sigma_{\max}$ są to pierwiastki (wartości własne) krakowianu **S**. Można sprawdzić, że krakowian **A** spełnia warunek ortogonalności $\mathbf{A}^2=\tau$, gdzie

$$\tau=\begin{Bmatrix} 1 & 0\\ 0 & 1\end{Bmatrix} \tag{28}$$

W przypadku szczególnym $\sigma=0$, tj.

$$\det\mathbf{S}=\begin{vmatrix} s_x^2 & s_{yx}\\ s_{xy} & s_y^2\end{vmatrix}=0 \tag{29}$$

Układ równań /18/ jest na ogół sprzeczny. Wówczas proste główne /24/ stają się prostymi regresji

$$\Delta y = \frac{s_{xy}}{s_x^2} \Delta x$$
$$\Delta x = \frac{s_{xy}}{s_y^2} \Delta y \qquad (30)$$

Niesprzeczność występuje w szczególnym przypadku, gdy współczynnik korelacji

$$r = \frac{s_{xy}}{s_x s_y} = \pm 1 \qquad (31)$$

gdzie s_x, s_y są to odchylenia standardowe $s_x = \sqrt{s_x^2}$, $s_y = \sqrt{s_y^2}$.

Niesprzeczność występuje więc w przypadku det $\mathbf{S} = 0$ $(r^2 = 1)$, czyli dla zależności funkcyjnej $\Delta y = \pm \Delta x$. Proste regresji /30/ wtedy pokrywają się i tworzą z osią x kąt 45° lub 135°.

Podwójny pierwiastek $\sigma_{\min} = \sigma_{\max} = \frac{1}{2}(s_x^2 + s_y^2)$ oznacza $r = 0$. Jeśli $s \to +\infty$, to $a \to 1$ i $b \to 0$ – prosta jest równoległa do osi x $(r = 0)$. Dla $s \to -\infty$, $a \to 0$, $b \to 1$ – prosta jest równoległa do osi x $(r = 0)$.

Miarą δ^2 rozproszenia punktów danego zbioru $\mathbf{X}$ względem prostej $\sigma = \sigma_{\min}$ jest wariancja odległości

$$\delta^2 = \frac{1}{n} \sum_{i=1}^{n} (a_2 \Delta x_i + b_2 \Delta y_i)^2 = \sigma_{\min} \qquad (32)$$

Miarą względną zaś może być stosunek

$$\frac{\delta}{\delta_o} \qquad (33)$$

gdzie $\delta_0^2 = \bar{x}^2 + \bar{y}^2$ jest odległością początku współrzędnych $\bar{\mathbf{X}}$ i $\mathbf{O}$. W przypadku $\sigma = 0$ $(r^2 = 1)$, odchylenie główne $\delta = \sqrt{\sigma_{\min}} = 0$, wszystkie punkty leżą na prostej.

Jeżeli odchylenie główne równe jest odległości punktu $(\bar{x}, \bar{y})$ od początku współrzędnych $(0, 0)$, to $\delta / \delta_0 = 1$.

Ten sam rezultat można uzyskać sprowadzając formę kwadratową $\Delta\mathbf{X} \cdot \mathbf{S} \cdot \Delta\mathbf{X}$ do postaci kanonicznej $\mathbf{Z} \cdot \mathbf{s} \cdot \mathbf{Z}$, gdzie

$$\mathbf{s} = \begin{Bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{Bmatrix} \qquad (34)$$

przez odpowiedni obrót (o kąt φ) układu współrzędnych $\Delta\mathbf{X}$ dookoła punktu $\bar{\mathbf{X}}$

$$\mathbf{Z} = \Delta\mathbf{X} \cdot \mathbf{A} \qquad (35)$$

przy czym $\mathbf{A}$ jest krakowianem obrotowym.

Powinien być spełniony warunek

$$\Delta\mathbf{X}\cdot\mathbf{S}\cdot\Delta\mathbf{X}=\mathbf{Z}\cdot\mathbf{s}\cdot\mathbf{Z}=\Delta\mathbf{X}(\tau\mathbf{A}\,\mathbf{s}\,\tau\mathbf{A})\,\Delta\mathbf{X}$$

z którego wynika, że

$$\mathbf{S}=\tau\mathbf{A}\cdot\mathbf{s}\cdot\tau\mathbf{A} \tag{36}$$

gdzie τ jest krakowianem jednostkowym.

Po elementarnych przekształceniach mamy

$$\mathbf{s}=\mathbf{A}\cdot\mathbf{S}\cdot\mathbf{A} \tag{37}$$

Odjęcie od obu stron ostatniej równości krakowianu $\sigma\tau$(σ – jest dowolną liczbą rzeczywistą) i uwzględnienie warunku $\mathbf{A}^2=\tau$ daje

$$\mathbf{A}\cdot(\mathbf{S}-\sigma\tau)\cdot\mathbf{A}=\begin{Bmatrix}\sigma_1-\sigma & 0\\ 0 & \sigma_2-\sigma\end{Bmatrix} \tag{38}$$

Dla $\sigma=\sigma_j$ otrzymuje się układ równań jednorodnych /18/

$$\mathbf{A}_j\cdot(\mathbf{S}-\sigma\tau)=0 \qquad j=1,2 \tag{39}$$

A więc $\mathbf{A}_j$ jest j-tą kolumną krakowianu /27/.

Zarówno kolumny, jak i wiersze krakowianu **A** określają wektory wzajemnie prostopadłe

$$\mathbf{A}_j\cdot\mathbf{A}_l=\begin{matrix}0\\1\end{matrix}, \quad \text{gdy} \quad \begin{matrix}j\neq l\\ j=l\end{matrix} \tag{40}$$

Liczby σ_1, σ_2 są pierwiastkami trójmianu /21/.

Dokonując obrotu układu współrzędnych o kąt φ dookoła środka $\bar{\mathbf{X}}$, zbiór **X** ulega transformacji

$$\mathbf{Z}=\Delta\mathbf{X}\cdot\mathbf{A}=\begin{Bmatrix}z_{11} & z_{21} & \cdots & z_{n1}\\ z_{12} & z_{22} & \cdots & z_{n2}\end{Bmatrix}$$

przy czym

$$\begin{aligned}z_1&=a_{11}\Delta x+a_{12}\Delta y=\mathbf{A}_1\Delta\mathbf{X}\\ z_2&=a_{21}\Delta x+a_{22}\Delta y=\mathbf{A}_2\Delta\mathbf{X}\end{aligned} \tag{41}$$

Można sprawdzić, że

$$\mathbf{A}=\begin{Bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{Bmatrix} \tag{42}$$

Osie **Z** nowych współrzędnych pokrywają się z kierunkami głównymi zbioru **X**.

