

Mateusz John

**ESTYMACJA I TESTOWANIE
MACIERZY KOWARIANCJI
NALEŻĄCYCH DO PODPRZESTRZENI
KWADRATOWYCH**

Praca przygotowana
w Instytucie Matematyki
Politechniki Poznańskiej
pod kierunkiem
dr hab. inż. Katarzyny Filipiak
z Instytutu Matematyki
Politechniki Poznańskiej
i przedłożona jako rozprawa doktorska
Radzie naukowej dyscyplin
matematyka i informatyka
Uniwersytetu im. Adama Mickiewicza
w Poznaniu

Dziedzina: nauki ścisłe i przyrodnicze

Dyscyplina: matematyka

Poznań, 2024

Podziękowania

W pierwszej kolejności chciałbym podziękować Rodzicom i Dziadkowi za wsparcie na każdym kroku podczas pracy naukowej jak i podczas pisania niniejszej rozprawy. Za codzienną opiekę i poświęcenia w dbaniu o mój naukowy rozwój. Dziękuję za wszelką otrzymaną pomoc i motywację do ciągłej wyężonej pracy.

Równie mocno pragnę podziękować mojej Pani Promotor, dr hab. inż. Katarzynie Filipiak za nieocenioną pomoc w czasie pisania mojej rozprawy. Pomoc Pani Promotor okazała się być nieoceniona, a za poświęcony czas i cierpliwość jestem bardzo wdzięczny. Dziękuję za każdą, nawet drobną sugestię, przyczyniającą się do mojego rozwoju naukowego.

Dziękuję również wszystkim uczestnikom Środowiskowego Seminarium z Algebry Liniowej i jej Zastosowań, organizowanego przez Pana prof. Augustyna Markiewicza, Panią dr hab. inż. Katarzynę Filipiak z Politechniki Poznańskiej, a także Pana prof. Jana Hauke z Uniwersytetu im. Adama Mickiewicza w Poznaniu, za każdą uwagę, wsparcie i dobre słowo, które okazało się być cenne. Dziękuję również uczestnikom międzynarodowej grupie badawczej "Multivariate and Mixed Linear Models", organizowanej przez Pana prof. Augustyna Markiewicza oraz Panią dr hab. inż. Katarzynę Filipiak, jak i osobom, z którymi jestem współautorem prac naukowych. Dziękuję za każdą wskazówkę, życzliwość i możliwość rozwoju naukowego.

Dziękuję także byłym i obecnym Koleżankom i Kolegom z Instytutu Matematyki Politechniki Poznańskiej za wsparcie, zrozumienie i motywację do wytrwałej pracy.

Spis treści

Wykaz ważniejszych skrótów i oznaczeń	9
Wstęp	11
1 Wprowadzenie	15
1.1 Model	15
1.2 Struktury kowariancyjne	19
1.3 Testowanie hipotez	24
1.4 Wybrane przekształcenia macierzy i ich własności	28
1.5 Przykład rzeczywisty	34
2 Estymacja macierzy kowariancji o zadanej strukturze, należącej do podprzestrzeni kwadratowej	37
2.1 Estymacja macierzy kowariancji z wykorzystaniem operatora rzutu	38
2.2 Estymacja szczególnych struktur macierzy kowariancji	41
2.2.1 Standardowy model wielowymiarowy	41
2.2.2 Model podwójnie wielowymiarowy	43
2.2.3 Model krzywych wzrostu	46
3 Testowanie hipotez dotyczących struktur kowariancji należących do podprzestrzeni kwadratowych w modelu podwójnie wielowymiarowym	49
3.1 Model i hipoteza	49
3.2 Statystyki testowe	51
3.2.1 Test ilorazu wiarygodności	52
3.2.2 Test wynikowy Rao	61

3.3	Porównanie LRT i RST – badania symulacyjne	68
3.3.1	Zbieżność do rozkładu granicznego	68
3.3.2	Analiza mocy testów	71
3.4	Przykład rzeczywisty	72
4	Testowanie niezależności międzyklasowej w modelu z macierzą kowa- riancji o blokowej strukturze kompletnej symetrii	77
4.1	Model i hipotezy	78
4.2	Statystyki testowe	79
4.2.1	Test F	79
4.2.2	Test Roya	80
4.2.3	Test ilorazu wiarygodności	81
4.2.4	Test wynikowy Rao	84
4.2.5	Test Walda	91
4.3	Badania symulacyjne	93
4.3.1	Zbieżność do rozkładu granicznego	95
4.3.2	Odporność testów na zaburzenia rozkładu normalnego	98
4.3.3	Analiza mocy testów	99
4.3.4	Przykład rzeczywisty	103
5	Podsumowanie	109
	Bibliografia	113

Streszczenie

Celem rozprawy doktorskiej jest zaproponowanie testów dotyczących struktur kowariancyjnych w modelach podwójnie wielowymiarowych. Ze względu na hierarchiczny charakter rozważanych eksperymentów odpowiednimi strukturami są macierze blokowe. W pracy rozważa się struktury blokowe należące do podprzestrzeni kwadratowych. Proponowane testy obejmują test ilorazu wiarygodności, test wynikowy Rao oraz test Walda. Wymienione testy porównywane są ze sobą ze względu na szybkość zbieżności do granicznego rozkładu chi-kwadrat oraz moc. Do porównania użyto metod symulacyjnych. Ponadto, ponieważ w każdym rozważanym teście istotną rolę odgrywają estymatory największej wiarygodności nieznanymi parametrami, w pracy pokazano, że estymatory te można uzyskać przez rzutowanie na odpowiednią podprzestrzeń kwadratową. Przedstawione wyniki zilustrowano na przykładzie danych rzeczywistych.

Abstract

The aim of the Ph. D. dissertation is to propose tests for covariance structures in doubly multivariate models. Due to the hierarchical nature of the considered experiments, block matrices are appropriate structures. The work considers block structures belonging to the quadratic subspaces. The proposed tests include the likelihood ratio test, the Rao score test and the Wald test. The mentioned tests are compared with each other in terms of the speed of convergence to the limiting chi-square distribution and the power. For the comparison simulation methods were used. Moreover, since the maximum likelihood estimators of unknown parameters have an important role in each considered test, the work shows that these estimators can be obtained by projecting onto an appropriate quadratic subspace. Presented results are illustrated using real data example.

Wykaz ważniejszych skrótów i oznaczeń

$\text{vec}(\bullet)$	–	operator wektoryzowania macierzy
\otimes	–	iloczyn Kroneckera
$\text{BTr}(\bullet)$	–	operator śladu blokowego
$\text{BSum}(\bullet)$	–	operatr sumy blokowej
$\text{PTr}(\bullet)$	–	operator śladu częściowego
$\langle \bullet, \bullet \rangle$	–	iloczyn skalarny
\mathbf{P}_A	–	operator rzutu ortogonalnego na przestrzeń kolumn macierzy \mathbf{A} ; $\mathbf{P}_A = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$
\mathbf{Q}_A	–	operator rzutu ortogonalnego na dopełnienie ortogonalne przestrzeni kolumn macierzy \mathbf{A} ; $\mathbf{Q}_A = \mathbf{I} - \mathbf{P}_A$
Σ, Ω, Ξ	–	macierze kowariancji
μ	–	wektor średnich
$\mathbf{K}_{m,n}$	–	macierz komutacji stopnia mn
\mathbf{D}_m	–	macierz duplikacji wymiaru $m \times m(m+1)/2$
I, BI	–	struktura proporcjonalna do macierzy jednostkowej i jej blokowy odpowiednik
D, BD	–	struktura diagonalna i blokowo diagonalna
CS, BCS	–	struktura kompletnej symetrii i blokowej kompletnej symetrii
CT, BCT	–	struktura kołowej macierzy Toeplitza i blokowej kołowej macierzy Toeplitza
LRT	–	test ilorazu wiarygodności
RST	–	test wynikowy Rao
WT	–	test Walda
FT	–	F-test
RLRT	–	test Roya
\mathbf{s}	–	wektor wynikowy
\mathbf{F}	–	macierz informacji Fishera
f_E	–	entropijna funkcja straty
η	–	skorygowana rozbieżność

Wstęp

W obecnych czasach, w wielu dziedzinach nauki, gromadzone są ogromne zestawy danych. Podczas eksperymentów w fizyce, ekonomii, biologii, genetyce, informatyce itd. na każdej jednostce eksperymentalnej mierzonych jest wiele cech, przy czym pomiary te mogą być powtarzane w różnych lokalizacjach, punktach czasowych, na różnych wysokościach, czy też głębokościach. Właściwym modelem eksperymentów, w którym mierzonych jest wiele cech, jest model wielowymiarowy. W przypadku, gdy pomiary są powtarzane, do analizy eksperymentu wykorzystuje się model podwójnie wielowymiarowy. W obu przypadkach zależności między cechami i/lub powtórzeniami wyrazić można za pomocą szczególnej struktury kowariancyjnej.

W wielu eksperymentach liczba jednostek eksperymentalnych jest mniejsza od liczby badanych cech, czego konsekwencją jest nieefektywna/błędna analiza statystyczna. Mówi się wówczas, że wystąpił problem wysokiej-wymiarowości danych (ang. *high-dimensionality*). W takich przypadkach klasyczne metody wnioskowania zazwyczaj zawodzą. Jednym ze sposobów rozwiązania problemu wysoko-wymiarowości jest nałożenie pewnych ograniczeń na rozważany model statystyczny, skutkujący zmniejszeniem liczby nieznanymi parametrów. W literaturze zazwyczaj ograniczenia te dotyczą wartości oczekiwanej lub macierzy kowariancji. W niniejszej pracy rozważony jest ten drugi przypadek. Istotne jest tutaj zweryfikowanie, jaka struktura kowariancyjna najlepiej odzwierciedla zależności pomiędzy zmiennymi i w jaki sposób ją wyestymować. Do identyfikacji struktury kowariancyjnej można wykorzystać metody graficzne, algebraiczne lub statystyczne, czyli testowanie hipotez.

