12. ROZWIĄZYWANIE RÓWNAŃ LINIOWYCH

<u>Lemat 3</u>

$$A_{1[m_1,n_1]}\Xi_{[n_1,m_2]}A_{2[m_2,n_2]} = 0_{[m_1,n_2]} \iff \bigvee_{R_{[n_1,m_2]}} \Xi = R - H_1RL_2$$

gdzie

$$H_1 = \mathcal{X}_1A_1\,, \qquad L_2 = A_2\mathcal{X}_2$$

a $\mathcal{X}_1$, $\mathcal{X}_2$ są dowolnie ustalonymi g-odwrotnościami odpowiednio macierzy A_1 i A_2.

Dowód

1) $A_1\Xi A_2 = 0 \Rightarrow \mathcal{X}_1A_1\Xi A_2\mathcal{X}_2 = 0 \Rightarrow H_1\Xi L_2 = 0 \Rightarrow$

$$\Rightarrow \Xi = \Xi - H_1\Xi L_2 \Rightarrow \bigvee_{R=\Xi} \Xi = R - H_1RL_2$$

2) $\bigvee_R \Xi = R - H_1RL_2 \Rightarrow A_1\Xi A_2 = A_1RA_2 - A_1H_1RL_2A_2 =$

$$= A_1RA_2 - A_1RA_2 = 0$$

na mocy wzorów (49).

W niektórych przypadkach wygodniej jest posługiwać się zamiast lematem 3 lematem następującym.

<u>Lemat 4</u>

$$A_{1[m_1,n_1]}\Xi_{[n_1,m_2]}A_{2[m_2,n_2]} = 0_{[m_1,n_2]} \iff \bigvee_{P_{[n_1,m_2]},\,Q_{[n_1,m_2]}} \Xi =$$
$$= (I_{[n_1]} - H_1)P + Q(I_{[m_2]} - L_2)$$

gdzie H_1, L_2, $\mathcal{X}_1$, $\mathcal{X}_2$ mają znaczenie jak w lemacie 3.

Dowód

1) $A_1 \Xi A_2 = 0 \Rightarrow \mathcal{X}_1 A_1 \Xi A_2 \mathcal{X}_2 = 0 \Rightarrow H_1 \Xi L_2 = 0 \Rightarrow$

$$\Rightarrow \Xi = \Xi - H_1\Xi + H_1\Xi - H_1\Xi L_2 \Rightarrow$$

$$\Rightarrow \Xi = (I_{[n_1]} - H_1)\Xi + H_1 \Xi (I_{[m_2]} - L_2) \Rightarrow$$

$$\Rightarrow \bigvee_{P=\Xi, Q=H_1\Xi} \Xi = (I_{[n_1]} - H_1)P + Q(I_{[m_2]} - L_2)$$

2) $\bigvee_{P,Q} \Xi = (I_{[n_1]} - H_1)P + Q(I_{[m_2]} - L_2) \Rightarrow$

$$\Rightarrow A_1\Xi A_2 = (A_1 - A_1H_1)PA_2 + A_1Q(A_2 - L_2A_2) = 0$$

na mocy wzorów (49).

Twierdzenie 30

Warunkiem koniecznym i wystarczającym niesprzeczności równania

(88) $$A_{1[m_1,n_1]} \Xi_{[n_1,m_2]} A_{2[m_2,n_2]} = T_{[m_1,n_2]}$$

jest

(89) $$L_1TH_2 = T$$

gdzie

$$L_1 = A_1\mathcal{X}_1, \quad H_2 = \mathcal{X}_2A_2$$

a $\mathcal{X}_1, \mathcal{X}_2$ są dowolnie ustalonymi g-odwrotnościami odpowiednio macierzy A_1 i A_2.

Dowód

1) Załóżmy najpierw, że równanie (88) jest niesprzeczne, czyli istnieje taka macierz Ξ, która spełnia to równanie. Wtedy

$$A_1\Xi A_2 = T \Rightarrow A_1\mathcal{X}_1A_1\Xi A_2\mathcal{X}_2A_2 = A_1\mathcal{X}_1T\mathcal{X}_2A_2 = T \Rightarrow L_1TH_2 = T$$

2) Załóżmy teraz, że jest spełniony warunek (89). Wtedy

$$L_1TH_2 = T \Rightarrow A_1\mathcal{X}_1T\mathcal{X}_2A_2 = T \Rightarrow \bigvee_{\Xi = \mathcal{X}_1T\mathcal{X}_2} A_1\Xi A_2 = T$$

co kończy dowód.