Zgodnie z wyrażeniem /19/, pierwiastki σ_1, σ_2 są to wariancje odległości punktów empirycznych $\mathbf{X}_i$ od prostych $\sigma=\sigma_1=\sigma_{\max}$ i $\sigma=\sigma_2=\sigma_{\min}$

$$\sigma_1=\frac{1}{n}\sum_{i=1}^{n} z_{i1}^2 \tag{43}$$

$$\sigma_2=\frac{1}{n}\sum_{i=1}^{n} z_{i2}^2 \tag{43}$$

Wykażemy, że kowariancja zmiennych $\mathbf{Z}=\begin{Bmatrix} z_1 \\ z_2 \end{Bmatrix}$ równa jest zeru:

$$\sigma_{12}=\frac{1}{n}\sum_{i=1}^{n} z_{i1} z_{i2}=\frac{1}{n}\sum_{i=1}^{n} A_1\cdot(\tau\Delta \mathbf{X}_i)^2\cdot A_2=A_1\cdot\mathbf{S}\cdot A_2=\sigma_1 A_1\cdot A_2=0 \tag{44}$$

bowiem $\mathbf{A}_1\cdot\mathbf{S}=\sigma_1\mathbf{A}_1$, a $\mathbf{A}_1\cdot\mathbf{A}_2=0$.

Stąd wniosek, że zmienne z_1, z_2 nie są skorelowane. Krakowian kowariancji zmiennych $\mathbf{Z}$ ma postać

$$\mathbf{s}=\begin{Bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{Bmatrix}$$

Dla zmiennych normowanych w sposób

$$z_1'=\frac{z_1}{\sqrt{\sigma_1}} \qquad z_2'=\frac{z_2}{\sqrt{\sigma_2}} \tag{45}$$

krakowian kowariancji $\mathbf{S}=\boldsymbol{\tau}$.

Wtedy trójmian /21/ upraszcza się do

$$(1-\sigma)^2=0 \tag{46}$$

i ma podwójny pierwiastek $\sigma_1=\sigma_2=1$.

W układzie współrzędnych standaryzowanych $x'=\frac{\Delta x}{s_x}$, $y'=\frac{\Delta y}{s_y}$ krakowian $\mathbf{A}$ przybiera postać

$$\mathbf{A}=\frac{1}{\sqrt{2}}\begin{Bmatrix} 1 & -1 \\ 1 & 1 \end{Bmatrix} \tag{47}$$

tzn.

$$\mathbf{A}=\begin{Bmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45 \end{Bmatrix}$$

Wynika to bezpośrednio z równania

$$\begin{aligned}(1-\sigma)a_{11}+ra_{12}&=0\\ ra_{21}+(1-\sigma)a_{22}&=0\end{aligned} \tag{48}$$

Przyjmując $a_{12}=a_{22}=1$ z pierwszego równania mamy $a_{11}=1$ i $a_{2\mathrm{r}}=-1$, gdyż $\sigma_1=1+r$, a $\sigma_2=1-r$.

We współrzędnych standaryzowanych, prosta główna jest nachylona pod kątem 45° do osi x.

Można sprawdzić, że również dla zmiennych

$$\left\{\begin{matrix} x'' \\ y'' \end{matrix}\right\} = \left\{\begin{matrix} x' \\ y' \end{matrix}\right\} \cdot \left\{\begin{matrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{matrix}\right\}$$

krakowian kowariancji jest jednostkowy.

Przykład 1

Niech zbiór **X** stanowią średnie 10-letnie (okres 1951 - 1960) wartości wilgotności bezwzględnej z 60 stacji meteorologicznych o danej szerokości geograficznej. Dla porównania weźmiemy pod uwagę półrocze ciepłe (IV - IX) — zbiór $\mathbf{X}_l$ i półrocze chłodne (X - III) — zbiór $\mathbf{X}_z$. Środki tych zbiorów określa tabela 1.

Tabela 1

Zmienne	x	Półrocze ciepłe	Półrocze chłodne	
Szerokość geograficzna	φ	51,8217	51,8217	1°
Wilgotność bezwzględna powietrza	ρ	9,2355	4,7538	g/m³

Krakowiany kowariancji mają postać

$$\begin{array}{cc} \text{IV - IX} & \text{X - III} \\ \mathbf{S}=\left\{\begin{matrix} 2{,}545 & \cdot \\ 0{,}1502 & 0{,}4801 \end{matrix}\right\}, & \mathbf{S}=\left\{\begin{matrix} 2{,}545 & \cdot \\ 0{,}2212 & 0{,}1570 \end{matrix}\right\}. \end{array}$$

Ich pierwiastki

$$\begin{array}{cc} \text{IV - IX} & \text{X - III} \\ \mathbf{s}=\left\{\begin{matrix} 2{,}556 & 0 \\ 0 & 0{,}4692 \end{matrix}\right\}, & \mathbf{s}=\left\{\begin{matrix} 2{,}566 & 0 \\ 0 & 0{,}1367 \end{matrix}\right\} \end{array}$$

spełniają równania

IV - IX $\quad \sigma^2-3{,}0251\sigma+1{,}19929\ =0$

X - III $\quad \sigma^2-2{,}7020\sigma+0{,}334936=0$

Kierunki główne danych zbiorów wyrażają krakowiany

$$\begin{array}{cc} \text{IV - IX} & \text{X - III} \\ \mathbf{A}=\left\{\begin{matrix} 0{,}9974 & 0{,}07214 \\ 0{,}07214 & -0{,}9974 \end{matrix}\right\}, & \mathbf{A}=\left\{\begin{matrix} 0{,}9958 & 0{,}09143 \\ 0{,}09143 & -0{,}9958 \end{matrix}\right\} \end{array}$$

A więc proste główne należące do mniejszych pierwiastków $\sigma_{\min}$ mają równania

IV - IX $\rho = 0{,}07233\varphi + 5{,}487$

X - III $\rho = 0{,}09182\varphi - 0{,}004185$

Równania zaś prostych regresji mają postać

IV - IX $\rho = 0{,}05900\varphi + 6{,}178$

$\varphi = 0{,}3128\rho + 48{,}93$

X - III $\rho = 0{,}08688\varphi + 0{,}2514$

$\varphi = 1{,}408\rho + 45{,}13$

Rozproszenie punktów empirycznych $\mathbf{X}_i = \begin{Bmatrix} \varphi_i \\ \rho_i \end{Bmatrix}$ względem prostych głównych charakteryzują odchylenia odległości

IV - IX $\delta = \sqrt{0{,}4692} = 0{,}6850$

X - III $\delta = \sqrt{0{,}1367} = 0{,}3698$

Jak widać jest ono znacznie większe w półroczu ciepłym niż chłodnym.