Głównymi celami niniejszej rozprawy jest zaproponowanie metod estymacji struktur kowariancyjnych należących do podprzestrzeni kwadratowych oraz testów statystycznych pozwalających na odpowiedni dobór struktury kowariancyjnej. Omówione testy zostały porównane między sobą oraz z testami statystycznymi rozważanymi w literaturze ze względu na szybkość zbieżności, odporność na zaburzenie normalności rozkładu obserwacji oraz moc.

Rozprawa składa się z czterech rozdziałów i podsumowania. Pierwszy rozdział zawiera wprowadzenie do rozważanych zagadnień. Zaprezentowano w nim najważniejsze definicje, które mają istotne znaczenie w dalszej części pracy, zaczynając od zdefiniowania modelu i przedstawienia struktur kowariancyjnych należących do podprzestrzeni kwadratowych, obejmujących między innymi blokową strukturę diagonalną, blokową strukturę kompletnej symetrii, czy też blokową strukturę kołowej macierzy Toeplitza. Przedstawiono również postać statystyk testowych takich jak statystyka ilorazu wiarygodności, statystyka wynikowa Rao oraz statystyka Walda. Zdefiniowano także wykorzystywane w rozprawie wybrane przekształcenia macierzy takie jak operatory śladu blokowego, sumy blokowej, czy śladu częściowego, jak również zasady różniczkowania macierzy. Na koniec opisano rzeczywisty eksperyment, który posłużył do zilustrowania głównych wyników rozprawy.

Kolejne rozdziały rozprawy zawierają główne wyniki pracy. W Rozdziale 2 dla modelu z nieznanymi wartością oczekiwaną i macierzą kowariancji podano warunki konieczne i dostateczne na to, aby estymator największej wiarygodności liniowej struktury kowariancyjnej należącej do podprzestrzeni kwadratowej miał jawną postać oraz udowodniono, że estymator ten może być reprezentowany przez rzut estymatora największej wiarygodności macierzy kowariancji bez struktury na przestrzeń macierzy o rozważanej strukturze. Przedstawione twierdzenia zilustrowano przykładami estymacji struktur kowariancyjnych w standardowym modelu wielowymiarowym, modelu podwójnie wielowymiarowym oraz modelu krzywych wzrostu.

W Rozdziale 3 rozważono standardowy model podwójnie wielowymiarowy, dla którego macierz kowariancji o strukturze blokowej należy do podprzestrzeni kwadratowej. Zaproponowano testy statystyczne weryfikujące należąca do komutatywnej podprzestrzeni kwadratowej strukturę zależności między- i wewnątrzklasowych. Wyprowadzono postać statystyki ilorazu wiarygodności wraz z jej dokładnym rozkładem prawdopodobieństwa, a także wyznaczono postać statystyki wynikowej Rao. Przy wykorzystaniu badań symulacyjnych porównano szybkość zbieżności powyższych statystyk testowych do rozkładu granicznego oraz moce obu testów. Na koniec

przedstawione testy zilustrowano przykładem rzeczywistym.

Rozdział 4 dotyczy standardowego modelu podwójnie wielowymiarowego z blokową strukturą kompletnej symetrii macierzy kowariancji. Dla testu hipotezy o niezależności międzyklasowej wyprowadzono statystykę ilorazu wiarygodności wraz z jej dokładnym rozkładem, wyznaczono statystykę testu wynikowego Rao oraz testu Walda. W badaniach symulacyjnych porównano wspomniane statystyki testowe ze statystykami istniejącymi w literaturze pod kątem szybkości zbieżności do rozkładu granicznego, odporności na zaburzenie normalności rozkładu obserwacji oraz moc. Przedstawione testy zilustrowano danymi rzeczywistymi.

Ostatnia część rozprawy to podsumowanie i analiza najważniejszych wniosków z pracy.

1 Wprowadzenie

1.1 Model

W pracy zostały rozważone eksperymenty, w których na n jednostkach doświadczalnych mierzonych jest m cech, które mogą być powtórzone q -krotnie w czasie, na różnych głębokościach, wysokościach, czy lokalizacjach. Odpowiednimi modelami takich eksperymentów może być standardowy model wielowymiarowy (gdy $m = 1$ lub $q = 1$) bądź model podwójnie wielowymiarowy ($m, q \geq 2$).

Niech $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{im})'$ będzie m -wymiarowym wektorem obserwacji cech na i -tej jednostce doświadczalnej, gdzie $i = 1, \dots, n$. Dla standardowego modelu wielowymiarowego (z $q = 1$) macierz obserwacji przedstawić można w postaci

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}'_1 \\ \mathbf{y}'_2 \\ \vdots \\ \mathbf{y}'_n \end{pmatrix}_{n \times m}.$$

W przypadku modeli podwójnie wielowymiarowych macierz obserwacji jest naturalnym rozszerzeniem przypadku wielowymiarowego, gdzie zamiast jednej obserwacji każdej cechy (na każdym obiekcie) pomiary powtarzane są q -krotnie, np. w różnych lokalizacjach. Wówczas obserwacje mogą być wyrażone poprzez trójindeksową macierz \mathcal{Y} (tensor 3-go rzędu) wymiaru $n \times m \times q$, jak na Rysunku 1.

$$\mathcal{Y} = \begin{pmatrix} y_{11k} & \cdots & y_{1qk} \\ \vdots & \ddots & \vdots \\ y_{n1k} & \cdots & y_{nqk} \end{pmatrix} \quad k = 1, \dots, m$$

Rysunek 1: Wizualizacja tensora $\mathcal{Y} \in \mathbb{R}_{n \times q \times m}$.

Tensor \mathcal{Y} można również przedstawić jak na Rysunku 2: $\mathcal{Y}^{(1)}$ przedstawia n macierzy \mathbf{Y}_i wymiaru $m \times q$, ustawionych jedna pod drugą, $\mathcal{Y}^{(2)}$ przedstawia q macierzy \mathbf{Y}_j wymiaru $n \times m$ ustawionych jedna obok drugiej; natomiast $\mathcal{Y}^{(3)}$ przedstawia m macierzy \mathbf{Y}_k o wymiarach $n \times q$ ułożonych jedna za drugą. Wizualizacje te pozwalają w prosty sposób sprowadzić tensor \mathcal{Y} do postaci macierzowej, to znaczy wektoryzując macierze \mathbf{Y}_i , \mathbf{Y}_j oraz \mathbf{Y}_k z powyższych wizualizacji oraz transponując odpowiednie wektory, macierz trójindeksowa \mathcal{Y} może zostać przekształcona do macierzy o wymiarach odpowiednio: $n \times mq$, $q \times nm$, $m \times nq$. Istotnie,

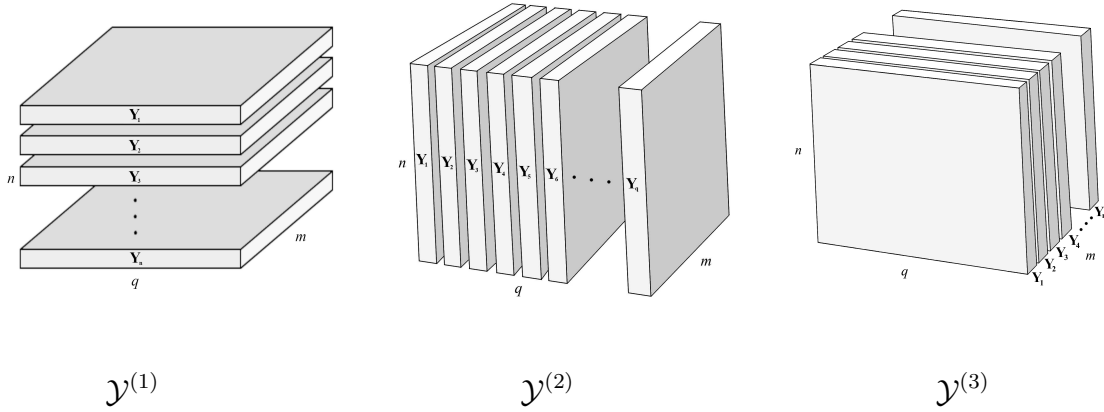
$$\mathcal{Y}^{(1)} : \begin{pmatrix} \text{vec}' \mathbf{Y}_1 \\ \text{vec}' \mathbf{Y}_2 \\ \vdots \\ \text{vec}' \mathbf{Y}_n \end{pmatrix} \in \mathbb{R}_{n \times mq}, \quad \mathbf{Y}_i \in \mathbb{R}_{m \times q}, i = 1, \dots, n,$$

$$\mathcal{Y}^{(2)} : \begin{pmatrix} \text{vec}' \mathbf{Y}_1 \\ \text{vec}' \mathbf{Y}_2 \\ \vdots \\ \text{vec}' \mathbf{Y}_q \end{pmatrix} \in \mathbb{R}_{q \times nm}, \quad \mathbf{Y}_j \in \mathbb{R}_{n \times m}, j = 1, \dots, q,$$

$$\mathcal{Y}^{(3)} : \begin{pmatrix} \text{vec}' \mathbf{Y}_1 \\ \text{vec}' \mathbf{Y}_2 \\ \vdots \\ \text{vec}' \mathbf{Y}_m \end{pmatrix} \in \mathbb{R}_{m \times nq}, \quad \mathbf{Y}_k \in \mathbb{R}_{n \times q}, k = 1, \dots, m,$$

gdzie „vec” jest operatorem wektoryzowania macierzy, czyli ustawienia jej kolumn jedna pod drugą.

