

Analiza dyskryminacyjna danych podwójnie wielowymiarowych

Mirosław Krzyśko, Michał Skorzybut

Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza

XXXV Konferencja Statystyka Matematyczna
Wiśła 7-11.12.2009

Wprowadzenie (1)

Założmy, że danych jest K klas obiektów scharakteryzowanych wartościami p cech obserwowanych w T momentach czasowych (lub w T różnych warunkach fizycznych lub w T różnych miejscach). Dane takie nazywamy, danymi podwójnie wielowymiarowymi. Chcemy skonstruować możliwie dobry klasyfikator na bazie n_i obserwacji z przeszłości z i -tej klasy, $i = 1, 2, \dots, K$ (na bazie próby uczącej).

Niech \mathbf{x}_{ijk} będzie wektorem wartości p cech zaobserwowanych w k -tym momencie czasowym na j -tym obiekcie pochodzącym z i -tej klasy, $k = 1, 2, \dots, T$; $j = 1, 2, \dots, n_i$; $i = 1, 2, \dots, K$ oraz niech

$$\mathbf{x}_{ij} = (\mathbf{x}'_{ij1}, \mathbf{x}'_{ij2}, \dots, \mathbf{x}'_{ijT})'$$

będzie wektorem rozmiaru $pT \times 1$ obserwacji dokonanych na j -tym obiekcie pochodzącym z i -tej klasy. Zakładamy, że wszystkie obserwacje \mathbf{x}_{ij} są niezależne oraz

$$\mathbf{x}_{ij} \sim N_{pT}(\boldsymbol{\mu}_i, \boldsymbol{\Omega}_i),$$

gdzie $\boldsymbol{\Omega}_i > 0$, $j = 1, 2, \dots, n_i$, $i = 1, 2, \dots, K$.

Warunkiem koniecznym i dostatecznym na to, by estymator $\hat{\Omega}_i$, dodatnio określonej macierzy kowariancji Ω_i , był dodatnio określony z prawdopodobieństwem równym jedności jest, by

$$n_i > pT, \quad i = 1, 2, \dots, K.$$

Gdy $n_i \leq pT$, załóżmy że

$$\Omega_i = \mathbf{V}_i \otimes \Sigma_i,$$

gdzie $\mathbf{V}_i > 0$ oraz $\Sigma_i > 0$, $i = 1, 2, \dots, K$.

Macierze \mathbf{V}_i i Σ_i możemy estymować oddzielnie, a warunkiem koniecznym i dostatecznym na to, by estymatory tych macierzy były określone dodatnio z prawdopodobieństwem równym jedności jest, by $n_i > \max(p, T)$.

Jednakże pojawia się innego rodzaju trudność. Ponieważ

$$(c\mathbf{V}_i) \otimes (c^{-1}\boldsymbol{\Sigma}_i) = \mathbf{V}_i \otimes \boldsymbol{\Sigma}_i,$$

dla każdego rzeczywistego $c \neq 0$, to wszystkie parametry nie są zdefiniowane jednoznacznie. Stąd bez straty ogólności założymy, że dla $\boldsymbol{\Sigma}_i = (\sigma_{rs}^{(i)})$, $\sigma_{pp}^{(i)} = 1$, lub równoważnie dla $\mathbf{V}_i = (v_{rs}^{(i)})$, $v_{TT}^{(i)} = 1$.

Estymacja metodą największej wiarygodności parametrów μ_i , V_i oraz Σ_i , $i = 1, 2, \dots, K$.

Niech

$$\mathbf{X}_{ij} = (\mathbf{x}_{ij1}, \mathbf{x}_{ij2}, \dots, \mathbf{x}_{ijT}),$$

$$\bar{\mathbf{X}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{X}_{ij}, \quad \mathbf{X}_{ijc} = \mathbf{X}_{ij} - \bar{\mathbf{X}}_i \quad \text{oraz}$$

$$\mathbf{X}_{ijc} = \begin{pmatrix} \mathbf{X}_{ijc1} \\ \mathbf{X}_{ijcT} \end{pmatrix}, \quad j = 1, 2, \dots, n_i, \quad i = 1, 2, \dots, K.$$

$$\hat{\boldsymbol{\mu}}_i = \bar{\mathbf{x}}_i = \text{vec}(\bar{\mathbf{X}}_i), \quad (1)$$