2. ELIPSA PRAWDOPODOBIEŃSTW ZMIENNYCH METEOROLOGICZNYCH

Przyjmiemy, że w populacji parametry mają normalny rozkład gęstości prawdopodobieństwa

$$f(\mathbf{X}) = \frac{1}{2\pi} |\mathbf{\Psi}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})\mathbf{\Psi}^{-1}(\mathbf{X}-\boldsymbol{\mu})} \tag{49}$$

gdzie $\boldsymbol{\mu}$ jest środkiem populacji, a Ψ^{-1} — odwrotnością krakowianu kowariancji

$$\mathbf{\Psi} = \begin{Bmatrix} \Psi_{11} & \Psi_{21} \\ \Psi_{12} & \Psi_{22} \end{Bmatrix}$$

Liniami stałej gęstości prawdopodobieństwa są elipsy o równaniu

$$\Delta\mathbf{X} \cdot \mathbf{\Psi}^{-1} \cdot \Delta\mathbf{X} \overset{\text{def}}{=} c^2 = \text{const} \tag{50}$$

Obróćmy układ współrzędnych $\mathbf{Z} = \Delta\mathbf{X} \cdot \tilde{\mathbf{A}}$ tak, by forma kwadratowa c^2 miała postać kanoniczną

$$\mathbf{Z} \cdot \boldsymbol{\psi}^{-1} \cdot \mathbf{Z} = c^2 \tag{51}$$

przy czym

$$\boldsymbol{\psi} = \begin{Bmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{Bmatrix}$$

Z porównania wyrażeń /50/ i /51/ wynika, że

$$\begin{aligned} \mathbf{\Psi}^{-1} &= \tau\tilde{\mathbf{A}} \cdot \boldsymbol{\psi}^{-1} \cdot \tau\tilde{\mathbf{A}} \\ \boldsymbol{\psi}^{-1} &= \tilde{\mathbf{A}} \cdot \mathbf{\Psi}^{-1} \cdot \tilde{\mathbf{A}} \end{aligned} \tag{52}$$

a następnie

$$\boldsymbol{\psi} = (\tilde{\mathbf{A}} \cdot \mathbf{\Psi}^{-1} \cdot \tilde{\mathbf{A}})^{-1} = \tilde{\mathbf{A}} \cdot \mathbf{\Psi} \cdot \tilde{\mathbf{A}} \tag{53}$$

czyli

$$\tilde{\mathbf{A}}_j \cdot (\mathbf{\Psi} - \boldsymbol{\psi}_j \cdot \tau) = 0$$

Jak widać elementy krakowianów $\mathbf{\Psi}_i$ i

$$\boldsymbol{\psi}^{-1} = \begin{Bmatrix} \frac{1}{\psi_1^2} & 0 \\ 0 & \frac{1}{\psi_2^2} \end{Bmatrix}$$

spełniają równania

$$\det(\mathbf{\Psi} - \boldsymbol{\psi}_j^2 \tau) = 0 \quad \text{i} \quad \det\left(\mathbf{\Psi}^{-1} - \frac{1}{\psi_j^2}\tau\right) = 0 \tag{54}$$

A więc krakowian transformujący $\tilde{\mathbf{A}}$ wyraża kierunki główne populacji $\mathbf{X}$. Funkcję rozkładu gęstości zmiennej $\mathbf{X}$ można zatem wyrazić wzorem

$$f(\mathbf{Z}) = \frac{1}{2\pi\psi_1\psi_2} e^{-\frac{1}{2}\left(\frac{z_1^2}{\psi_1^2} + \frac{z_2^2}{\psi_2^2}\right)} \tag{55}$$

gdyż $\det \Psi = \det \boldsymbol{\psi} = \psi_1 \psi_2$.

Rozkład ten, to znaczy

$$f(\mathbf{Z}) = \frac{1}{\sqrt{2\pi\psi_1}} e^{-\frac{1}{2}\frac{z_1^2}{\psi_1^2}} \cdot \frac{1}{\sqrt{2\pi\psi_2}} e^{-\frac{1}{2}\frac{z_2^2}{\psi_2^2}} \tag{56}$$

wskazuje na niezależność parametrów $\mathbf{Z}$.

Elipsa prawdopodobieństw /51/ ma orientację zgodną z wektorami $\tilde{\mathbf{A}}$. Jej osie leżą na osiach głównych populacji $\mathbf{X}$ i mają długości $c\,\psi_1$, $c\,\psi_2$. Ma ona powierzchnię równą

$$D = 4\pi c\, \psi_1 \psi_2 = 4\pi c\, |\Psi|^{\frac{1}{2}} \tag{57}$$

Załóżmy teraz, że w populacji zmienne Z są niezależne i mają gęstość określoną wzorem

$$f(\mathbf{Z}) = \frac{1}{2\pi\psi_1\psi_2} e^{-\frac{1}{2}\frac{\Delta^2 z_1}{\psi_1^2} + \frac{\Delta^2 z_2}{\psi_2^2}} \tag{58}$$

Ponieważ punkty empiryczne $\mathbf{Z}_i = \begin{Bmatrix} z_{i1} \\ z_{i2} \end{Bmatrix}$ można traktować jako zdarzenia niezależne, wobec tego prawdopodobieństwo zrealizowania zbioru

$$\mathbf{Z}=\Delta\mathbf{X}\cdot\tilde{\mathbf{A}}=\begin{Bmatrix} z_{11} & z_{21} & \dots & z_{n1} \\ z_{12} & z_{22} & \dots & z_{n2}\end{Bmatrix} \tag{59}$$

jest równe

$$F(\mathbf{Z})=f(\mathbf{Z}_1)\cdot f(\mathbf{Z}_2)\dots f(\mathbf{Z}_n)$$

czyli

$$F(\mathbf{Z})=\frac{1}{(2\pi\psi_1\psi_2)}e^{-\frac{1}{2}\sum\limits_{i=1}^{n}\left(\frac{\Delta^2 z_{i1}}{\psi_1^2}+\frac{\Delta^2 z_{i2}}{\psi_2^2}\right)} \tag{60}$$