W dalszej części rozprawy wykorzystywana będzie transformacja przekształcająca w macierz tensor $\mathcal{Y}^{(1)}$. Dla ułatwienia zapisu oznaczana ona będzie przez \mathbf{Y} , a rozróżnienie między modelem wielowymiarowym a podwójnie wielowymiarowym wynikać będzie z kontekstu rozważanego zagadnienia.



Rysunek 2: Wizualizacja danych trójwymiarowych z wykorzystaniem macierzy.

Założmy, że każdy z n niezależnych wektorów losowych $\mathbf{y}_i = \text{vec } \mathbf{Y}_i$, $i = 1, \dots, n$, ma rozkład normalny z wartością oczekiwaną $\mathbb{E}(\mathbf{y}_i) = \text{vec } \mathbf{M}$, gdzie \mathbf{M} jest dowolną macierzą wymiaru $m \times q$, i z macierzą kowariancji (wariancji-kowariancji, dyspersji) $\mathbb{D}(\mathbf{y}_i) = \mathbf{\Omega}$, gdzie $\mathbf{\Omega} \in \mathbb{R}_{qm \times qm}$. Przyjmując oznaczenie $\text{vec } \mathbf{M} = \boldsymbol{\mu}$ możemy zapisać, że macierz obserwacji $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)'$ ma macierzowy rozkład normalny postaci

$$\mathbf{Y} \sim N_{n,mq}(\mathbf{1}_n \boldsymbol{\mu}', \mathbf{I}_n, \mathbf{\Omega}), \quad (1.1)$$

gdzie $\mathbf{1}_n$ jest n -wymiarowym wektorem o składowych równych 1, a \mathbf{I}_n jest macierzą jednostkową stopnia n . Zgodnie z definicją macierzowego rozkładu normalnego $\mathbb{D}(\mathbf{Y}) = \mathbb{D}(\text{vec } \mathbf{Y}) = \mathbf{\Omega} \otimes \mathbf{I}_n$, gdzie \otimes oznacza iloczyn Kroneckera; por. Kollo, von Rosen (2005). Ponadto, dla $q = 1$ (lub $m = 1$) model (1.1) redukuje się do standardowego modelu wielowymiarowego, który bez straty ogólności zapiszemy w postaci

$$\mathbf{Y} \sim N_{n,m}(\mathbf{1}_n \boldsymbol{\mu}', \mathbf{I}_n, \mathbf{\Omega}), \quad (1.2)$$

gdzie $\boldsymbol{\mu}$ jest wektorem m -wymiarowym, a $\mathbf{\Omega}$ macierzą stopnia m .

Głównym celem rozprawy jest estymacja i testowanie struktury macierzy kowariancji $\mathbf{\Omega}$. W modelu (1.1), w którym $\mathbf{\Omega}$ jest dowolną macierzą symetryczną, określoną

dotadnio, jednymi z najczęściej wykorzystywanych estymatorów macierzy kowariancji jest macierz kowariancji z próby, $\mathbf{S}_1 = \frac{1}{n-1} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$, oraz estymator największej wiarogodności (ang. *maximum likelihood estimator* - MLE) macierzy kowariancji $\mathbf{\Omega}$, postaci $\mathbf{S} = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$, gdzie $\mathbf{Q}_n = \mathbf{Q}_{1_n}$ z \mathbf{Q}_A będącym operatorem rzutu ortogonalnego na dopełnienie ortogonalne przestrzeni kolumn macierzy \mathbf{A} wymiaru $n_1 \times n_2$, to jest $\mathbf{Q}_A = \mathbf{I}_{n_1} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$.

Zauważmy, że jeśli $mq \geq n$, macierz \mathbf{S} jest osobliwa, natomiast dla n bliskiego mq jest macierzą źle uwarunkowaną numerycznie, co może prowadzić do błędnej analizy statystycznej, por. Ledoit, Wolf (2004). Jednym z rozwiązań tego problemu jest założenie, że macierz kowariancji posiada pewną, ściśle określoną, strukturę. Zabieg ten pozwala zredukować liczbę nieznanymi parametrów macierzy kowariancji, $mq(mq + 1)/2$, nawet do liczby niezależnej od liczby cech i powtórzeń. Jeśli dysponujemy informacjami a priori dotyczącymi struktury kowariancyjnej, związanymi przykładowo ze sposobem przeprowadzania doświadczenia, czy też wiedzą ekspercką, to bez problemu możemy dopasować odpowiednią strukturę. Jeśli jednak takiej wiedzy nie posiadamy, należy zaproponować metody identyfikacji takiej struktury; por. Filipiak i in. (2021). Szczególne struktury kowariancyjne, rozważane w niniejszej rozprawie, zostały przedstawione w Rozdziale 1.2.

W literaturze rozważa się również różne struktury wartości oczekiwanej, na przykład:

- model o stałym wektorze wartości oczekiwanych, $\boldsymbol{\mu}$, dla każdego powtórzenia

$$\mathbf{Y} \sim N_{n,mq}(\mathbf{1}_n \boldsymbol{\mu}' (\mathbf{1}'_q \otimes \mathbf{I}_m), \mathbf{I}_n, \mathbf{\Omega}),$$

gdzie $\boldsymbol{\mu}$ jest wektorem m -wymiarowym, por. Liang i in. (2021);

- model krzywych wzrostu, który bez straty ogólności możemy zapisać jako

$$\mathbf{Y} \sim N_{n,m}(\mathbf{X}\mathbf{B}\mathbf{Z}, \mathbf{I}_n, \mathbf{\Omega}), \tag{1.3}$$

gdzie \mathbf{B} jest macierzą nieznanymi parametrów wymiaru $p_1 \times p_2$, przy $p_1 < n$, natomiast \mathbf{X} i \mathbf{Z} są znanymi macierzami układu, wymiaru $n \times p_1$ i $p_2 \times m$, odpowiednio, por. Kollo, von Rosen (2005).

1.2 Struktury kowariancyjne

W niniejszej rozprawie uwzględnione zostaną struktury kowariancyjne, które należą do kwadratowej podprzestrzeni \mathcal{V} . Przypomnijmy, że przestrzeń liniowa \mathcal{V} jest podprzestrzenią kwadratową, jeśli $\Omega \in \mathcal{V}$ implikuje, że $\Omega^2 \in \mathcal{V}$; Seely (1971). Należy również wspomnieć, że podprzestrzeń kwadratowa jest komutatywna, jeżeli dwie dowolne macierze należące do \mathcal{V} komutują.

Jedną z najszerszych podprzestrzeni kwadratowych tworzą macierze symetryczne, jednak ze względu na możliwą dużą liczbę nieznanych parametrów, przedmiotem zainteresowania w niniejszej rozprawie będą pewne szczególne, znane z literatury, struktury kowariancyjne (należące do podprzestrzeni przestrzeni macierzy symetrycznych), między innymi macierz proporcjonalna do macierzy jednostkowej, macierz diagonalna, macierz kompletnej symetrii czy kołowa macierz Toeplitza, zdefiniowane poniżej. Macierze kowariancji o zadanej strukturze oznaczane będą w pracy przez Σ i indeksowane będą odpowiednim skrótem odnoszącym się do rozważanej struktury. Dla uproszczenia, definiując kolejne struktury kowariancyjne przyjmijmy, że $m = 1$.

Najprostszą strukturą kowariancyjną jest macierz proporcjonalna do macierzy jednostkowej (ang. *spherical structure*, *sphericity*) postaci

$$\Sigma_{\text{I}} = c_0 \mathbf{I}_q,$$

gdzie $c_0 \in \mathbb{R}^+$. Jednym z jej uogólnień jest macierz o strukturze diagonalnej (D) postaci

$$\Sigma_{\text{D}} = \text{Diag}(c_1, c_2, \dots, c_q),$$

gdzie $c_i \in \mathbb{R}^+$ dla $i = 1, 2, \dots, q$.

Strukturę Σ_I można również uogólnić na przypadek, gdzie poza główną przekątną nie mamy wartości zerowych, a pewną stałą wartość. Taką strukturę, nazywaną strukturą kompletnej symetrii (ang. *compound symmetry* - CS), można zapisać jako

$$\Sigma_{CS} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \dots & \dots & \rho \\ \rho & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \rho & \dots & \rho & \rho & 1 \end{pmatrix} = \sigma^2 [(1 - \rho)\mathbf{I}_q + \rho\mathbf{1}_q\mathbf{1}'_q], \quad (1.4)$$

gdzie $\sigma^2 \in \mathbb{R}^+$ oraz, w celu zapewnienia dodatniej określoności macierzy o strukturze CS, $\rho \in (-1/(q - 1), 1)$. Ponieważ kowariancja nie zależy od odległości między kolejnymi pomiarami (punktami czasowymi, lokalizacjami itp.), strukturę CS nazywa się również przemienną (ang. *exchangeable*); por. m.in. Koziół i in. (2017), Roy i in. (2018), Žežula i in. (2018). Macierz o strukturze CS można także przedstawić wykorzystując operatory $\mathbf{P}_q = \mathbf{P}_{1_q}$, gdzie $\mathbf{P}_A = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ jest operatorem rzutu ortogonalnego na przestrzeń kolumn macierzy \mathbf{A} oraz $\mathbf{Q}_q = \mathbf{I}_q - \mathbf{P}_q$ jako

$$\Sigma_{CS} = \gamma\mathbf{P}_q + \delta\mathbf{Q}_q, \quad (1.5)$$

gdzie $\gamma, \delta > 0$ oraz, na podstawie parametryzacji (1.4), $\gamma = \sigma^2[1 + (q - 1)\rho]$, $\delta = \sigma^2(1 - \rho)$.