$$\hat{\mathbf{V}}_i = \frac{1}{pn_i} \begin{bmatrix} \sum_{j=1}^{n_i} \mathbf{x}'_{ijc1} \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{x}_{ijc1} & \sum_{j=1}^{n_i} \mathbf{x}'_{ijc1} \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{x}_{ijcT} \\ \sum_{j=1}^{n_i} \mathbf{x}'_{ijcT} \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{x}_{ijc1} & \sum_{j=1}^{n_i} \mathbf{x}'_{ijcT} \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{x}_{ijcT} \end{bmatrix} = \frac{1}{pn_i} \sum_{j=1}^{n_i} \mathbf{x}'_{ijc} \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{x}_{ijc}, \quad (2)$$

$$\sum_{j=1}^{n_i} \mathbf{x}'_{ijcT} \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{x}_{ijcT} = pn_i, \quad (3)$$

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{Tn_i} \sum_{j=1}^{n_i} \mathbf{x}_{ijc} \hat{\mathbf{V}}_i^{-1} \mathbf{x}'_{ijc}, \quad i = 1, 2, \dots, K. \quad (4)$$

Równania (2) i (4) z warunkiem pobocznym (3) rozwiązujemy iteracyjnie.

Jeżeli $n_i > \max(p, T)$, to iteracyjny proces rozwiązań układu (2) i (4) z warunkiem pobocznym (3) jest zbieżny do jedynego maksimum.

Krok 1. Za wartość początkową macierzy $\hat{\Sigma}_i$ przyjmujemy

$$\tilde{\Sigma}_i = \frac{1}{Tn_i} \sum_{j=1}^{n_i} \sum_{k=1}^T (\mathbf{x}_{ijk} - \bar{\mathbf{x}}_{ik})(\mathbf{x}_{ijk} - \bar{\mathbf{x}}_{ik})' \quad (5)$$

gdzie

$$\bar{\mathbf{x}}_{ik} = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{ijk}, \quad i = 1, 2, \dots, K, \quad k = 1, 2, \dots, T.$$

Krok 2. Zastępując w równaniu (2) $\hat{\Sigma}_i$ przez $\tilde{\Sigma}_i$ postaci (5) wyliczamy macierz $\hat{\mathbf{V}}_i$ i wszystkie jej elementy $\hat{v}_{rs}^{(i)}$ zastępujemy elementami $\hat{v}_{rs}^{(i)} / \hat{v}_{TT}^{(i)}$.

Krok 3. Wyliczamy macierz $\hat{\Sigma}_i$ z równania (4) wykorzystując macierz \hat{V}_i otrzymaną w Kroku 2.

Krok 4. Powtarzamy Kroki 2 i 3 do osiągnięcia zbieżności, przy czym wybieramy następującą regułę stopu. Wyliczamy dwie macierze:

(a) macierz różnic między dwoma kolejnymi rozwiązaniami równania (2) oraz

(b) macierz różnic między dwoma kolejnymi rozwiązaniami równania (4).

Kontynuujemy kolejne kroki iteracyjne aż do momentu, gdy maksymalna bezwzględna wartość elementów macierzy (a) i (b) będą mniejsze od wcześniej obranej wartości.

Twierdzenie 2.

Jeżeli $n_i > \max(p, T)$, to klasyfikator bazujący na próbach uczących wielkości n_1, n_2, \dots, n_K z odpowiednich klas ma postać:

$$\hat{g}_1(\mathbf{x}) = \arg \max_{1 \leq i \leq K} \hat{\delta}_{i1}(\mathbf{x}),$$

gdzie

$$\hat{\delta}_{i1}(\mathbf{x}) = -\frac{1}{2} \left[p \ln |\hat{\mathbf{V}}_i| + T \ln |\hat{\mathbf{\Sigma}}_i| + (\mathbf{x} - \bar{\mathbf{x}}_i)' (\hat{\mathbf{V}}_i^{-1} \otimes \hat{\mathbf{\Sigma}}_i^{-1}) (\mathbf{x} - \bar{\mathbf{x}}_i) - 2 \ln \hat{\pi}_i \right]$$

jest kwadratową funkcją klasyfikującą, $\bar{\mathbf{x}}_i$, $\hat{\mathbf{V}}_i$, $\hat{\mathbf{\Sigma}}_i$ są estymatorami największej wiarygodności przy dodatkowym warunku $v_{TT}^{(i)} = 1$ oraz $\hat{\pi}_i = n_i / \sum_{j=1}^K n_j$ jest oceną prawdopodobieństwa a priori π_i przynależności obserwacji \mathbf{x} do i -tej klasy, $i = 1, 2, \dots, K$.