Estymatory $\hat{\psi}$ parametrów rozkładu ψ otrzymuje się z warunków maksimum funkcji wiarygodności $L=F(\mathbf{Z})$

$$\ln L=\ln F(\mathbf{Z})=\max \tag{61}$$

Z równań wiarygodności

$$\frac{\partial \ln L}{\partial \psi_1}=0, \quad \frac{\partial \ln L}{\partial \Delta\psi_2}=0, \tag{62}$$

wynika, że

$$\hat{\psi}_1^2=\sigma_1=\frac{1}{n}\sum_{i=1}^{n} z_{i1}^2$$
$$\hat{\psi}_2^2=\sigma_2=\frac{1}{n}\sum_{i=1}^{n} z_{i2}^2 \tag{63}$$

A więc według zasady największej wiarygodności w populacji rozkład gęstości prawdopodobieństwa określa wzór

$$f(\mathbf{X})=\frac{1}{2\pi\sqrt{\sigma_1\sigma_2}}e^{-\frac{1}{2}\left(\frac{z_1^2+z_2^2}{\sigma_1\sigma_2}\right)} \tag{64}$$

Ponieważ zmienne $\frac{z_1}{\sqrt{\sigma_1}}$, $\frac{z_2}{\sqrt{\sigma_2}}$ są niezależne i mają normalne rozkłady gęstości prawdopodobieństwa o parametrach (0, 1) [3], wobec tego zmienna

$$\frac{z_1^2}{\sigma_1}+\frac{z_2^2}{\sigma_2}=c^2 \tag{65}$$

ma rozkład Chi – kwadrat χ^2 o dwóch stopniach swobody

$$f(c)=\tfrac{1}{2}e^{-\frac{c^2}{2}} \tag{66}$$

Prawdopodobieństwo położenia punktu **X** wewnątrz elipsy $c=c_g$ jest równe

$$P(c<c_g) \stackrel{\text{def}}{=} p = \int_0^{c_g} e^{-\frac{c^2}{2}} dc^2 = 1 - e^{-\frac{1}{2}c_g^2} \tag{67}$$

co jest zgodne ze wzorami [5, s. 431].

Prawdopodobieństwo położenia **X** w pierścieniu elipsy $c_1 < c < c_2$ wynosi

$$P(c_1 \leqslant c \leqslant c_2) = \tfrac{1}{2} \int_{c_1}^{c_2} e^{-\frac{1}{2}c^2} dc^2 = e^{-\frac{1}{2}c_1^2} - e^{\frac{1}{2}c_2^2} \tag{68}$$

Odwrotnie, przyjmując poziom ufności p można znaleźć przedział ufności $(0, c_g)$

$$c_g = ln \frac{1}{(1-p)^2} \tag{69}$$

Z prawdopodobieństwem p wiadomo, że punkt **X** leży wewnątrz elipsy o półosiach $c\sqrt{\sigma_1}$, $c\sqrt{\sigma_2}$.

We współrzędnych Z funkcja wiarygodności wyraża się

$$ln\, L = -n\, ln\, 2 - n\, ln(\psi_1 \psi_2) - \frac{1}{2} \sum_{i=1}^{n} \left(\frac{z_{i1}^2}{\psi_1^2} + \frac{z_{i2}^2}{\psi_2^2} \right) \tag{70}$$

Jak wiadomo [4], estymatory największej wiarygodności są najefektywniejsze i posiadają tutaj dwuwymiarowy asymptotycznie normalny rozkład o krakowianie kowariancji

$$\boldsymbol{\Lambda} = -E \left\{ \frac{\partial\, ln\, L}{\partial \psi_j \partial \psi_l} \right\}^{-1} \tag{71}$$

gdzie E jest symbolem uśrednienia. Ze względu na $E(\sigma_1) = \psi_1^2$ i $E(\sigma_2) = \psi_2^2$ otrzymuje się

$$\boldsymbol{\Lambda} = \left\{ \begin{array}{|cc|} \dfrac{2n}{\psi_1^2} & 0 \\ 0 & \dfrac{2n}{\psi_2^2} \end{array} \right\}^{-1} = \left\{ \begin{array}{|cc|} \dfrac{\psi_1^2}{2n} & 0 \\ 0 & \dfrac{\psi_2^2}{2n} \end{array} \right\} \tag{72}$$

Oznacza to, że estymatory $\hat{\psi}_1$, $\hat{\psi}_2$ mają rozkład prawdopodobieństwa o gęstości

$$f(\hat{\psi}) = \frac{n}{\pi \psi_1 \psi_2} e^{-n\left(\frac{z_1^2}{\psi_1^2} + \frac{z_2^2}{\psi_2^2}\right)} \tag{73}$$

skąd

$$\begin{aligned} f(\hat{\psi}_1) & \left(\frac{n}{\pi}\right)^{\frac{1}{2}} \frac{1}{\psi_1} e^{-n\frac{z_1^2}{\psi_1}} \\ f(\hat{\psi}_2) &= \left(\frac{n}{\pi}\right)^{\frac{1}{2}} \frac{1}{\psi_2} e^{-n\frac{z_2^2}{\psi_2}} \end{aligned} \tag{74}$$

Przyjmując odpowiedni poziom ufności p można w znany sposób znaleźć przedziały ufności dla parametrów populacji ψ_1, ψ_2. Zawsze musimy jednak zastąpić w powyższych rozkładach nieznane wariancje ψ_1^2, ψ_2^2 pierwiastkami σ_1, σ_2 z próby **X**.

Przykład 2

Napiszemy równania elipsy prawdopodobieństw i wyznaczymy jej powierzchnię według danych z przykładu 1.

Przyjmując $c=1$ mamy

IV - IX
$$\frac{z_1^2}{1{,}599^2}+\frac{z_2^2}{0{,}6850^2}=1$$

X - III
$$\frac{z_1^2}{1{,}602^2}+\frac{z_2^2}{0{,}3698^2}=1$$

Ich powierzchnie wynoszą: w półroczu ciepłym $D=3{,}442$, w półroczu chłodnym $D=1{,}861$.