Twierdzenie 31

Jeśli równanie (88) jest niesprzeczne, to

$$A_1 \Xi A_2 = T \Leftrightarrow \bigvee_{R_{[n_1, m_2]}} \Xi = \mathcal{X}_1 T \mathcal{X}_2 + R - H_1 R L_2$$

gdzie H_1, L_2, $\mathcal{X}_1$, $\mathcal{X}_2$ mają znaczenie jak w lemacie 3.

Dowód

1) Załóżmy najpierw, że macierz Ξ spełnia równanie (88) czyniące zadość warunkowi (89). Wtedy

$$A_1 \Xi A_2 = T \quad \text{i} \quad A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 = T$$

skąd

$$A_1(\Xi - \mathcal{X}_1 T \mathcal{X}_2) A_2 = 0$$

i na mocy lematu 3

$$\bigvee_R \Xi = \mathcal{X}_1 T \mathcal{X}_2 + R - H_1 R L_2$$

2) Załóżmy teraz, że

$$\bigvee_R \Xi = \mathcal{X}_1 T \mathcal{X}_2 + R - H_1 R L_2$$

Wtedy na mocy (89) i (49)

$$A_1 \Xi A_2 = A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 + A_1 R A_2 - A_1 H_1 R L_2 A_2 =$$

$$= L_1 T H_2 + A_1 R A_2 - A_1 R A_2 = L_1 T H_2 = T$$

czyli Ξ jest rozwiązaniem równania (88).

W niektórych przypadkach jest wygodniej posługiwać się zamiast twierdzeniem 31 twierdzeniem następującym.

Twierdzenie 32

Jeśli równanie (88) jest niesprzeczne, to

$$A_1 \Xi A_2 = T \Leftrightarrow \bigvee_{P_{[n_1, m_2]}, Q_{[n_1, m_2]}} \Xi = \mathcal{X}_1 T \mathcal{X}_2 + (I_{[n_1]} - H_1) P + Q (I_{[m_2]} - L_2)$$

gdzie H_1, L_2, $\mathcal{X}_1$, $\mathcal{X}_2$ mają znaczenie jak w lemacie 3.

Dowód

Dowód twierdzenia 32 jest analogiczny do dowodu twierdzenia 31, z tą tylko różnicą, że w miejsce lematu 3 wykorzystuje się lemat 4.

Rozpatrzymy jeszcze kilka przypadków szczególnych, jako wnioski z udowodnionych lematów i twierdzeń. Symbolem $\mathcal{X}$ oznaczamy tu dowolnie ustaloną g-odwrotność macierzy A, ponadto $H = \mathcal{X}A$ i $L = A\mathcal{X}$.

Wniosek 1

$$A_{[m,n]}\Xi_{[n,p]} = 0 \Leftrightarrow \bigvee_{P_{[n,p]}} \Xi = (I_{[n]} - H)P$$

Wniosek 2

Warunkiem koniecznym i wystarczającym niesprzeczności równania

(90) $$A_{[m,n]}\Xi_{[n,p]} = T_{[m,p]}$$

jest

(91) $$LT = T$$

czyli

(92) $$(I_{[m]} - L)T = 0$$

Wniosek 3

Jeśli równanie (90) jest niesprzeczne, to

(93) $$A\Xi = T \Leftrightarrow \bigvee_{P_{[n,p]}} \Xi = \mathcal{X}T + (I_{[n]} - H)P$$

Wnioski 1, 2 i 3 otrzymujemy z lematu 4, twierdzeń 30 i 32, podstawiając na A_2 macierz jednostkową.

Wniosek 4

$$\Xi_{[q,m]}A_{[m,n]} = 0 \Leftrightarrow \bigvee_{Q_{[q,m]}} \Xi = Q(I_{[m]} - L)$$

Wniosek 5

Warunkiem koniecznym i wystarczającym niesprzeczności równania

$$(94)\qquad \Xi_{[q,m]} A_{[m,n]} = T_{[q,n]}$$

jest

$$(95)\qquad TH = T$$

czyli

$$(96)\qquad T(I_{[n]} - H) = 0$$

Wniosek 6

Jeśli równanie (94) jest niesprzeczne, to

$$(97)\qquad \Xi A = T \Leftrightarrow \bigvee_{Q_{[q,m]}} \Xi = T\mathcal{X} + Q(I_{[m]} - L)$$

Wnioski 4, 5 i 6 otrzymujemy z lematu 4, twierdzeń 30 i 32, podstawiając na A_1 macierz jednostkową.