Model złożonej symetrii (1)

Założmy teraz, że macierze \mathbf{V}_i mają strukturę złożonej symetrii tj.

$$\mathbf{V}_i = (1 - \rho_i)^{-1}[(1 - \rho_i)\mathbf{I}_T + \rho_i\mathbf{1}_T\mathbf{1}'_T],$$

gdzie

$$-\frac{1}{T-1} < \rho_i < 1, \quad i = 1, 2, \dots, K.$$

Wówczas

$$|\mathbf{V}_i| = (1 - \rho_i)[1 + (T - 1)\rho_i]$$

oraz

$$\mathbf{V}_i^{-1} = \mathbf{I}_T - \frac{\rho_i}{1 + (T - 1)\rho_i}\mathbf{1}_T\mathbf{1}'_T, \quad i = 1, 2, \dots, K.$$

Model złożonej symetrii (2)

Zakładając, że $\sigma_{pp}^{(i)} = 1$ otrzymujemy następujący układ równań największej wiarygodności:

$$\hat{\Sigma}_i = \frac{1}{Tn_i} \sum_{j=1}^{n_i} \mathbf{X}_{ijc} \hat{\mathbf{V}}_i^{-1} \mathbf{X}'_{ijc},$$

$$\sum_{j=1}^{n_i} \mathbf{X}_{ijcp} \hat{\mathbf{V}}_i^{-1} \mathbf{X}'_{ijcp} = Tn_i,$$

$$\hat{\rho}_i = \frac{b_i - pTn_i}{p(T-1)Tn_i + b_i},$$

gdzie

$$b_i = \sum_{j=1}^{n_i} \text{tr}(\mathbf{1}_T \mathbf{1}'_T \mathbf{A}_j^{(i)}),$$

$$\mathbf{A}_j^{(i)} = \mathbf{X}'_{ijc} \hat{\Sigma}_i^{-1} \mathbf{X}_{ijc}, \quad i = 1, 2, \dots, K.$$

Model złożonej symetrii (3)

Układ ten rozwiązujemy iteracyjnie. Ocenę $\hat{\mathbf{V}}_i$ uzyskujemy z macierzy \mathbf{V}_i po zastąpieniu ρ_i przez $\hat{\rho}_i$, $i = 1, 2, \dots, K$. Ponieważ $\hat{\Sigma}_i > 0$ z pr. 1, to $\mathbf{A}_j^{(i)} = \mathbf{X}'_{ijc} \hat{\Sigma}_i^{-1} \mathbf{X}_{ijc} \geq 0$ z pr. 1 oraz $\mathbf{1}_T \mathbf{1}'_T \mathbf{A}_j^{(i)} \geq 0$ i $\text{tr}(\mathbf{1}_T \mathbf{1}'_T \mathbf{A}_j^{(i)}) \geq 0$ z pr. 1 (Abadir i Magnus (2005), s. 338).

Stąd zawsze $b_i \geq 0$ w prawdopodobieństwie równym 1. Jeżeli $b_i > 0$, to

$$\hat{\rho}_i \in \left(-\frac{1}{T-1}, 1 \right).$$

Jako wartość startową parametru $\hat{\rho}_i$ przyjmujemy

$$\hat{\rho}_i^{(0)} = (\mathbf{1}'_T \mathbf{H}_i \mathbf{1}_T - \text{tr}(\mathbf{H}_i)) / (T(T-1)),$$

gdzie

$$\mathbf{H}_i = \frac{1}{p} \sum_{j=1}^p \mathbf{H}_{ij}$$

a \mathbf{H}_{ij} jest macierzą korelacji z n_i elementowej próby z i -tej klasy między T punktami czasowymi dla j -tej cechy, $i = 1, 2, \dots, K$, $j = 1, 2, \dots, p$.

Twierdzenie 3.