Przybliżonymi krakowianami kowariancji estymatorów $\boldsymbol{\psi}$ są

IV - IX, X - III

$$\Lambda=\begin{Bmatrix} 0{,}0213 & 0 \\ 0 & 0{,}0040 \end{Bmatrix}, \quad \Lambda=\begin{Bmatrix} 0{,}0213 & 0 \\ 0 & 0{,}001139 \end{Bmatrix}$$

Z prawdopodobieństwem 95% wiadomo, że w zbiorowości generalnej półosie elipsy prawdopodobieństw leżą w przedziałach

IV - IX
$$1{,}3130 \leqslant \psi_1 \leqslant 1{,}8850$$
$$0{,}4000 \leqslant \psi_2 \leqslant 0{,}9720$$

X - III
$$1{,}3160 \leqslant \psi_1 \leqslant 1{,}8880$$
$$0{,}3036 \leqslant \psi_2 \leqslant 0{,}4360$$

3. ELIPSOIDA BŁĘDÓW

Załóżmy, że parametry meteorologiczne mają w populacji rozkład o gęstości prawdopodobieństwa

$$f(\mathbf{X})=\frac{1}{2\pi}|\Psi|^{-\frac{1}{2}}e^{-\frac{1}{2}\Delta\mathbf{X}\cdot\Psi^{-1}\cdot\Delta\mathbf{X}} \tag{75}$$

Prawdopodobieństwo zrealizowania zbioru **X** jest równe

$$F(\mathbf{X})=f(\mathbf{X}_1)f(\mathbf{X}_2)\dots f(\mathbf{X}_n)$$

to znaczy

$$F(\mathbf{X})=\frac{1}{(2\pi)^n|\Psi|^{-\frac{n}{2}}}e^{-\frac{1}{2}\sum_{i=1}^{n}\Delta\mathbf{X}_i\cdot\Psi^{-1}\cdot\Delta\mathbf{X}_i} \tag{76}$$

Funkcję wiarygodności $L=F(\mathbf{X})$ wyrazimy w postaci

$$ln\,L=-n\,ln\,2-\frac{n}{2}ln(\psi_1\psi_2)-\frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu})\cdot\boldsymbol{\Psi}^{-1}\cdot(\mathbf{X}_i-\boldsymbol{\mu}) \tag{77}$$

Zamiast parametrów $\boldsymbol{\mu}=\begin{Bmatrix}\mu_1\\ \mu_2\end{Bmatrix}$ i Ψ weźmiemy pod uwagę $\boldsymbol{\mu}$, $\boldsymbol{\psi}$ i α, przy czym α jest współczynnikiem kierunkowym prostej głównej należącej do mniejszego pierwiastka $\sigma_2=\sigma_{\min}$. Krakowian kowariancji Ψ daje się przekształcić

$$\boldsymbol{\Psi}=\frac{1}{1+\alpha^2}\begin{Bmatrix}\psi_1^2+\alpha^2\psi_2^2 & \alpha(\psi_2^2-\psi_1^2)\\ \alpha(\psi_2^2-\psi_1^2) & \psi_2^2+\alpha^2\psi_1^2\end{Bmatrix} \tag{78}$$

czyli

$$\boldsymbol{\Psi}^{-1}=\frac{1}{1+\alpha^2}\begin{Bmatrix}\frac{1}{\psi_1^2}+\alpha^2\frac{1}{\psi_2^2} & \alpha\left(\frac{1}{\psi_2^2}-\frac{1}{\psi_1^2}\right)\\ \alpha\left(\frac{1}{\psi_2^2}-\frac{1}{\psi_1^2}\right) & \frac{1}{\psi_2^2}+\alpha^2\frac{1}{\psi_1^2}\end{Bmatrix} \tag{79}$$

Należy zauważyć, że postać krakowianów Ψ i Ψ^{-1} jest podobna, z tym że w drugim przypadku zamiast pierwiastków ψ_1, ψ_2 występują ich odwrotności.

Estymatory $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\psi}}$, $\hat{\alpha}$, znajdziemy z warunków największej wiarygodności

$$\nabla L=0$$

gdzie

$$\boldsymbol{\nabla}=\begin{Bmatrix}\frac{\partial}{\partial\mu}\\ \vdots\\ \frac{\partial}{\partial\alpha}\end{Bmatrix}\overset{\text{def}}{=}\begin{Bmatrix}\boldsymbol{\nabla}_\mu\\ \boldsymbol{\nabla}_\psi\\ \boldsymbol{\nabla}_\alpha\end{Bmatrix} \tag{80}$$

jest krakowianowym operatorem różniczkowania.

Różniczkując funkcję wiarygodności /77/ przy uwzględnieniu wyrażenia /79/ mamy

$$\boldsymbol{\nabla}_\mu\,ln\,L=\sum_{i=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu})\cdot\boldsymbol{\Psi}^{-1}=0 \tag{81}$$

skąd $\boldsymbol{\mu}=\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i=\mathbf{X}$.

$$\frac{ln\,L}{\partial\psi_1}=-\frac{n}{\psi_1}-\frac{1}{2}\sum_{i=1}^{n}\Delta\mathbf{X}_i\frac{\partial\boldsymbol{\Psi}^{-1}}{\partial\psi_1}\Delta\mathbf{X}_i=-\frac{n}{\psi_1}+\frac{n}{\psi_1^3}\frac{1}{(1+\alpha^2)}(s_x^2-2\alpha s_{xy}+\alpha s_y^2)=0 \tag{82}$$

$$\frac{\partial\ln L}{\partial\psi_2}=\frac{n}{\psi_2}-\frac{1}{2}\sum_{i=1}^{n}\Delta\mathbf{X}_i\frac{\partial\boldsymbol{\Psi}^{-1}}{\partial\psi_2}\Delta\mathbf{X}_i=-\frac{n}{\psi_2}+\frac{n}{\psi_2^2}\frac{1}{1+\alpha^2}(\alpha\ s_x^2+2\alpha s_{xy}+s_y^2)=0 \tag{83}$$

$$\frac{\partial\ln L}{\partial\alpha}=-\frac{1}{2}\sum_{i=1}^{n}\Delta\mathbf{X}_i\frac{\partial\boldsymbol{\Psi}^{-1}}{\partial\alpha}\Delta\mathbf{X}_i=\frac{1}{1+\alpha^2}\left(\frac{1}{\psi_2^2}-\frac{1}{\psi^2}\right)s_{xy}(\alpha^2-s\alpha-1)=0 \tag{84}$$

gdzie

$$s=\frac{s_x^2-s_y^2}{s_{xy}}$$

Z ostatniego równania wiarygodności ($\psi_1 \neq \psi_2$)

$$\alpha^2 - s\alpha - 1 = 0 \tag{85}$$

wynika, że estymatorem α jest współczynnik kierunkowy prostej głównej $\sigma=\sigma_{\min}$ równy

$$\hat{\alpha}=\tfrac{1}{2}(s-\sqrt{s^2+4})=\alpha_0 \tag{86}$$

Po uwzględnieniu /86/ równania /82/ przybierają postać

$$\begin{aligned} -\frac{n}{\psi_1}+\frac{n}{\psi_1^3}\sigma_1 &= 0 \\ -\frac{n}{\psi_2}+\frac{n}{\psi_2^3}\sigma_2 &= 0 \end{aligned} \tag{87}$$