Przykład 4

Rozwiążemy układ równań

$$-2x + y + 5z = 1$$

$$3x + 2y + 4z = 5$$

Układ ten można napisać w postaci

$$A\Xi = T$$

gdzie

$$A = \begin{bmatrix} -2 & 1 & 5 \\ 3 & 2 & 4 \end{bmatrix}, \qquad \Xi = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \qquad T = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

W przykładzie 1 obliczyliśmy g-odwrotność macierzy A

$$\mathcal{X} = \begin{bmatrix} 10 & 7 \\ -39 & -26 \\ 12 & 8 \end{bmatrix}$$

skąd

$$H = \mathcal{X}A = \begin{bmatrix} 10 & 7 \\ -39 & -26 \\ 12 & 8 \end{bmatrix} \begin{bmatrix} -2 & 1 & 5 \\ 3 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 24 & 78 \\ 0 & -91 & -299 \\ 0 & 28 & 92 \end{bmatrix}$$

$$L = A\mathcal{X} = \begin{bmatrix} -2 & 1 & 5 \\ 3 & 2 & 4 \end{bmatrix} \begin{bmatrix} 10 & 7 \\ -39 & -26 \\ 12 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Ponieważ $LT = T$, więc na mocy wniosku 2 dany układ równań jest niesprzeczny. Wobec tego na mocy wniosku 3 rozwiązanie tego układu można napisać w postaci

$$\Xi = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10 & 7 \\ -39 & -26 \\ 12 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} + \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 24 & 78 \\ 0 & -91 & -299 \\ 0 & 28 & 92 \end{bmatrix} \right) P$$

gdzie P jest dowolną macierzą postaci

$$P = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Mamy stąd

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 45 - 24b - 78c \\ -169 + 92b + 299c \\ 52 - 28b - 91c \end{bmatrix}$$

czyli

$$\begin{aligned} x &= 45 - 24b - 78c \\ y &= -169 + 92b + 299c \\ z &= 52 - 28b - 91c \end{aligned}$$

gdzie b i c są dowolnymi liczbami.

Kładąc $4b + 13c = d$ można rozwiązanie napisać w postaci

$$\begin{aligned} x &= 45 - 6d \\ y &= -169 + 23d \\ z &= 52 - 7d \end{aligned}$$

gdzie d jest dowolną liczbą.

O możliwości przedstawienia rozwiązania w zależności od jednego tylko parametru informuje nas rząd macierzy $I_{[n]} - H$. Mianowicie na mocy własności 11 - wiedząc z przykładu 1, że $\rho(A) = 2$ - mamy

$$\rho(I_{[n]} - H) = 1$$

Wobec tego wszystkie trzy kombinacje liniowe parametrów a, b, c, jakie otrzymujemy z mnożenia $(I_{[n]} - H)P$, muszą być liniowo zależne od jednej z nich.

Twierdzenie 33

Warunkiem koniecznym i wystarczającym niesprzeczności układu równań

$$(98)\qquad \begin{aligned} A_{1[m_1, n_1]}\Xi_{[n_1, m_2]} &= T_{1[m_1, m_2]} \\ \Xi_{[n_1, m_2]}A_{2[m_2, n_2]} &= T_{2[n_1, n_2]} \end{aligned}$$

jest niesprzeczność każdego z tych równań i $A_1T_2 = T_1A_2$, czyli na mocy wniosków 2 i 5

$$(99)\qquad L_1T_1 = T_1, \qquad T_2H_2 = T_2, \qquad A_1T_2 = T_1A_2$$

gdzie $L_1 = A_1\mathcal{X}_1$, $H_2 = \mathcal{X}_2A_2$ a $\mathcal{X}_1$ i $\mathcal{X}_2$ są g-odwrotnościami odpowiednio macierzy A_1 i A_2.

Dowód

1) Załóżmy najpierw, że istnieje macierz Ξ spełniająca układ równań (98). Zatem na mocy wniosków 2 i 5 są spełnione pierwsze dwa z warunków (99). Mnożąc obie strony pierwszego z równań (98) prawostronnie przez A_2, a obie strony drugiego lewostronnie przez A_1, otrzymujemy

$$A_1\Xi A_2 = T_1A_2, \qquad A_1\Xi A_2 = A_1T_2$$

a stąd ostatni z warunków (99).