Jeżeli $\mathbf{V}_i > 0$ ma strukturę symetrii złożonej, $\boldsymbol{\Sigma}_i > 0$, oraz $n_i > \max(p, T)$, to klasyfikator bazujący na próbach uczących wielkości n_1, n_2, \dots, n_K z odpowiednich klas ma postać:

$$\hat{g}_2(\mathbf{x}) = \arg \max_{1 \leq i \leq K} \hat{\delta}_{i2}(\mathbf{x}),$$

gdzie

$$\begin{aligned} \hat{\delta}_{i2}(\mathbf{x}) = & -\frac{1}{2} \{ p \ln \{ (1 - \hat{\rho}_i)^{-1} [1 + (T - 1) \hat{\rho}_i] \} + T \ln |\hat{\boldsymbol{\Sigma}}_i| + \\ & + (\mathbf{x} - \bar{\mathbf{x}}_i)' \left[(\mathbf{I}_T - \frac{\hat{\rho}_i}{1 + (T - 1) \hat{\rho}_i} \mathbf{I}_T \mathbf{I}'_T) \otimes \hat{\boldsymbol{\Sigma}}_i^{-1} \right] (\mathbf{x} - \bar{\mathbf{x}}_i) - 2 \ln \hat{\pi}_i \}, \end{aligned}$$

jest kwadratową funkcją klasyfikującą, gdzie $\hat{\pi}_i = n_i / \sum_{j=1}^K n_j$.

Model autoregresji rzędu pierwszego (1)

Założmy teraz, że macierze kowariancyjne \mathbf{V}_i mają strukturę autoregresyjną rzędu pierwszego AR(1), tj.

$$\mathbf{V}_i = (1 - \rho_i^2)^{-1} (\rho_i^{|r-s|})_{r,s=1}^T,$$

gdzie $\rho_i \in (-1, 1)$, $i = 1, 2, \dots, K$.

Wówczas

$$\begin{aligned} |\mathbf{V}_i| &= (1 - \rho_i^2)^{-1}, \\ \mathbf{V}_i^{-1} &= \mathbf{I}_T + \rho_i^2 \mathbf{K}_1 - \rho_i \mathbf{K}_2, \end{aligned}$$

gdzie

$$\mathbf{K}_1 = \text{diag}(0, \mathbf{1}'_{T-2}, 0)$$

oraz \mathbf{K}_2 jest macierzą trójdagonalną stopnia T z zerami na głównej przekątnej oraz jedynekami na nadprzekątnej i podprzekątnej.

Model autoregresji rzędu pierwszego (2)

Jeżeli założymy, że $\sigma_{pp}^{(i)} = 1$, to równania największej wiarygodności są postaci

$$\hat{\Sigma}_i = \frac{1}{Tn_i} \sum_{j=1}^{n_i} \mathbf{x}_{ijc} \hat{\mathbf{V}}_i^{-1} \mathbf{x}'_{ijc},$$

$$\sum_{j=1}^{n_i} \mathbf{x}_{ijcp} \hat{\mathbf{V}}_i^{-1} \mathbf{x}'_{ijcp} = Tn_i,$$

$$2c_{i1}\hat{\rho}_i^3 - c_{i2}\hat{\rho}_i^2 - 2(pn_i + c_{i1})\hat{\rho}_i + c_{i2} = 0,$$

gdzie

$$c_{i1} = \sum_{j=1}^{n_i} \text{tr}(\mathbf{K}_1 \mathbf{A}_j^{(i)}), \quad c_{i2} = \sum_{j=1}^{n_i} \text{tr}(\mathbf{K}_2 \mathbf{A}_j^{(i)}), \quad i = 1, 2, \dots, K.$$

Model autoregresji rzędu pierwszego (3)

Układ ten rozwiązujemy iteracyjnie. Ocenę $\hat{\mathbf{V}}_i$ uzyskujemy z macierzy \mathbf{V}_i po zastąpieniu ρ_i przez $\hat{\rho}_i$, $i = 1, 2, \dots, K$.

Wartość startową ρ_i wyliczamy następująco. Wyliczamy macierz

$$\mathbf{H}_i = \frac{1}{p} \sum_{j=1}^p \mathbf{H}_{ij},$$

gdzie \mathbf{H}_{ij} jest macierzą korelacji z n_i elementowej próby z i -tej klasy pomiędzy T punktami czasowymi dla j -tej cechy, $i = 1, 2, \dots, K$, $j = 1, 2, \dots, p$.