Wówczas

$$\sum_{i=1}^{n} \Delta \mathbf{X}_i \frac{\partial \psi^{-1}}{\partial \psi_j} \Delta \mathbf{X}_i = \frac{2n\sigma_j}{\psi_j^3} \tag{88}$$

gdyż $\det \dfrac{\partial \boldsymbol{\psi}^{-1}}{\partial \psi_j} = 0$

Ostatecznie otrzymuje się $\hat{\psi}_1^2=\sigma_1$, $\hat{\psi}_2^2=\sigma_2$. A więc $\hat{\boldsymbol{\Psi}}=\mathbf{S}$, $\hat{\boldsymbol{\Psi}}^{-1}=\mathbf{S}^{-1}$. Zgodnie z powyższymi stwierdzeniami estymatory $\hat{\mathbf{U}}$

$$\Delta\mathbf{U}=\hat{\mathbf{U}}-\mathbf{U}=\begin{Bmatrix}\hat{\mu}_1\\ \hat{\mu}_2\\ \hat{\psi}_1\\ \hat{\psi}_2\\ \hat{\alpha}\end{Bmatrix}-\begin{Bmatrix}\mu_1\\ \mu_2\\ \psi_1\\ \psi_2\\ \alpha\end{Bmatrix}\overset{def}{=}\begin{Bmatrix}\Delta\mu\\ \Delta\psi\\ \Delta\alpha\end{Bmatrix} \tag{89}$$

będące rozwiązaniem równań największej wiarygodności (prawdopodobieństwa) mają 5-wymiarowy asymptotycznie normalny rozkład gęstości prawdopodobieństwa

$$f(\hat{\mathbf{U}})=\frac{1}{(2\pi)^{\frac{5}{2}}}|\mathbf{\Lambda}|^{-\frac{1}{2}}e^{-\frac{1}{2}\Delta\mathbf{U}\cdot\mathbf{\Lambda}^{-1}\cdot\Delta\mathbf{U}} \tag{90}$$

gdzie krakowian kowariancji $\mathbf{\Lambda}$ określony jest wzorem

$$\mathbf{\Lambda}=-(E\nabla^2 \ln L)^{-1} \tag{91}$$

przy czym

$$\nabla^2 = \nabla = \begin{Bmatrix}\nabla_\mu^2 & & \\ \nabla_\mu\nabla_\psi & \nabla_\psi^2 & \\ \nabla_\mu\nabla_\alpha & \nabla_\psi\nabla_\alpha & \nabla_\alpha^2\end{Bmatrix} \tag{92}$$

Elementy krakowianu $E\nabla^2 \ln L$ wynoszą

1) $$E\nabla_\mu^2 \ln L = -n\Psi^{-1} \tag{93}$$

2) $$\nabla_\mu \nabla_\psi \ln L = 0, \qquad \nabla_\mu \nabla_s \ln \alpha = 0$$

3) $$E(\nabla_\psi^2 \ln L) = \begin{Bmatrix} \dfrac{2n}{\psi_1^2} & 0 \\ 0 & \dfrac{2n}{\psi_2^2} \end{Bmatrix} = 2n\psi^{-1}$$

4) $$E(\nabla_\psi \nabla_\alpha \ln L) = 0$$

5) $$\nabla_\alpha^2 \ln L = -\frac{n}{(1+\alpha^2)^2}\left(\frac{1}{\psi_2^2} - \frac{1}{\psi_1^2}\right) s_{xy}(2\alpha - s)$$

$$E(\nabla_\alpha^2 \ln L) = -n\,\frac{(\psi_1^2 - \psi_2^2)^2}{\psi_1^2 \psi_2^2}\,\frac{1}{(1+\alpha^2)^2}$$

bo

$$Es_{xy}(2\alpha - s) = -(\psi_1^2 - \psi_2^2)$$

Zatem krakowian kowariancji Λ^{-1} wyraża się

$$\Lambda^{-1} = -\begin{Bmatrix} n\Psi^{-1} & & \\ & 2n\psi^{-1} & \\ & & n\,\dfrac{(\psi_1^2 - \psi_2^2)^2}{\psi_1^2 \psi_2^2}\,\dfrac{1}{(1+\alpha^2)^2} \end{Bmatrix} \tag{94}$$

skąd

$$\Lambda = \begin{Bmatrix} \dfrac{\Psi_{11}^2}{n} & \dfrac{\Psi_{21}}{n} & & & \\ \dfrac{\Psi_{12}}{n} & \dfrac{\Psi_{22}^2}{n} & & & \\ & & \dfrac{\psi_1^2}{2n} & & \\ & & & \dfrac{\psi_2^2}{2n} & \\ & & & & \dfrac{1}{n}\,\dfrac{\psi_1^2 \psi_2^2}{(\psi_1^2 - \psi_2^2)^2}\,(1+\alpha^2)^2 \end{Bmatrix} \tag{95}$$

Biorąc pod uwagę /95/ można napisać

$$f(\hat{\mathrm{U}}) = f(\hat{\mu})f(\hat{\psi})f(\hat{\alpha}) \tag{96}$$

gdzie

$$f(\hat{\mu}) = \frac{\sqrt{n}}{2\pi} |\Psi|^{-\frac{1}{2}} e^{-\frac{1}{2} n \Delta\mu \cdot \Psi^{-1} \cdot \Delta\mu} \tag{97}$$

$$f(\hat{\psi}) = \frac{n}{\pi \psi_1 \psi_2} e^{-n \Delta\psi \cdot \psi^{-1} \cdot \Delta\psi} \tag{98}$$

$$f(\alpha) = \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} \frac{\psi_1 \psi_2}{\psi_1^2 - \psi_2^2} (1+\alpha^2) e^{-\frac{n}{2} \frac{(\psi_1^2-\psi_2^2)^2}{\psi_1^2 \psi_2^2} \frac{\Delta^2\alpha}{(1+\alpha^2)^2}}$$

Według powyższych rozkładów można wyznaczyć przedziały ufności dla parametrów μ, ψ, α. Z punktu widzenia badań klimatologicznych najistotniejsza jest znajomość przedziału ufności dla współczynnika kierunkowego α prostej głównej. $\psi = \psi_{min}$. Tak jak poprzednio należy wstawić w miejsce nieznanych parametrów populacji ψ, α przybliżone wartości $\mathbf{s}$, α_0 (estymatory), wyznaczone z próby $\mathbf{X}$.