2) Załóżmy teraz, że są spełnione warunki (99) i rozważmy macierz

$$\Xi = \mathcal{X}_1 T_1 + T_2 \mathcal{X}_2 - H_1 T_2 \mathcal{X}_2$$

gdzie $H_1 = \mathcal{X}_1 A_1$. Mamy

$$A_1 \Xi = A_1 \mathcal{X}_1 T_1 + A_1 T_2 \mathcal{X}_2 - A_1 H_1 T_2 \mathcal{X}_2 =$$

$$= L_1 T_1 + A_1 T_2 \mathcal{X}_2 - A_1 T_2 \mathcal{X}_2 = L_1 T_1 = T_1$$

$$\Xi A_2 = \mathcal{X}_1 T_1 A_2 + T_2 \mathcal{X}_2 A_2 - H_1 T_2 \mathcal{X}_2 A_2 =$$

$$= \mathcal{X}_1 A_1 T_2 + T_2 H_2 - H_1 T_2 H_2 =$$

$$= H_1 T_2 + T_2 - H_1 T_2 = T_2$$

Zatem macierz Ξ spełnia układ (98) i tym samym układ ten jest niesprzeczny.

Twierdzenie 34

Jeśli układ równań (98) jest niesprzeczny, to macierz Ξ jest jego rozwiązaniem wtedy i tylko wtedy, gdy można ją przedstawić w postaci

$$(100) \quad \Xi = \mathcal{X}_1 T_1 + T_2 \mathcal{X}_2 - H_1 T_2 \mathcal{X}_2 + (I_{[n_1]} - H_1) R (I_{[m_2]} - L_2)$$

gdzie - jak poprzednio - $H_1 = \mathcal{X}_1 A_1$, $L_2 = A_2 \mathcal{X}_2$, $\mathcal{X}_1$ i $\mathcal{X}_2$ są g-odwrotnościami odpowiednio macierzy A_1 i A_2, a $R = R_{[n_1, m_2]}$ jest dowolną macierzą o podanych wymiarach.

Dowód

1) Załóżmy najpierw, że macierz Ξ jest rozwiązaniem układu (98). Wtedy

$$(I_{[n_1]} - H_1)\Xi(I_{[m_2]} - L_2) = \Xi - H_1 \Xi - \Xi L_2 + H_1 \Xi L_2 =$$

$$= \Xi - \mathcal{X}_1 A_1 \Xi - \Xi A_2 \mathcal{X}_2 + H_1 \Xi A_2 \mathcal{X}_2 =$$

$$= \Xi - \mathcal{X}_1 T_1 - T_2 \mathcal{X}_2 + H_1 T_2 \mathcal{X}_2$$

Wobec tego

$$\Xi = \mathcal{X}_1 T_1 + T_2\mathcal{X}_2 - H_1 T_2 \mathcal{X}_2 + (I_{[n_1]} - H_1)\Xi(I_{[m_2]} - L_2)$$

czyli macierz Ξ można przedstawić w postaci (100).

2) Załóżmy teraz, że macierz Ξ jest dana wzorem (100). Wtedy

$$A_1\Xi = A_1\mathcal{X}_1 T_1 + A_1 T_2\mathcal{X}_2 - A_1 H_1 T_2\mathcal{X}_2 + (A_1 - A_1 H_1)R(I_{[m_2]} - L_2) =$$

$$= A_1\mathcal{X}_1 T_1 + A_1 T_2\mathcal{X}_2 - A_1 H_1 T_2\mathcal{X}_2 =$$

$$= L_1 T_1 + A_1 T_2\mathcal{X}_2 - A_1 T_2\mathcal{X}_2 = L_1 T_1 = T_1$$

$$\Xi A_2 = \mathcal{X}_1 T_1 A_2 + T_2\mathcal{X}_2 A_2 - H_1 T_2\mathcal{X}_2 A_2 + (I_{[n_1]} - H_1)R(A_2 - L_2 A_2) =$$

$$= \mathcal{X}_1 T_1 A_2 + T_2\mathcal{X}_2 A_2 - H_1 T_2\mathcal{X}_2 A_2 =$$

$$= \mathcal{X}_1 A_1 T_2 + T_2 H_2 - H_1 T_2 H_2 =$$

$$= H_1 T_2 + T_2 - H_1 T_2 = T_2$$

Zatem macierz (100) spełnia układ równań (98)..