Niech $\rho_{1*}^{(i)}$ oznacza wartość średnią elementów pierwszej nadprzekątnej macierzy \mathbf{H}_i , niech $\rho_{2*}^{(i)}$ oznacza wartość średnią elementów drugiej nadprzekątnej macierzy \mathbf{H}_i , i tak dalej. Za wartość startową parametru ρ_i przyjmujemy

$$\hat{\rho}_i^{(0)} = \left[\frac{1}{T-1} \sum_{m=1}^{T-1} (\rho_{m*}^{(i)})^{(T-1)/m} \right]^{1/(T-1)}, \quad i = 1, 2, \dots, K.$$

Twierdzenie 4.

Jeżeli $\mathbf{V}_i > 0$ ma strukturę autokorelacji rzędu pierwszego, $\boldsymbol{\Sigma}_i > 0$, oraz $n_i > \max(p, T)$, to klasyfikator bazujący na próbach uczących wielkości n_1, n_2, \dots, n_K z odpowiednich klas ma postać:

$$\hat{g}_3(\mathbf{x}) = \arg \max_{1 \leq i \leq K} \hat{\delta}_{i3}(\mathbf{x}),$$

gdzie

$$\begin{aligned} \hat{\delta}_{i3}(\mathbf{x}) = & -\frac{1}{2} \{ T \ln |\hat{\boldsymbol{\Sigma}}_i| - p \ln(1 - \hat{\rho}_i^2) + \\ & + (\mathbf{x} - \bar{\mathbf{x}}_i)' \left[(\mathbf{I}_T + \hat{\rho}_i^2 \mathbf{K}_1 - \hat{\rho}_i \mathbf{K}_2) \otimes \hat{\boldsymbol{\Sigma}}_i^{-1} \right] (\mathbf{x} - \bar{\mathbf{x}}_i) - 2 \ln \hat{\pi}_i \}, \end{aligned}$$

gdzie $\hat{\pi}_i = n_i / \sum_{j=1}^K n_j$, $i = 1, 2, \dots, K$.

Testy (1)

Chcemy zweryfikować hipotezę

$$H_{0i} : \Omega_i = \mathbf{V}_i \otimes \Sigma_i \quad \text{przeciwko hipotezie}$$

$$H_{ai} : \Omega_i \quad \text{ma dowolną strukturę kowariancyjną.}$$

Przy prawdziwości hipotezy alternatywnej H_{ai} i warunku $n_i > pT$, estymator największej wiarygodności macierzy Ω_i ma postać

$$\hat{\Omega}_i = \frac{1}{n_i} \sum_{j=1} n_j (\text{vec} \mathbf{X}_{jc}) (\text{vec} \mathbf{X}_{jc})'.$$

Wówczas

$$-2 \ln \Lambda_{i1} = n_i (p \ln |\hat{\mathbf{V}}_i| + T \ln |\hat{\Sigma}_i| - \ln |\hat{\Omega}_i|),$$

gdzie Λ_{i1} jest ilorazem wiarygodności, natomiast $\hat{\mathbf{V}}_i$ oraz $\hat{\Sigma}_i$ są estymatorami największej wiarygodności uzyskanymi przy warunku $v_{TT}^{(i)} = 1$ lub $\sigma_{pp}^{(i)} = 1$ oraz $n_i > \max(p, T)$, $i = 1, 2, \dots, K$. Mamy

$$-2 \ln \Lambda_{i1} \sim \chi_{\frac{1}{2}pT(pT+1) - \frac{1}{2}p(p+1) - \frac{1}{2}T(T+1)+1}^2, \quad i = 1, 2, \dots, K.$$

Chcemy zweryfikować hipotezę

$H_{0i} : \mathbf{V}_i$ ma strukturę złożonej symetrii, przeciwko

$H_{ai} : \mathbf{V}_i$ ma dowolną strukturę kowariancyjną.

Mamy

$$-2 \ln \Lambda_{i2} = n_i [p(\ln |\tilde{\mathbf{V}}_i| - \ln |\hat{\mathbf{V}}_i|) + T(\ln |\tilde{\boldsymbol{\Sigma}}_i| - \ln |\hat{\boldsymbol{\Sigma}}_i|)],$$

gdzie Λ_{i2} jest ilorazem wiarygodności oraz $\tilde{\mathbf{V}}_i$ i $\tilde{\boldsymbol{\Sigma}}_i$ są ocenami największej wiarygodności macierzy \mathbf{V}_i i $\boldsymbol{\Sigma}_i$ przy założeniu, że prawdziwa jest hipoteza zerowa H_{0i} , natomiast $\hat{\mathbf{V}}_i$ i $\hat{\boldsymbol{\Sigma}}_i$ są ocenami największej wiarygodności macierzy \mathbf{V}_i i $\boldsymbol{\Sigma}_i$ przy założeniu, że prawdziwa jest hipoteza alternatywna H_{ai} . Ponadto wszystkie te oceny uzyskano przy założeniu, że $v_{TT}^{(i)} = 1$ lub $\sigma_{pp}^{(i)} = 1$ oraz $n_i > \max(p, T)$.