Kolejną interesującą nas strukturą jest struktura kołowej macierzy Toeplitza (ang. *circular Toeplitz* - CT), postaci

$$\Sigma_{CT} = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \dots & \rho_2 & \rho_1 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \ddots & \rho_3 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \ddots & \rho_4 & \rho_3 \\ \rho_3 & \rho_2 & \rho_1 & 1 & \ddots & \rho_5 & \rho_4 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \rho_2 & \rho_3 & \rho_4 & \ddots & \ddots & 1 & \rho_1 \\ \rho_1 & \rho_2 & \rho_3 & \ddots & \ddots & \rho_1 & 1 \end{pmatrix}$$

$$= \begin{cases} \sigma^2 \left[\mathbf{I}_q + \sum_{i=1}^{\lfloor q/2 \rfloor} \rho_i (\mathbf{C}^i + \mathbf{C}'^i) \right] & \text{dla } q \text{ nieparzystego} \\ \sigma^2 \left[\mathbf{I}_q + \sum_{i=1}^{q/2-1} \rho_i (\mathbf{C}^i + \mathbf{C}'^i) + \rho_{q/2} \mathbf{C}^{q/2} \right] & \text{dla } q \text{ parzystego,} \end{cases} \quad (1.6)$$

gdzie $\rho_i \in \mathbb{R}$, $i = 1, \dots, \lfloor q/2 \rfloor$ lub $i = 1, \dots, q/2 - 1$, są dobrane tak, aby zapewnić dodatnią określoność macierzy Σ_{CT} , symbol $\lfloor \cdot \rfloor$ oznacza część całkowitą z liczby, natomiast \mathbf{C} jest macierzą permutacyjną generującą macierze kołowe, to znaczy macierzą o elementach równych zero za wyjątkiem pierwszej przekątnej pod przekątną główną i prawego górnego narożnika, gdzie elementy są równe 1; por. Olkin, Press (1969). Notacja \mathbf{C}^i oznacza i -tą potęgę macierzy \mathbf{C} . Rozważana struktura może być stosowana do opisu zjawisk, dla których zależność pomiędzy równoodległymi od siebie punktami czasowymi/lokalizacjami jest taka sama.

Na podstawie pracy Seely (1971, Lemat 6) wiadomo, że warunkiem koniecznym i wystarczającym, aby przestrzeń \mathcal{V} była komutatywną podprzestrzenią kwadratową, jest istnienie bazy ortogonalnej $\{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_\nu\}$, gdzie każda macierz \mathbf{V}_i jest idempotentna, tzn. $\mathbf{V}_i^2 = \mathbf{V}_i$, oraz $\mathbf{V}_i \mathbf{V}_j = \mathbf{0}$ dla każdego $i \neq j$, $i, j = 1 \dots, \nu$. Z uwagi na to, że wszystkie wspomniane struktury kowariancyjne rozważane w pracy należą do komutatywnej podprzestrzeni kwadratowej, mogą być zapisane jako kombinacje liniowe macierzy bazowych, tzn.

$$\Sigma = \sum_{i=1}^{\nu} \lambda_i \mathbf{V}_i. \quad (1.7)$$

Jeden z możliwych sposobów konstrukcji macierzy \mathbf{V}_i , $i = 1, \dots, \nu$, bazuje na rozkładzie spektralnym. Jeżeli przez \mathbf{G} oznaczymy macierz, której kolumny stanowią ortonormalne wektory własne (ℓ_k , $k = 1, \dots, q$) rozważanej struktury kowariancyjnej oraz przez \mathbf{G}_i , $i = 1, \dots, \nu$, $\nu \leq q$, oznaczymy podmacierze macierzy \mathbf{G} składające się z wektorów własnych odpowiadających tej samej wartości własnej, to $\mathbf{V}_i = \mathbf{G}_i \mathbf{G}'_i$. Dla wymienionych struktur rozkłady spektralne prezentują się następująco:

- Σ_I : $\mathbf{G} = \mathbf{I}_q$, $\nu = 1$, a stąd $\mathbf{V}_1 = \mathbf{I}_q$;
- Σ_D : $\mathbf{G} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_q) = \mathbf{I}_q$, $\nu = q$, a stąd $\mathbf{V}_i = \mathbf{e}_i \mathbf{e}_i'$, $i = 1, \dots, \nu$, gdzie \mathbf{e}_i jest wektorem o i -tej składowej równej 1 i pozostałych składowych równych 0;
- Σ_{CS} : $\mathbf{G} = (\frac{1}{\sqrt{q}}\mathbf{1}_q : \mathbf{G}_{q-1})$, $\nu = 2$, a stąd $\mathbf{V}_1 = \frac{1}{q}\mathbf{1}_q\mathbf{1}_q'$, $\mathbf{V}_2 = \mathbf{G}_{q-1}\mathbf{G}_{q-1}'$, gdzie \mathbf{G}_{q-1} składa się z $q - 1$ wektorów własnych, ortogonalnych do $\mathbf{1}_q$; por. Olkin, Press (1969). Warto zauważyć, że $\mathbf{U}_1 = \mathbf{P}_q$ oraz $\mathbf{U}_2 = \mathbf{Q}_q$;
- Σ_{CT} : $\mathbf{G} = (\frac{1}{\sqrt{q}}\mathbf{1}_q, \ell_2, \dots, \ell_q)$, $\nu = \lfloor \frac{q}{2} + 1 \rfloor$, a stąd otrzymujemy $\mathbf{V}_1 = \frac{1}{q}\mathbf{1}_q\mathbf{1}_q'$,
 $\mathbf{V}_i = (\ell_i : \ell_{q-i+2})(\ell_i : \ell_{q-i+2})'$, $i = 2, \dots, \nu - 1$,
 $\mathbf{V}_\nu = \begin{cases} (\ell_\nu : \ell_{\nu+1})(\ell_\nu : \ell_{\nu+1})' & \text{dla } q \text{ nieparzystego} \\ \ell_\nu \ell_\nu' & \text{dla } q \text{ parzystego,} \end{cases}$
gdzie j -ta składowa wektora ℓ_i jest postaci

$$\ell_i^j = \frac{1}{\sqrt{q}} \left\{ \cos \left[\frac{2\pi}{q}(i-1)(j-1) \right] + \sin \left[\frac{2\pi}{q}(i-1)(j-1) \right] \right\}; \quad (1.8)$$

por. Basilevsky (1983), Olkin, Press (1969).

Warto nadmienić, że w przypadku macierzy $\mathbf{\Omega}$ nieposiadającej struktury, zazwyczaj jako bazę ortonormalną przyjmuje się zbiór macierzy generujących macierze symetryczne. Wówczas macierze $\mathbf{A}_{ii} = \mathbf{E}_{ii}$ dla $i = 1, \dots, q$ oraz $\mathbf{A}_{ij} = \frac{1}{\sqrt{2}}(\mathbf{E}_{ij} + \mathbf{E}_{ij}')$ dla $i < j$ oraz $i, j = 1, \dots, q$, gdzie \mathbf{E}_{ij} jest macierzą zerową z wyjątkiem jednego elementu, równego 1, znajdującego się w i -tym wierszu i j -tej kolumnie, stanowią bazę ortonormalną macierzy symetrycznej. Ponieważ przestrzeń macierzy symetrycznych nie jest komutatywna, macierze \mathbf{A}_{ij} , $i, j = 1, \dots, q$, nie są ortogonalne w ścisłym sensie.

Naturalnymi uogólnieniami powyższych struktur kowariancyjnych na modele podwójnie wielowymiarowe, gdzie $m > 1$, są ich blokowe odpowiedniki, to znaczy struktury, w których elementy skalarne poszczególnych macierzy Σ zastąpione zostały przez macierze stopnia m . Otrzymujemy wówczas odpowiednio:

- macierz blokowo diagonalną z takimi samymi blokami, $\Sigma_{BI} = \mathbf{I}_q \otimes \Delta$, gdzie Δ jest macierzą stopnia m określoną dodatnio;
- macierz o blokowej strukturze diagonalnej (ang. *block diagonal* - BD):

$$\Sigma_{BD} = \text{BDiag}(\Delta_1, \Delta_2, \dots, \Delta_q),$$

gdzie Δ_i , $i = 1, \dots, q$, są macierzami stopnia m określonymi dodatnio;

- macierz o blokowej strukturze kompletnej symetrii (ang. *block compound symmetry* - BCS):

$$\Sigma_{BCS} = (\mathbf{1}_q \mathbf{1}'_q - \mathbf{I}_q) \otimes \Gamma_1 + \mathbf{I}_q \otimes \Gamma_0 \quad (1.9a)$$

$$= \mathbf{Q}_q \otimes \Delta_1 + \mathbf{P}_q \otimes \Delta_2, \quad (1.9b)$$

gdzie macierz Γ_0 , stopnia m , jest symetryczna i określona dodatnio, Γ_1 , również stopnia m , jest symetryczna i dobrana tak, aby macierz Σ_{BCS} była określona dodatnio, natomiast w zapisie wykorzystującym rozkład spektralny macierze

$$\Delta_1 = \Gamma_0 - \Gamma_1 \quad \text{oraz} \quad \Delta_2 = \Gamma_0 + (q - 1)\Gamma_1 \quad (1.10)$$

są symetrycznymi macierzami określonymi dodatnio;

- macierz o blokowej strukturze kołowej macierzy Toeplitza (ang. *block circular Toeplitz* - BCT):

$$\Sigma_{BCT} = \begin{cases} \mathbf{I}_q \otimes \Gamma_0 + \sum_{i=1}^{\lfloor q/2 \rfloor} (\mathbf{C}^i + \mathbf{C}'^i) \otimes \Gamma_i & \text{dla } q \text{ nieparzystego} \\ \mathbf{I}_q \otimes \Gamma_0 + \sum_{i=1}^{q/2-1} (\mathbf{C}^i + \mathbf{C}'^i) \otimes \Gamma_i + \mathbf{C}^{q/2} \otimes \Gamma_{q/2} & \text{dla } q \text{ parzystego,} \end{cases}$$

gdzie symetryczna, określona dodatnio macierz Γ_0 stopnia m oraz symetryczne macierze Γ_i , $i = 1, 2, \dots, \lfloor q/2 \rfloor$ lub $(q/2 - 1)$ również stopnia m są dobrane tak, aby Σ_{BCT} była określona dodatnio.