Za pomocą macierzy H-symetrycznych i L-symetrycznych można rozwiązywać w sposób przybliżony również równania sprzeczne. Do tego celu można posługiwać się następującym twierdzeniem.

Twierdzenie 35

Norma

$$\left\| A_1 \Xi A_2 - T \right\|$$

osiąga minimum wtedy i tylko wtedy, gdy

$$\Xi = \mathcal{X}_1 T \mathcal{X}_2 + (I_{[n_1]} - H_1)P + Q(I_{[m_2]} - L_2) \tag{101}$$

gdzie $\mathcal{X}_1$ jest g-odwrotnością L-symetryczną macierzy A_1, $\mathcal{X}_2$ jest g-odwrotnością H-symetryczną macierzy A_2, a pozostałe symbole mają znaczenie takie jak w twierdzeniu 32.

Dowód

Połóżmy

$$\Xi = \mathcal{X}_1 T \mathcal{X}_2 + \Lambda$$

Na mocy (66) i (70) macierze $\mathcal{X}_1$ i $\mathcal{X}_2$ spełniają warunki

$$A_1^* A_1 \mathcal{X}_1 = A_1^*, \qquad \mathcal{X}_2 A_2 A_2^* = A_2^*$$

Wobec tego mamy

$$\left\| A_1 \Xi A_2 - T \right\|^2 = \left\| A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 - T + A_1 \Lambda A_2 \right\|^2 =$$

$$= \left\| A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 - T \right\|^2 + \left\| A_1 \Lambda A_2 \right\|^2 +$$

$$+ 2 \operatorname{Tr} \operatorname{Re} (A_2^* \Lambda^* A_1^* A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 - A_2^* \Lambda^* A_1^* T) =$$

$$= \left\| A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 - T \right\|^2 + \left\| A_1 \Lambda A_2 \right\|^2 +$$

$$+ 2 \operatorname{Tr} \operatorname{Re} (\mathcal{X}_2 A_2 A_2^* \Lambda^* A_1^* A_1 \mathcal{X}_1 T - A_2^* \Lambda^* A_1^* T) =$$

$$= \left\| A_1 \mathcal{X}_1 T \mathcal{X}_2 A_2 - T \right\|^2 + \left\| A_1 \Lambda A_2 \right\|^2$$

A zatem powyższa norma osiąga minimum wtedy i tylko wtedy, gdy $A_1 \Lambda A_2 = 0$, czyli na mocy lematu 4 wtedy i tylko wtedy, gdy macierz Ξ ma postać (101).

Z twierdzenia 35 wynika, że wzór (101) podaje wszystkie rozwiązania równania

$$A_1 \Xi A_2 = T \tag{102}$$

najlepsze w sensie normy, nawet w przypadkach, gdy równanie (102) jest sprzeczne.

Twierdzenie 35 można zatem uważać za uogólnienie twierdzenia 32.

Z twierdzenia 35 wynikają przypadki szczególne, które podamy w formie wniosków.

Wniosek 7

Norma

$$\| A\Xi - T \|$$

osiąga minimum wtedy i tylko wtedy, gdy

$$\Xi = \mathcal{X}T + (I - H)P \tag{103}$$

gdzie $\mathcal{X}$ jest g-odwrotnością L-symetryczną macierzy A, $H = \mathcal{X}A$ a jest dowolną macierzą o odpowiednich wymiarach.

Wniosek 8

Norma

$$\| \Xi A - T \|$$

osiąga minimum wtedy i tylko wtedy, gdy

$$\Xi = T\mathcal{X} + Q(I - L)$$

gdzie $\mathcal{X}$ jest g-odwrotnością H-symetryczną macierzy A, a Q jest dowolną macierzą o odpowiednich wymiarach.

Wniosek 7 można uważać za uogólnienie wniosku 3, a wniosek 8 na uogólnienie wniosku 6.

13. UOGÓLNIENIE ODLEGŁOŚCI MAHALANOBISA

Niech będzie dana populacja, w której bada się n cech, i niech C będzie macierzą kowariancji tych cech. Niech

$$Z_1 = \begin{bmatrix} z_{11} \\ \vdots \\ z_{n1} \end{bmatrix}, \quad Z_2 = \begin{bmatrix} z_{12} \\ \vdots \\ z_{n2} \end{bmatrix} \tag{104}$$

będą dwoma dowolnymi punktami n-wymiarowej przestrzeni euklidesowej o współrzędnych rzeczywistych, będących wartościami tych cech.