Mamy

$$-2 \ln \Lambda_{i2} \sim \chi_{\frac{1}{2}T(T+1)-1}^2.$$

Testy (3)

Chcemy zweryfikować hipotezę

$H_{0i} : \mathbf{V}_i$ ma strukturę autoregresyjną pierwszego rzędu, przeciwko

$H_{ai} : \mathbf{V}_i$ ma dowolną strukturę kowariancyjną.

Mamy

$$-2 \ln \Lambda_{i3} = n_i [p(\ln |\bar{\mathbf{V}}_i| - \ln |\hat{\mathbf{V}}_i|) + T(\ln |\bar{\boldsymbol{\Sigma}}_i| - \ln |\hat{\boldsymbol{\Sigma}}_i|)],$$

gdzie Λ_{i3} jest ilorazem wiarygodności oraz $\bar{\mathbf{V}}_i$ i $\bar{\boldsymbol{\Sigma}}_i$ są ocenami największej wiarygodności macierzy \mathbf{V}_i i $\boldsymbol{\Sigma}_i$ przy założeniu, że prawdziwa jest hipoteza zerowa H_{0i} , natomiast $\hat{\mathbf{V}}_i$ i $\hat{\boldsymbol{\Sigma}}_i$ są ocenami największej wiarygodności macierzy \mathbf{V}_i i $\boldsymbol{\Sigma}_i$ przy założeniu, że prawdziwa jest hipoteza alternatywna H_{ai} . Ponadto wszystkie te oceny uzyskano przy założeniu, że $v_{TT}^{(i)} = 1$ lub $\sigma_{pp}^{(i)} = 1$ oraz $n_i > \max(p, T)$.

Mamy

$$-2 \ln \Lambda_{i3} \sim \chi_{\frac{1}{2}T(T+1)-1}^2.$$

Przykład 1.

Przeprowadzono eksperyment laboratoryjny, badający efekt wzrostu zmodyfikowanej trawy odmiany Paspalum zainfekowanej grzybami i hodowanej w czterech różnych temperaturach (14, 18, 22, 26°C). Tę zmodyfikowaną grupę porównano z grupą kontrolną. Każda z 6 doniczek z trawą podlegała pomiarowi wartości trzech cech:

x_1 = świeża masa korzeni (g),

x_2 = maksymalna długość korzenia (mm),

x_3 = świeża masa części nadziemnych trawy (g).





Dane zostały zaczerpnięte z monografii Sebera (1984). W naszym przypadku $K = 2$, $p = 3$, $T = 4$, $n_1 = n_2 = 6$. Ponieważ $n_i < pT$, to przyjmujemy, że $\Omega_i = \mathbf{V}_i \otimes \Sigma_i$, $i = 1, 2$. Uzyskano następujące wyniki klasyfikacji:

	1	2
1	6	0
2	0	6

Przykład 2.

Dane dotyczą pomiarów dwóch cech mierzonych w trzech momentach czasowych u każdej z sześciu osób, w każdej z trzech klas (2 grupy chorobowe i grupa kontrolna) i pochodzą z podręcznika SAS Institute Inc (1990, example 9, p. 988). W naszym przypadku $K = 3$, $p = 2$, $T = 3$, $n_1 = n_2 = n_3 = 6$. Uzyskano następujące wyniki klasyfikacji:

	1	2	3
1	6	0	0
2	1	5	0
3	0	0	6

-  Abadir K.M. and Magnus J.R., Matrix Algebra, Cambridge University Press, New York 2005.
-  SAS procedures guide, Version 6, SAS Institute Inc., Cary NC, 1990.
-  Seber G.A.F., Multivariate Observations, Wiley, New York 1984.
-  Srivastava M.S., von Rosen T., and von Rosen D., Models with a Kronecker product covariance structure: estimation and testing, Math. Methods Stat. 17(4) 2008, 357–370.