W ogólności rozważane struktury blokowe można przedstawić jako

$$\Sigma = \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \Delta_i, \quad (1.11)$$

gdzie $\mathbf{V}_i, i = 1, \dots, \nu$, są macierzami bazowymi stopnia q komutatywnej podprzestrzeni kwadratowej \mathcal{V} , oraz $\Delta_i, i = 1, \dots, \nu$, są symetrycznymi określonymi dodatnio macierzami stopnia m . Zauważmy, że podprzestrzenie macierzy o strukturach rozważanych w niniejszej rozprawie w przypadku modeli podwójnie wielowymiarowych należą do podprzestrzeni kwadratowych, lecz nie komutatywnych.

Struktury reprezentowane przez macierze bazowe $\mathbf{V}_i, i = 1, \dots, \nu$, przedstawiają zależności między wektorami obserwacji związanymi z powtórzeniami. W niniejszej rozprawie będziemy je nazywać zależnościami międzyklasowymi. W przypadku, gdy na macierze reprezentowane przez Δ_i nałożymy jedną z rozważanych struktur, będziemy mówić o zależnościami wewnątrzklasowych. W celu sprecyzowania rozważanej struktury używać będziemy notacji z podwójnym indeksem dolnym, na przykład, macierze o strukturze BCS, dla których podbloki są macierzami stopnia m o strukturze CS oznaczane będą przez $\Sigma_{\text{BCS_CS}}$, natomiast macierze o strukturze BCT, dla których podbloki są macierzami stopnia m o strukturze CS oznaczane będą przez $\Sigma_{\text{BCT_CS}}$.

1.3 Testowanie hipotez

Testowaniem hipotez związanych ze strukturą macierzy kowariancji w modelach wielowymiarowych zajmowano się już w pierwszej połowie ubiegłego wieku. Testy statystyczne dotyczące macierzy kowariancji o strukturze proporcjonalnej do macierzy jednostkowej zaproponowane zostały w pracy Mauchly (1940), Girshick (1941) oraz Gleser (1966). W literaturze zajmowano się również pozostałymi strukturami kowariancyjnymi rozważanymi i przedstawionymi w niniejszej pracy. Macierz kowariancji o strukturze CS była testowana w pracy Wilks (1946). Podobne zagadnienie dotyczące średniej dla modelu wielowymiarowego i struktury CS rozważał Geisser (1963), który

zapropował statystykę testową składającą się z liniowej kombinacji dwóch niezależnych rozkładów F. Zapropowanie statystyk testowych dla hipotez badających, czy macierz kowariancji posiada konkretną strukturę zostało również przedstawione w pracy Olkin, Press (1969). Testowali oni, czy macierz kowariancji ma strukturę D lub CS przeciwko strukturze CT, a także czy posiada strukturę CT przeciwko dowolnej macierzy bez założonej struktury. Natomiast w pracy Krishnaiah, Lee (1976) rozważono zagadnienie testowania hipotezy dla macierzy kowariancji posiadającej strukturę liniową. Co więcej, przy założeniu, że macierz kowariancji posiada strukturę CT, autorzy sformułowali statystyki testowe dla testów wartości średniej sprawdzające, czy średnie wynoszą zero, czy też są równe sobie wobec hipotezy alternatywnej, że średnie są dowolne. Zależność kołowa została również rozważona w pracach Eaton (1983) oraz Nahtman, von Rosen (2008).

W literaturze rozważano również testowanie struktur kowariancyjnych dla modeli podwójnie wielowymiarowych. Bardzo często interesujące zmienne można w naturalny sposób podzielić na pewne grupy, z pewnymi wzorcami zależności dla zmiennych z różnych grup (zależność międzyklasowa) i/lub pewnymi wzorcami zależności dla zmiennych w obrębie każdej grupy (zależność wewnątrzklasowa). Przykładowo, kołową zależność w blokach macierzy, która prowadzi do struktury BCT, rozważano w pracy Olkin (1973). Co więcej, założenie równej zależności między dowolną parą zmiennych z różnych grup skutkuje strukturą BCS, wprowadzoną przez Rao (1945, 1953), gdzie badany był problem dyskryminacji grup odmiennych genetycznie. Badania struktury kowariancji BCS rozważane były również w pracach Arnold (1973, 1979). Zastosowanie struktury BCS w wielowymiarowych zagadnieniach danych przedziałowych pokazano w Hao i in. (2015). W pracy Roy i in. (2015) podano procedurę testową dla równości wektorów średnich dla danych podwójnie wielowymiarowych mających strukturę BCS.

W literaturze możemy znaleźć wiele przykładów testowania struktur kowariancyjnych statystykami uwzględnionymi również w niniejszej pracy. Test ilorazu wiarygodności (LRT) i test wynikowy Rao (RST) do testowania struktury kowariancji

BCS przedstawiono w pracach Roy, Leiva (2011), Roy i in. (2018) oraz Filipiak, Klein (2021). Statystyki testowe dla testów sprawdzających równość wektorów średnich, przy założeniu modelu podwójnie wielowymiarowego ze strukturą kowariancyjną BCS, zaproponowano w pracach Žežula i in. (2018) oraz Zmysłony i in. (2018). Asymptotyczny rozkład normalny LRT przy założeniu wymiaru każdego z bloków i wielkości próby dążącej do nieskończoności został zaprezentowany w Sun, Xie (2020). Niedawno Liang i in. (2021) wyprowadzili test LRT do jednoczesnego testowania równości wektorów średnich i równości podbloków macierzy kowariancji o strukturze BCS, posiadających strukturę CS lub CT, w przeciwieństwie do podbloków bez założonej struktury.

W niniejszej rozprawie rozważać będziemy ogólną hipotezę dotyczącą struktury macierzy kowariancji postaci

$$H_0 : \Sigma \in \mathcal{W}_0 \quad \text{vs} \quad H_1 : \Sigma \in \mathcal{W}_1 \setminus \mathcal{W}_0, \quad (1.12)$$

gdzie \mathcal{W}_0 jest pewną komutatywną podprzestrzenią kwadratową, \mathcal{W}_1 jest podprzestrzenią kwadratową (niekoniecznie komutatywną), natomiast symbol \setminus w hipotezie alternatywnej oznacza przestrzeń \mathcal{W}_1 z wykluczeniem przestrzeni \mathcal{W}_0 i służy do zapewnienia rozłączności hipotez. Dla uproszczenia zapisu, struktury rozważane w hipotezach zerowej i alternatywnej będziemy oznaczać odpowiednio przez Σ_0 i Σ_1 .

W niniejszej rozprawie rozważymy m.in. test LRT, test RST oraz test Walda (WT). Przy założeniu modelu (1.1) w każdym przypadku istotną rolę odgrywa funkcja wiarygodności $L(\boldsymbol{\mu}, \boldsymbol{\Omega}; \mathbf{x})$, której logarytm naturalny ma postać

$$\ln L(\boldsymbol{\mu}, \boldsymbol{\Omega}) = -\frac{nmq}{2} \ln(2\pi) - \frac{n}{2} \ln |\boldsymbol{\Omega}| - \frac{1}{2} \text{Tr} [(\mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu}') \boldsymbol{\Omega}^{-1} (\mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu}')'] .$$

Iloraz wiarygodności, w przypadku, gdy testujemy strukturę kowariancyjną, definiuje się jako

$$\Lambda = \frac{|\widehat{\Sigma}_0|^{n/2}}{|\widehat{\Sigma}_1|^{n/2}}, \quad (1.13)$$

gdzie $\widehat{\Sigma}_0$ oraz $\widehat{\Sigma}_1$ są estymatorami największej wiarygodności macierzy kowariancji odpowiednio przy prawdziwości hipotezy zerowej i alternatywnej; por. Anderson (2003).

Na podstawie (1.13) zauważmy, że statystyka testowa LRT, która jest zależna od estymatorów największej wiarygodności zarówno przy prawdziwości hipotezy zerowej jak i alternatywnej, przyjmuje postać

$$\text{LRT} = -2 \ln \Lambda = -2(\ln |\hat{\Sigma}_0| - \ln |\hat{\Sigma}_1|).$$

Statystyka testu RST, jest postaci

$$\text{RST} = \mathbf{s}'(\hat{\boldsymbol{\theta}}_0) \mathbf{F}^{-1}(\hat{\boldsymbol{\theta}}_0) \mathbf{s}(\hat{\boldsymbol{\theta}}_0), \quad (1.14)$$

gdzie $\mathbf{s}(\boldsymbol{\theta})$ jest wektorem wynikowym, którego elementami są pochodne cząstkowe pierwszego rzędu zlogarytmowanej funkcji wiarygodności ze względu na wektor parametrów $\boldsymbol{\theta}$ przy założeniu prawdziwości hipotezy alternatywnej, $\mathbf{F}(\boldsymbol{\theta})$ jest macierzą informacji Fishera o wektorze parametrów $\boldsymbol{\theta}$ zawartej w \mathbf{Y} , natomiast $\hat{\boldsymbol{\theta}}_0$ jest estymatorem największej wiarygodności wektora $\boldsymbol{\theta}$ przy prawdziwości hipotezy zerowej; por. Rao (2005). W odróżnieniu od statystyki LRT, statystyka RST jest funkcją estymatora największej wiarygodności wektora parametrów $\boldsymbol{\theta}$ tylko przy prawdziwości hipotezy zerowej.