P. C. Mahalanobis [4] wprowadził wielkość

$$(105)\qquad D^2 = (Z_1 - Z_2)'C^{-1}(Z_1 - Z_2)$$

jako odległość między dwiema populacjami o równych macierzach kowariancji C, gdy punkty (104) są środkami ciężkości tych populacji, tzn. punktami o współrzędnych będących średnimi wartościami cech w każdej z dwu populacji.

Definicja (105) wymaga założenia, że macierz kowariancji C jest nieosobliwa. Tymczasem w wielu zastosowaniach napotykamy macierze kowariancji osobliwe. W niniejszym paragrafie uogólnimy pojęcie odległości Mahalanobisa na wszystkie możliwe macierze kowariancji C([5]).

<u>Uogólnioną odległością Mahalanobisa</u> nazwiemy odległość euklidesową punktów w układzie głównych komponent, a więc w układzie współrzędnych nieskorelowanych, po ich unormowaniu. Przez unormowanie komponenty y_k $(k = 1, \ldots, n)$ będziemy rozumieli zastąpienie jej przez $y_k/\sqrt{\lambda_k}$, gdy $\lambda_k > 0$ i pozostawienie jej bez zmiany, gdy $\lambda_k = 0$, gdzie λ_k jest odpowiednią wartością własną macierzy kowariancji C i zarazem wariancją komponenty k-ej. Gdy $\lambda_k = 0$, mamy $y_k = 0$ z prawdopodobieństwem 1.

Innymi słowy, jeżeli

$$Y_1 = \begin{bmatrix} y_{11} \\ \vdots \\ y_{n1} \end{bmatrix} \qquad Y_2 = \begin{bmatrix} y_{12} \\ \vdots \\ y_{n2} \end{bmatrix}$$

są obrazami punktów (104) w układzie głównych komponent, czyli

$$(106)\qquad Y_1 = P'(Z_1 - \bar{Z}), \qquad Y_2 = P'(Z_2 - \bar{Z}),$$

gdzie $\bar{Z}$ jest środkiem ciężkości danej populacji a

$$P = \begin{bmatrix} P_1 \ldots P_n \end{bmatrix} = \begin{bmatrix} p_{11} & \ldots & p_{1n} \\ \vdots & \vdots\vdots\vdots & \vdots \\ p_{n1} & \ldots & p_{nn} \end{bmatrix}, \quad PP' = P'P = I$$

(I macierz jednostkowa)

jest tzw. macierzą modalną, utworzoną z wektorów własnych $P_1, \ldots, P_n$ macierzy kowariancji C, to uogólnioną odległością Mahalanobisa punktów (104) będziemy nazywać liczbę

(107)
$$D = \sqrt{\left(\frac{y_{11}}{\sqrt{\lambda_1}} - \frac{y_{12}}{\sqrt{\lambda_1}}\right)^2 + \ldots + \left(\frac{y_{r1}}{\sqrt{\lambda_r}} - \frac{y_{r2}}{\sqrt{\lambda_r}}\right)^2 + (y_{r+1,1} - y_{r+1,2})^2 + \ldots + (y_{n1} - y_{n2})^2} =$$
$$= \sqrt{\frac{(y_{11} - y_{12})^2}{\lambda_1} + \ldots + \frac{(y_{r1} - y_{r2})^2}{\lambda_r}}$$

gdzie r jest liczbą niezerowych wartości własnych macierzy C, tzn. $\lambda_k > 0$ dla $k = 1, \ldots, r$ i $\lambda_k = 0$ dla $k = r + 1, \ldots, n$. Liczba r jest zarazem rzędem macierzy kowariancji C.