Elipsoidę błędów parametrów U można wyznaczyć sprowadzając jej równanie

$$\Delta \mathbf{U} \cdot \mathbf{\Lambda}^{-1} \cdot \Delta \mathbf{U} = C^2 = \text{const} \tag{99}$$

do postaci kanonicznej

$$\Delta \mathbf{u} \cdot \boldsymbol{\lambda}^{-1} \cdot \Delta \mathbf{u} = C^2 \tag{100}$$

gdzie

$$\boldsymbol{\lambda} = \begin{Bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_5^2 \end{Bmatrix} \tag{101}$$

według przekształcenia ortogonalnego

$$\mathbf{u} = \Delta \mathbf{U} \cdot \mathbf{B} \tag{102}$$

Podobnie jak przedtem związek między krakowianami Λ, λ i B ma postać

$$\begin{aligned} \mathbf{\Lambda}^{-1} &= \tau \mathbf{B} \cdot \boldsymbol{\lambda}^{-1} \cdot \tau \mathbf{B} \\ \boldsymbol{\lambda}^{-1} &= \mathbf{B} \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{B} \end{aligned} \tag{103}$$

Odejmując od obu stron ostatniej równości krakowian przekątny $\frac{1}{\lambda_j} \boldsymbol{\tau}$ otrzymuje się

$$\mathbf{B}_j \cdot \left(\mathbf{\Lambda}^{-1} - \frac{1}{\lambda_j^2} \boldsymbol{\tau} \right) = \mathbf{O} \tag{104}$$

Powyższy układ równań jednorodnych ma rozwiązanie niezerowe $\mathbf{B}_j \neq 0$ gdy

$$\det \left(\mathbf{\Lambda}^{-1} - \frac{1}{\lambda_j^2} \boldsymbol{\tau} \right) = 0 \tag{105}$$

czyli

$$\det (\mathbf{\Lambda} - \lambda_j^2 \boldsymbol{\tau}) = 0$$

gdyż według wzoru /103/

$$\boldsymbol{\lambda}=(\mathbf{B}\cdot\boldsymbol{\Lambda}^{-1}\cdot\mathbf{B})^{-1}=\mathbf{B}\cdot\boldsymbol{\Lambda}\cdot\mathbf{B} \tag{106}$$

oraz

$$\det\left(\boldsymbol{\Lambda}^{-1}-\frac{1}{\lambda_j^2}\,\boldsymbol{\tau}\right)=-\frac{1}{\lambda_j^2}\det\boldsymbol{\Lambda}^{-1}\det(\boldsymbol{\Lambda}-\lambda_j^2\,\boldsymbol{\tau})$$

A więc liczby λ_j są to pierwiastki krakowianu kowariancji $\boldsymbol{\Lambda}$. Wyznaczyć je należy z równania

$$\det(\boldsymbol{\Lambda}-\lambda_j^2\,\boldsymbol{\tau})=\det\left(\frac{1}{n}\,\Psi-\lambda_j^2\,\boldsymbol{\tau}\right)\det\left(\frac{1}{2n}\,\psi-\lambda_j\,\boldsymbol{\tau}^2\right)(\Lambda_{55}-\lambda_j^2)=0 \tag{107}$$

skąd

$$\boldsymbol{\lambda}=\left\{\begin{array}{ccccc} \frac{1}{n}\,\psi_1^2 & & & & \\ & \frac{1}{n}\,\psi_2^2 & & & \\ & & \frac{1}{2n}\,\psi_1^2 & & \\ & & & \frac{1}{2n}\,\psi_2^2 & \\ & & & & \frac{1}{n}\frac{\psi_1^2\psi_2^2}{(\psi_1^2-\psi_2^2)}\,(1+\alpha^2)^2 \end{array}\right\} \tag{108}$$

Elipsoida błędów ma równanie

$$\frac{u_1^2}{\lambda_1^2}+\frac{u_2^2}{\lambda_2^2}+\ldots+\frac{u_5^2}{\lambda_5^2}=C^2 \tag{109}$$

Jej objętość jako miara dokładności oszacowania parametrów **U** jest równa

$$G=\frac{2}{3}\,\pi\left(\frac{C}{n}\right)^{\frac{5}{2}}\psi_1^3\,\psi_2^3\,\frac{1+\alpha^2}{\psi_1^2-\psi_2^2} \tag{110}$$

Jeśli $n\to\infty$, wtedy $G\to 0$ jest miarą ufności, gdyż można wyznaczyć prawdopodobieństwa, z którymi oceny U będą leżały wewnątrz tej elipsoidy. Ma ona orientację taką jak wektory własne **B**. Jej półosie λ_j są to pierwiastki kwadratowe z wariancji zmiennych u_j.

$$\boldsymbol{\Omega}=\frac{1}{n}\sum_{i=1}^{n}(\tau\mathbf{u})^2=\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}\cdot(\tau\Delta\mathbf{U}_i)^2\cdot\mathbf{B} \tag{111}$$

$$=\mathbf{B}\cdot\boldsymbol{\Lambda}\cdot\mathbf{B}=\boldsymbol{\lambda}$$

Rozkład gęstości /90/ można więc wyrazić wzorem

$$f(\mathbf{u})=\left(\frac{1}{2\pi}\right)^{\frac{5}{2}}\frac{1}{\lambda_1\,\lambda_2\ldots\lambda_5}\,e^{-\frac{1}{2}\left(\frac{u_1^2}{\lambda_1^2}+\frac{u_2^2}{\lambda^2}+\ldots+\frac{u_5^2}{\lambda_5^2}\right)} \tag{112}$$

Zatem zmienna $c^2=\mathbf{u}\cdot\boldsymbol{\lambda}^{-1}\cdot\mathbf{u}$, będąca sumą kwadratów zmiennych niezależnych $\mathbf{U}$ o rozkładzie normalnym standaryzowanym ma rozkład Chi – kwadrat χ^2 o 5 stopniach swobody

$$f_5(C^2)=\frac{1}{2^{\frac{5}{2}}\Gamma(\frac{5}{2})}C^3e^{-\frac{C^2}{2}} \tag{113}$$

Prawdopodobieństwo tego, że błąd $\Delta\mathbf{U}$ leży wewnątrz elipsoidy $C=C_g$ jest równe

$$P(C<C_g)=\frac{1}{2^{\frac{5}{2}}\Gamma(\frac{5}{2})}\int\limits_0^{c_g}C^3e^{-\frac{C^2}{2}}dC^2 \tag{114}$$

$$=\frac{\gamma\left(\frac{5}{n},\frac{C}{2}\right)}{\Gamma(\frac{5}{2})}$$

gdzie Γ, γ są symbolami kompletnej i niekompletnej funkcji Gamma.