Zauważmy, że hipotezy dotyczące struktur kowariancyjnych są hipotezami złożonymi, które w ogólności wyrazić można jako $\mathbf{w}(\boldsymbol{\theta}) = \mathbf{c}$, gdzie \mathbf{w} jest r -wymiarową funkcją wektorową wektora parametrów $\boldsymbol{\theta}$, przy czym r jest nie większe niż wymiar $\boldsymbol{\theta}$, natomiast \mathbf{c} jest r -wymiarowym wektorem stałych. Wówczas statystyka testowa Walda dla hipotezy $H_0 : \mathbf{w}(\boldsymbol{\theta}) = \mathbf{c}$, $H_1 : \mathbf{w}(\boldsymbol{\theta}) \neq \mathbf{c}$, na podstawie Rao (2005), jest postaci

$$\text{WT} = (\mathbf{w}(\hat{\boldsymbol{\theta}}) - \mathbf{c})' \mathbf{A}^{-1}(\hat{\boldsymbol{\theta}}) (\mathbf{w}(\hat{\boldsymbol{\theta}}) - \mathbf{c}), \quad (1.15)$$

gdzie $\mathbf{A}(\boldsymbol{\theta}) = \mathbf{W}(\boldsymbol{\theta}) \mathbf{F}^{-1}(\boldsymbol{\theta}) \mathbf{W}'(\boldsymbol{\theta})$ wraz z macierzą $\mathbf{W}(\boldsymbol{\theta})$ będącą macierzą pochodnych funkcji $\mathbf{w}(\boldsymbol{\theta})$ względem składowych $\boldsymbol{\theta}$, $\mathbf{F}(\boldsymbol{\theta})$ jest macierzą informacji Fishera o wektorze parametrów $\boldsymbol{\theta}$ oraz $\hat{\boldsymbol{\theta}}$ jest estymatorem największej wiarygodności wektora $\boldsymbol{\theta}$ przy prawdziwości hipotezy alternatywnej.

Przedstawione powyżej testy, określane mianem "świętej trójcy", wraz ze wzrostem wielkości próby dążą do rozkładu chi-kwadrat z liczbą stopni swobody równą

różnicy liczby nieznanych parametrów przy założeniu hipotezy alternatywnej i liczby nieznanych parametrów przy założeniu hipotezy zerowej; Rao (2005).

Omówione w pracy testy zostaną porównane m.in. pod kątem ich mocy, czyli prawdopodobieństwa uniknięcia błędu drugiego rodzaju. Empiryczna moc testu statystycznego obliczana jest w niniejszej pracy z wykorzystaniem badań symulacyjnych jako stosunek liczby odrzuceń hipotezy zerowej przy prawdziwości hipotezy alternatywnej do liczby wszystkich powtórzeń.

W analizie mocy testu istotną rolę odgrywa rozbieżność między hipotezami. Ponieważ w pracy zakładamy normalność modelu, jako miara rozbieżności wykorzystana zostanie entropijna funkcja straty (rozbieżność Kullbacka-Leiblera, por. Stein, 1956; Lin i in., 2014), reprezentująca rozbieżność między dwoma rozkładami normalnymi różniącymi się macierzą kowariancji, postaci

$$f_E(\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1) = \text{Tr}(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - \ln |\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0| - mq \quad (1.16)$$

Przy zadanej macierzy $\boldsymbol{\Sigma}_1$, wielkość rozbieżności między hipotezami wyznacza macierz $\boldsymbol{\Sigma}_0$, dla której

$$\zeta(\tilde{\boldsymbol{\Sigma}}_0) = \min_{\boldsymbol{\Sigma}_0 \in \mathcal{W}_0} f_E(\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1).$$

Zauważmy, że entropijna funkcja straty nie jest ograniczona z góry, dlatego do porównania mocy testów wykorzystamy skorygowaną rozbieżność postaci

$$\eta(\tilde{\boldsymbol{\Sigma}}_0) = 1 - \frac{1}{1 + \zeta(\tilde{\boldsymbol{\Sigma}}_0)},$$

sprowadzającą wartości rozbieżności do przedziału $[0, 1)$.

1.4 Wybrane przekształcenia macierzy i ich własności

W niniejszej pracy wykorzystywane są liczne własności macierzy i wektorów, dlatego podamy teraz niezbędne definicje i własności operatorów macierzowych, bazujące głównie na pracy Magnus, Neudecker (1986).

Należy przypomnieć, że operator „vec” służy do wektoryzowania macierzy poprzez ustawienie jej kolumn jedna pod drugą. Poniżej zostały przedstawione jego własności.

Własności 1.1. (*operatora vec i iloczynu Kroneckera*)

Niech \mathbf{a} , \mathbf{b} będą dwoma wektorami oraz niech macierze \mathbf{A}, \mathbf{B} , \mathbf{C} będą wymiarów, pozwalających na wykonanie poniższych działań. Wówczas

- (i) $\text{vec}(\mathbf{ab}') = \mathbf{b} \otimes \mathbf{a}$;
- (ii) $\text{vec}' \mathbf{A} \cdot \text{vec} \mathbf{B} = \text{Tr}(\mathbf{A}'\mathbf{B})$;
- (iii) $\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) \text{vec} \mathbf{B}$;
- (iv) $\text{Tr}(\mathbf{A} \otimes \mathbf{B}) = \text{Tr} \mathbf{A} \cdot \text{Tr} \mathbf{B}$;
- (v) dla kwadratowych macierzy \mathbf{A}, \mathbf{B} stopnia odpowiednio m oraz q
 $|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^q |\mathbf{B}|^m$.

Definicja 1.1. Niech \mathbf{A} będzie dowolną macierzą wymiaru $m \times q$. Macierz permutacyjną $\mathbf{K}_{m,q}$ wymiaru $mq \times mq$ taką, że

$$\mathbf{K}_{m,q} \text{vec} \mathbf{A} = \text{vec} \mathbf{A}' ,$$

nazywamy macierzą komutacyjną.

Z powyższej definicji wynika bezpośrednio, że jeśli \mathbf{A} jest macierzą symetryczną stopnia m , to $\mathbf{K}_{m,m} = \mathbf{I}_{m^2}$. Ponadto, dla macierzy \mathbf{A} i \mathbf{B} wymiaru odpowiednio $m \times q$ oraz $n \times p$, prawdziwa jest zależność

$$\mathbf{K}_{n,m}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{p,q}, \quad (1.17)$$

a w konsekwencji, ponieważ $\mathbf{K}'_{p,q} = \mathbf{K}_{q,p}$,

$$\mathbf{K}_{n,m}(\mathbf{A} \otimes \mathbf{B})\mathbf{K}_{q,p} = \mathbf{B} \otimes \mathbf{A}. \quad (1.18)$$

Definicja 1.2. Niech \mathbf{A} będzie kwadratową, symetryczną macierzą stopnia m . Wówczas $\text{vech } \mathbf{A}$ jest operatorem wektoryzowania dolnego trójkąta macierzy \mathbf{A} , a zatem wektorem $m(m+1)/2$ -wymiarowym, powstającym przez zapisanie kolumn dolnego trójkąta macierzy \mathbf{A} jedna pod drugą.

Definicja 1.3. Niech \mathbf{A} będzie dowolną kwadratową i symetryczną macierzą stopnia m . Macierz \mathbf{D}_m wymiaru $m^2 \times \frac{1}{2}m(m+1)$ taką, że

$$\mathbf{D}_m \text{vech } \mathbf{A} = \text{vec } \mathbf{A},$$

nazywamy macierzą duplikacji.

W pracy będziemy wykorzystywać również następującą własność zaczerpniętą z pracy Filipiak i in. (2016, Lemat 1).

Lemat 1.1. Dla dowolnej symetrycznej macierzy \mathbf{A} stopnia m prawdziwe jest

$$\begin{aligned} \mathbf{D}'_m &= (\text{vec}' \mathbf{A} \otimes \mathbf{D}'_m)(\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m)(\text{vec } \mathbf{A}^{-1} \otimes \mathbf{I}_{m^2}) \\ &= (\text{vec}' \mathbf{A} \otimes \mathbf{D}'_m)(\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m)(\mathbf{I}_{m^2} \otimes \text{vec } \mathbf{A}^{-1}). \end{aligned}$$

Niech \mathbf{D}_m^+ będzie odwrotnością Moore'a - Penrose'a macierzy duplikacji \mathbf{D}_m . Wówczas $\mathbf{D}_m \mathbf{D}_m^+ = \mathbf{N}_m$, gdzie $\mathbf{N}_m = \frac{1}{2}(\mathbf{I}_{m^2} + \mathbf{K}_{m,m})$. Ponadto, dla dowolnej symetrycznej macierzy \mathbf{A} stopnia m ,

$$\mathbf{N}_m \text{vec } \mathbf{A} = \text{vec } \mathbf{A}.$$

Co więcej, z pracy Filipiak i in. (2016) wiadomo, że dla dowolnej, nieosobliwej macierzy \mathbf{A} stopnia m prawdziwa jest równość

$$\left[\mathbf{D}'_m (\mathbf{A}^{-1} \otimes \mathbf{A}^{-1}) \mathbf{D}_m \right]^{-1} = \mathbf{D}_m^+ (\mathbf{A} \otimes \mathbf{A}) \mathbf{D}_m'^+. \quad (1.19)$$

W niniejszej pracy wykorzystano również operatory śladu blokowego. Poniższa definicja oraz własności śladu blokowego zostały przedstawione za pracą Filipiak i in. (2018a), gdzie wspomniany operator jest zdefiniowany dla macierzy prostokątnych.