Wzór (107) można napisać w postaci

$$D^2 = (Y_1 - Y_2)' M (Y_1 - Y_2)$$

gdzie

(108)
$$M = \begin{bmatrix} \frac{1}{\lambda_1} & & & & \\ & \ddots & & & \\ & & \frac{1}{\lambda_r} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}$$

i wobec (106) otrzymujemy

$$D^2 = (Z_1 - Z_2)' PMP' (Z_1 - Z_2)$$

Podstawiając

(109)
$$C^+ = PMP'$$

otrzymujemy ostateczny wzór na uogólnioną odległość Mahalanobisa

(110)
$$D^2 = (Z_1 - Z_2)' C^+ (Z_1 - Z_2)$$

Wykażemy, że C^+ jest odwrotnością Moore'a-Penrosego macierzy kowariancji C, tzn. spełnia następujące warunki:

$$CC^+C = C$$
$$C^+CC^+ = C^+$$
$$(C^+C)' = C^+C$$
$$(CC^+)' = CC^+$$

W tym celu zauważmy najpierw, że macierzą kowariancji dla głównych komponent jest

$$S = E(\eta\eta') = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} \tag{111}$$

gdzie η jest wektorem losowym zapisanym kolumnowo, którego współrzędnymi są główne komponenty traktowane jako zmienne losowe. Jeśli ζ oznacza wektor losowy cech wyjściowych, również zapisany kolumnowo, a $\bar{\zeta} = E\zeta$, to analogicznie do (106) jest

$$\eta = P'(\zeta - \bar{\zeta}), \quad \text{czyli} \quad \zeta - \bar{\zeta} = P\eta$$

i wobec tego

$$C = E(\zeta - \bar{\zeta})(\zeta - \bar{\zeta})' =$$
$$= E(P\eta\eta'P') = PE(\eta\eta')P'$$

a stąd

$$C = PSP' \tag{112}$$

Ponadto na mocy (108) i (111) jest

$$SM = MS, \quad SMS = S, \quad MSM = M. \tag{113}$$

Opierając się na wzorach (112) i (113) otrzymujemy

$$CC^+C = PSP'PMP'PSP' = PSMSP' = PSP' = C$$
$$C^+CC^+ = PMP'PSP'PMP' = PMSMP' = PMP' = C^+$$
$$(C^+C)' = (PMP'PSP')' = (PMSP')' = PSMP' =$$
$$= PMSP' = PMP'PSP' = C^+C$$

$$(CC^+)' = (PSP'PMP')' = (PSMP')' = PMSP' =$$
$$= PSMP' = PSP'PMP' = CC^+$$

Tym samym wykazaliśmy, że - istotnie - macierz C^+ jest odwrotnością Moore a-Penrosego macierzy kowariancji C. Wynika stąd również, że w przypadku macierzy kowariancji C nieosobliwej jest $C^+ = C^{-1}$, czyli w tym przypadku wzór (110) można napisać w postaci (105). Dowodzi to, że wzór (110) jest istotnie uogólnieniem wzoru (105), czyli mamy do czynienia z uogólnieniem pojęcia odległości Mahalanobisa.

14. ZAKOŃCZENIE

W ostatnich dwu paragrafach podano przykłady zastosowań g-odwrotności i MP-odwrotności macierzy. Nie wyczerpuje to oczywiście wszystkich możliwości zastosowań tych odwrotności. Zamiarem autora było jedynie zilustrować przykładami działanie twierdzeń uprzednio udowodnionych dla łatwiejszego ich przyswojenia przez czytelnika.

Również algorytmy opisane w niniejszej pracy mają charakter jedynie przykładowy. Nie pretendują one bynajmniej do roli algorytmów numerycznie optymalnych.

Celem pracy jest spopularyzowanie pojęcia uogólnionych odwrotności i danie czytelnikowi możliwości operowania nimi, co w takich dziedzinach, jak algebra liniowa czy wielowymiarowa analiza statystyczna ma doniosłe znaczenie.

BIBLIOGRAFIA

[1] C. R. Rao. A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. J. Roy. Statist. Soc. 24, 152-158, 1962.

[2] E. H. Moore. General analysis. Philadelphia. Amer. Phil. Soc. 1935.

[3] R. Penrose. A generalized inverse for matrices. Proc. Camb. Phil. Soc., 51, 406-413, 1955.

[4] P. C. Mahalanobis. On the generalized distance in statistics. Proc.Nat. Inst. Sci. India, 12, 49-55, 1936.

[5] M. Warmuś. Uogólnienie odległości Mahalanobisa. Listy Biometryczne nr 30-33.

[6] G. Peters and J. H. Wilkinson. The least squares problem and pseudo-inverses. The Computer Journal, V. 13/3, 309-316, 1970.

[7] F. R. Gantmacher. Teorija matric. Wyd. II. Moskwa 1966.

SPIS TREŚCI