Odwrotnie, przyjmując odpowiedni poziom ufności p można wyznaczyć przedział ufności $(0, C_g)$, a tym samym długości poszczególnych osi elipsoidy dokładności i jej objętość. Najwygodniej jest posługiwać się tablicami rozkładu χ^2.

Przykład 3

Dla zbiorów $\mathbf{X}_l$, $\mathbf{X}_z$ (przykład 1) przybliżony krakowian kowariancji estymatorów $\hat{\mathbf{U}}$ ma postać

IV - IX

$$\boldsymbol{\Lambda}=\left\{\begin{matrix} 0{,}0424 & 0{,}0025 & & & \\ 0{,}0025 & 0{,}0080 & & & \\ & & 0{,}0213 & & \\ & & & 0{,}0040 & \\ & & & & 0{,}0047 \end{matrix}\right\}$$

X - III

$$\boldsymbol{\Lambda}=\left\{\begin{matrix} 0{,}0424 & 0{,}0037 & & & \\ 0{,}0037 & 0{,}0026 & & & \\ & & 0{,}0213 & & \\ & & & 0{,}0011 & \\ & & & & 0{,}0010 \end{matrix}\right\}$$

Skąd wynika, że

IV - IX

$$\boldsymbol{\lambda}=\left\{\begin{matrix} 0{,}0426 & & & & \\ & 0{,}0080 & & & \\ & & 0{,}0213 & & \\ & & & 0{,}0040 & \\ & & & & 0{,}0047 \end{matrix}\right\}$$

X - III

$$\lambda=\begin{Bmatrix} 0{,}0426 & & & & \\ & 0{,}0022 & & & \\ & & 0{,}0213 & & \\ & & & 0{,}0011 & \\ & & & & 0{,}0010 \end{Bmatrix}$$

Zatem elipsoida błędów $C=1$ ma równanie

IV - IX

$$\frac{u_1^2}{0{,}2064^2}+\frac{u_2^2}{0{,}0894^2}+\frac{u_3^2}{0{,}1456^2}+\frac{u_4^2}{0{,}0632^2}+\frac{u_5^2}{0{,}0316^2}=1$$

X - III

$$\frac{u_1^2}{0{,}02064^2}+\frac{u_2^2}{0{,}0469^2}+\frac{u_3^2}{0{,}1459^2}+\frac{u_4^2}{0{,}0332^2}+\frac{u_5^2}{0{,}0316^2}=1$$

Ich objętości są równe: w półroczu ciepłym $G=4{,}755\cdot 10^{-5}$, w półroczu chłodnym $G=6{,}480\cdot 10^{-6}$.

LITERATURA

[1] Banachiewicz T., *Rachunek krakowianowy,* Warszawa 1959.
[2] Boryczka J., *Badanie współzależności parametrów meteorologicznych metodą płaszczyzn głównych,* „Prace i Studia Instytutu Geografii" z. 7, Warszawa 1973.
[3] Deutsch R., *Teoria estymacji* (tłum. z ang.), Warszawa 1969.
[4] Kaczmarek Z., *Ocena parametrów rozkładu log-normalnego na podstawie związków korelacyjnych pomiędzy zmiennymi*", Przegląd Geofizyczny" z. 3 - 4, Warszawa 1967.
[5] *Krakowiany oraz ich zastosowanie w nauce i technice,* „ Zeszyty Naukowe Akademii Górniczo-Hutniczej im. Stanisława Staszica" nr 254, z. 18, Kraków 1971.
[6] Woldin B. G., Ganin M. P., Diner I. J., Komarow L. B., Swiesnikow A. A., Starobin K. B., *Problemy rachunku prawdopodobieństwa* (tłum. z ros.), Warszawa 1966.

Jerzy Boryczka

TWODIMENSIONAL PATTERN OF PROBABILITY OF OCCURRENCE OF METEOROLOGICAL ELEMENTS AND PHENOMENA

Summary

In his study the author embarked on establishing methods for defining the twodimensional pattern of the density of probability of meteorological variables. At the outset he presents a fairly simple method of defining the main straight line characterized by the fact, that the sum of the

squares of the distances of empiric points from this line is smallest. In contrast to the line of regressions, only one line exists which satisfies this condition. It is immaterial which one of the variables is treated as subordinate. The gauge indicating, how far the straight line is best adjusted with regard to the group of empiric points, is the variance in distance of the points from the straight line.

Chapter two deals with a method of plotting the ellipse of probability of atmospheric phenomena in a cracovian concept. In conclusion the author presents a way of how to construct an ellipsoid of errors in meteorological variables, starting out from equations of highest probability. He gives definite examples of how to apply the formulae he has developed, taking into consideration the pattern of absolute humidity determined for the 1951 to 1960 period for the area of Poland.

Ежи Борычка

ДВУИЗМЕРИТЕЛЬНОЕ РАСПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ МЕТЕОРОЛОГИЧЕСКИХ ЭЛЕМЕНТОВ И АТМОСФЕРНЫХ ЯВЛЕНИЙ

Краткое изложение

Главной целью работы являются методы определения двуизмерительного распределения густоты вероятности метеорологических переменных. С начала представлен довольно простой метод определения главной прямой, характеризующийся тем свойством, что сумма квадратов расстояний от нее эмпирических пунктов является наименьшей. В отличии от линии регрессии существует только одна прямая, которая исполняет это условие. Не то является существенным, которую из переменных трактуется как зависимую. Мерой приспособленности этой прямой к множеству эмпирических пунктов является варианция расстояния пунктов от прямой. Вторая глава трактует о способе определения эллипса вероятности атмосферных явлений в краковяном подходе. В конце представлен метод определения эллипсоида погрешностей метеорологичесих переменных, исходя из уравнений наибольшей достоверности. Приводятся конкретные примеры применения выведенных формул, принимая во внимание распределение абсолютной влажности на территории Польши (десятилетний период времени 1951 - 1960).