Definicja 1.4. Dla dowolnej macierzy blokowej \mathbf{A} wymiaru $m q \times n q$ operatorem śladu blokowego, oznaczonym przez $\text{BTr}_{m,n} \mathbf{A}$, nazywamy macierz stopnia q powstałą poprzez zsumowanie diagonalnych bloków wymiaru $m \times n$ macierzy \mathbf{A} , tj.

$$\text{BTr}_{m,n} \mathbf{A} = \sum_{i=1}^m \mathbf{A}_{ii}.$$

Jeśli $m = n$, wówczas dla uproszczenia zapisu zamiast $\text{BTr}_{m,n} \mathbf{A}$ będziemy stosować notację $\text{BTr}_m \mathbf{A}$.

Własności 1.2. (operatora śladu blokowego)

Niech \mathbf{A} będzie macierzą stopnia $m q$ oraz niech macierze \mathbf{B}, \mathbf{C} będą takiego stopnia, aby poniższe działania były zdefiniowane. Wówczas

- (i) $\text{vec}\{\text{BTr}_m[(\mathbf{B} \otimes \mathbf{I}_m) \mathbf{A} (\mathbf{C}' \otimes \mathbf{I}_m)]\} = [\text{vec}'(\mathbf{B}'\mathbf{C}) \otimes \mathbf{I}_{m^2}](\mathbf{I}_q \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_m) \text{vec} \mathbf{A}$;
- (ii) $\text{BTr}_{mq}(\mathbf{B} \otimes \mathbf{A}) = \text{Tr} \mathbf{B} \cdot \mathbf{A}$;
- (iii) $\text{Tr}(\text{BTr}_m \mathbf{A}) = \text{Tr}(\text{BTr}_q \mathbf{A}) = \text{Tr} \mathbf{A}$;
- (iv) $\text{BTr}_n[(\mathbf{I}_q \otimes \mathbf{B}) \mathbf{A} (\mathbf{I}_q \otimes \mathbf{C})] = \mathbf{B} \cdot \text{BTr}_m(\mathbf{A}) \cdot \mathbf{C}$;
- (v) dla dowolnych macierzy \mathbf{A}, \mathbf{B} , wymiaru odpowiednio $m q \times n q$, $n \times m$ prawdziwa jest zależność $\text{BTr}_q[\mathbf{A}(\mathbf{B} \otimes \mathbf{I}_q)] = \text{BTr}_q[(\mathbf{B} \otimes \mathbf{I}_q)\mathbf{A}]$.

W pracy będzie wykorzystywany również operator sumy blokowej, przedstawiony w pracy Roy i in. (2015) i zdefiniowany następująco.

Definicja 1.5. Dla dowolnej macierzy blokowej \mathbf{A} wymiaru $m q \times n q$ operatorem sumy blokowej, oznaczonym przez $\text{BSum}_{m,n} \mathbf{A}$, nazywamy macierz stopnia q powstałą poprzez zsumowanie wszystkich podbloków wymiaru $m \times n$ macierzy \mathbf{A} , tj.

$$\text{BSum}_{m,n} \mathbf{A} = \sum_{j=1}^n \sum_{i=1}^m \mathbf{A}_{ij}.$$

Jeśli $m = n$, wówczas dla uproszczenia zapisu zamiast $\text{BSum}_{m,n} \mathbf{A}$ będziemy stosować notację $\text{BSum}_m \mathbf{A}$.

W niniejszej rozprawie wykorzystamy także operator śladu częściowego, przedstawiony w pracy Filipiak i in. (2018a), którego definicja jest następująca.

Definicja 1.6. Dla dowolnej macierzy blokowej \mathbf{A} wymiaru $m q \times n q$ operatorem śladu częściowego, oznaczonym przez $\text{PTr}_q \mathbf{A}$, nazywamy macierz wymiaru $m \times n$ powstałą poprzez zastąpienie poszczególnych podbloków stopnia q macierzy \mathbf{A} ich śladami, tj.

$$\text{PTr}_q \mathbf{A} = (\text{Tr } \mathbf{A}_{ij})_{i,j=1,\dots,m}.$$

Własności 1.3. (operatora śladu częściowego)

Niech $\mathbf{A}, \mathbf{B}, \mathbf{C}$ będą dowolnymi macierzami wymiaru odpowiednio $m q \times n q$, $q \times m$ oraz $q \times n$. Wówczas,

- (i) $\text{vec}(\text{PTr}_q \mathbf{A}) = (\mathbf{I}_{mn} \otimes \text{vec}' \mathbf{I}_q)(\mathbf{I}_n \mathbf{K}_{m,q} \otimes \mathbf{I}_q) \text{vec } \mathbf{A}$;
- (ii) $\text{PTr}_q(\text{vec } \mathbf{B} \text{vec}' \mathbf{C}) = \mathbf{B}' \mathbf{C}$.

Operatorem odgrywającym istotną rolę w niniejszej pracy jest również operator rzutowania na przestrzeń macierzy o zadanej strukturze. W tym celu potrzebny będzie nam iloczyn skalarny macierzy.

Definicja 1.7. Iloczynem skalarnym dla macierzy \mathbf{A} i \mathbf{B} wymiaru $m \times n$, nazywamy wyrażenie

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{A} \mathbf{B}').$$

Oznaczmy przez \mathfrak{B} bazę ortonormalną macierzy ze zbioru struktur \mathcal{S} . Wówczas rzutem macierzy \mathbf{A} na r -wymiarową przestrzeń generowaną przez bazę \mathfrak{B} , nazywamy macierz postaci

$$\sum_{i=1}^r \langle \mathbf{A}, \mathbf{B}_i \rangle \mathbf{B}_i,$$

gdzie $\mathbf{B}_i \in \mathfrak{B}$ oraz $\langle \bullet, \bullet \rangle$ jest iloczynem skalarnym macierzy. Rzut macierzy \mathbf{A} na przestrzeń struktur kowariancyjnych można zdefiniować równoważnie jako

$$\text{argmin}_{\Sigma \in \mathcal{S}} \|\mathbf{A} - \Sigma\|_F,$$

gdzie $\|\mathbf{A}\|_F = \sqrt{\text{Tr}(\mathbf{A} \mathbf{A}')}$ oznacza normę Frobeniusa macierzy \mathbf{A} .

Istotnym elementem niniejszej pracy jest również różniczkowanie macierzy. Poniższe wzory przedstawione są za pracą Fackler (2005).

Lemat 1.2. Dla dowolnej macierzy \mathbf{X} oraz $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$ pochodną \mathbf{Y} względem \mathbf{X} definiujemy jako

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d \operatorname{vec} \mathbf{Y}}{d \operatorname{vec}' \mathbf{X}}.$$

Lemat 1.3. Niech macierze \mathbf{A}, \mathbf{B} będą odpowiednio wymiaru $m \times n$ oraz $p \times q$, a także niech macierz \mathbf{X} i wektor \mathbf{x} będą takiego wymiaru, aby przedstawione poniżej iloczyny były zdefiniowane. Dodatkowo niech a będzie stałą. Wówczas,

$$\begin{aligned} \frac{da\mathbf{A}}{da} &= \operatorname{vec} \mathbf{A}; \\ \frac{d\mathbf{x}'\mathbf{A}\mathbf{x}}{d\mathbf{A}} &= \mathbf{x}\mathbf{x}'; \\ \frac{d\mathbf{A}'\mathbf{A}}{d\mathbf{A}} &= (\mathbf{I}_{n^2} + \mathbf{K}_{n,n}) (\mathbf{I}_n \otimes \mathbf{A}'); \\ \frac{d\mathbf{A}\mathbf{A}'}{d\mathbf{A}} &= (\mathbf{I}_{m^2} + \mathbf{K}_{m,m}) (\mathbf{A} \otimes \mathbf{I}_m); \\ \frac{d\mathbf{A}\mathbf{X}\mathbf{B}}{d\mathbf{X}} &= \mathbf{B}' \otimes \mathbf{A}; \\ \frac{d\mathbf{A}^{-1}}{d\mathbf{A}} &= - \left[(\mathbf{A}^{-1})' \otimes \mathbf{A}^{-1} \right]; \\ \frac{d \ln |\mathbf{A}|}{d\mathbf{A}} &= \operatorname{vec}' (\mathbf{A}^{-1})'; \\ \frac{d \operatorname{Tr} (\mathbf{A}\mathbf{X})}{d\mathbf{X}} &= \operatorname{vec}' \mathbf{A}'; \\ \frac{d\mathbf{A} \otimes \mathbf{B}}{d\mathbf{A}} &= (\mathbf{I}_n \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_p) (\mathbf{I}_{mn} \otimes \operatorname{vec} \mathbf{B}) = (\mathbf{I}_{nq} \otimes \mathbf{K}_{m,p}) (\mathbf{I}_n \otimes \operatorname{vec} \mathbf{B} \otimes \mathbf{I}_m); \\ \frac{d\mathbf{A} \otimes \mathbf{B}}{d\mathbf{B}} &= (\mathbf{I}_n \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_p) (\operatorname{vec} \mathbf{A} \otimes \mathbf{I}_{pq}) = (\mathbf{K}_{p,q} \otimes \mathbf{I}_{mn}) (\mathbf{I}_q \otimes \operatorname{vec} \mathbf{A} \otimes \mathbf{I}_p); \\ \frac{d\mathbf{A}^i}{d\mathbf{A}} &= \sum_{j=1}^i (\mathbf{A}')^{i-j} \otimes \mathbf{A}^{j-1}. \end{aligned}$$

Dodatkowo, korzystając z pracy Kollo, von Rosen (2005), poniżej w postaci lematu została przedstawiona następująca własność.

Lemat 1.4. *Dla dowolnych macierzy \mathbf{X} , \mathbf{L} oraz $\mathbf{Y}(\mathbf{X})$ i macierzy duplikacji \mathbf{D} wymiaru takiego, że działanie $\mathbf{D}'\mathbf{Y}\mathbf{L}$ jest zdefiniowane, prawdziwa jest własność*

$$\frac{d\mathbf{D}'\mathbf{Y}\mathbf{L}}{d\mathbf{X}} = (\mathbf{L}' \otimes \mathbf{D}') \frac{d\mathbf{Y}}{d\mathbf{X}}.$$

Na koniec zauważmy, że reguła łańcuchowa różniczkowania funkcji złożonych może być również zastosowana w przypadku różniczkowania złożonych funkcji macierzowych, z uwzględnieniem kolejności różniczkowania, to znaczy

$$\frac{d\mathbf{Z}(\mathbf{Y}(\mathbf{X}))}{d\mathbf{X}} = \frac{d\mathbf{Z}}{d\mathbf{Y}} \cdot \frac{d\mathbf{Y}}{d\mathbf{X}}. \quad (1.20)$$

Ponadto, regułę łańcuchową wykorzystać można do różniczkowania macierzy symetrycznych, przez różniczkowanie nie ze względu na całą macierz symetryczną tylko jej dolny trójkąt. Zatem w przypadku, gdy \mathbf{X} jest macierzą symetryczną stopnia m ,

$$\frac{d\mathbf{Y}(\mathbf{X})}{d\mathbf{X}} = \frac{d \operatorname{vec} \mathbf{Y}}{d \operatorname{vec}' \mathbf{X}} \cdot \frac{d \operatorname{vec} \mathbf{X}}{d \operatorname{vech}' \mathbf{X}} = \frac{d \operatorname{vec} \mathbf{Y}}{d \operatorname{vec}' \mathbf{X}} \cdot \mathbf{D}_m;$$

por. Magnus, Neudecker (1986).

1.5 Przykład rzeczywisty

Do zilustrowania procedur testowych przedstawionych w niniejszej pracy posługiwać się będziemy danymi rzeczywistymi, pierwotnie przedstawionymi w pracy Liang i in. (2015), w której zaprezentowano hierarchiczny model ze strukturą BCT (w szczególności strukturą BCS). Zebrane dane dotyczą długości płatków kwiatów żyworódek. Pomiarów dokonano na 11 losowo wybranych roślinach z tej samej szklarni. Z każdej rośliny wybrano losowo 3 z 4 kwiatostanów, a następnie zmierzono długości ich płatków (każdy kwiatostan składa się z 4 płatków). Zatem w tym eksperymencie dysponujemy $n = 11$ obiektami, dla których zmierzono długość $m = 4$ płatków na każdym z $q = 3$ kwiatostanów.

Zakładając, podobnie jak Liang i in. (2015), że kowariancja między każdymi dwoma kwiatostanami jest taka sama, rozsądne jest przyjęcie struktury BCS dla macierzy kowariancji. Ponadto, ponieważ w każdym kwiecie układ płatków jest kołowy, Liang i in. (2015) dodatkowo założyli strukturę kowariancji CT między płatkami w każdym kwiecie.

Naszym celem jest przetestowanie hipotez związanych z niektórymi z omawianych struktur macierzy kowariancji długości płatków. Należy jednak zauważyć, że ponieważ w niniejszej rozprawie nie uwzględniono żadnej specjalnej struktury wartości oczekiwanej, otrzymane wyniki różnią się od tych przedstawionych przez Liang i in. (2021), gdzie zakładano, że wartość oczekiwana posiada strukturę $(\mathbf{1}_3 \otimes \boldsymbol{\mu})\mathbf{1}'_{11}$ gdzie $\boldsymbol{\mu} \in \mathbb{R}^4$.

2 Estymacja macierzy kowariancji o zadanej strukturze, należącej do podprzestrzeni kwadratowej

Zdefiniujmy następujący model mieszany:

$$\mathbf{y} = \mathbf{Z}_0\boldsymbol{\beta}_0 + \sum_{i=1}^{\nu-1} \mathbf{Z}_i\boldsymbol{\beta}_i + \mathbf{e}, \quad (2.1)$$

gdzie \mathbf{y} jest wektorem obserwacji, $\boldsymbol{\beta}_0$ nieznanym wektorem efektów stałych, $\boldsymbol{\beta}_i$, $i = 1, \dots, \nu$, nieznanymi wektorami efektów losowych, \mathbf{e} jest nieznanym wektorem błędów losowych, natomiast $\mathbf{Z}_0, \mathbf{Z}_i$, $i = 1, \dots, \nu - 1$, są macierzami układu odpowiednio dla efektów stałych i losowych. Przyjmując, że wszystkie składniki losowe modelu mają zerowe wartości oczekiwane i $\text{cov}(\mathbf{e}) = \sigma_0^2\mathbf{L}_0$, $\text{cov}(\boldsymbol{\beta}_i) = \sigma_i^2\mathbf{L}_i$, $\text{cov}(\mathbf{e}, \boldsymbol{\beta}_i) = 0$ dla $i = 1, \dots, \nu - 1$, gdzie σ_i^2 są nieznanymi parametrami, natomiast \mathbf{L}_i są znanym symetrycznymi, określonymi dodatnio macierzami, macierz kowariancji modelu (2.1) można wyrazić jako $\boldsymbol{\Xi} = \boldsymbol{\Xi}(\boldsymbol{\sigma}) = \sum_{i=0}^{\nu-1} \sigma_i^2\mathbf{L}_i$.

Rozważmy teraz model (1.2), w którym $\boldsymbol{\Omega} = \boldsymbol{\Sigma}$ ma strukturę liniową zależną od wektora nieznanych współczynników $\boldsymbol{\sigma}$, to znaczy $\boldsymbol{\Sigma} = \sum_{i=1}^{mq} \sigma_i^2\boldsymbol{\Sigma}_i$. Łatwo wówczas zauważyć, że za pomocą wektoryzacji można przekształcić wspomniany model do modelu analogicznego do (2.1) postaci

$$\text{vec } \mathbf{Y} \sim N_{nmq}((\mathbf{I}_{mq} \otimes \mathbf{X}) \text{vec } \boldsymbol{\beta}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n).$$

Analogię można zauważyć przyjmując $\text{vec } \mathbf{Y} = \mathbf{y}$, $\text{vec } \boldsymbol{\beta} = \boldsymbol{\beta}_0$, $\mathbf{I}_{mq} \otimes \mathbf{X} = \mathbf{Z}_0$ oraz $\boldsymbol{\Sigma}_i \otimes \mathbf{I}_n = \mathbf{L}_i$. Taka obserwacja pozwala nam zaadaptować metody estymacji w modelu mieszanym na model wielowymiarowy z liniową strukturą macierzy kowariancji, co zostanie przedstawione w kolejnych rozdziałach.

Wszystkie wyniki zaprezentowane w tym rozdziale znaleźć można w artykule Filipiak, John, Markiewicz (2020).

2.1 Estymacja macierzy kowariancji z wykorzystaniem operatora rzutu

Szatrowski (1980) dostrzegając podobieństwo między modelem wielowymiarowym z liniową strukturą kowariancji, a jednowymiarowym modelem mieszanym pokazał, przy założeniu znanej wartości oczekiwanej, że estymator największej wiarygodności macierzy kowariancji ma jawną reprezentację wtedy i tylko wtedy, gdy przestrzeń macierzy kowariancji jest podprzestrzenią kwadratową. Pokażemy, że warunek ten można rozszerzyć na przypadek, gdy wartość oczekiwana nie jest znana.

Rozważmy model postaci (2.1). Wiadomo, że jeśli rzut na przestrzeń wartości oczekiwanej, \mathbf{P}_X , komutuje z macierzą kowariancji, $\mathbf{\Xi}$, to równanie wiarygodności dla wartości oczekiwanej nie zależy od macierzy kowariancji, a zatem jego rozwiązaniem jest estymator najmniejszych kwadratów (ang. *ordinary least squares estimator*, OLSE). Rzeczywiście, w modelu z macierzą kowariancji $\mathbf{\Xi}$ równanie wiarygodności dla wartości oczekiwanej ma postać

$$\mathbf{X}'\mathbf{\Xi}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{0}.$$

Mnożąc równanie z lewej strony przez $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$ dostajemy równoważnie

$$\mathbf{P}_X\mathbf{\Xi}^{-1}\mathbf{y} = \mathbf{P}_X\mathbf{\Xi}^{-1}\mathbf{X}\boldsymbol{\beta}. \quad (2.2)$$

Zakładając, że rzut \mathbf{P}_X komutuje z macierzą kowariancji, oczywistym jest także fakt, iż komutuje z jej odwrotnością, czyli $\mathbf{P}_X\mathbf{\Xi}^{-1} = \mathbf{\Xi}^{-1}\mathbf{P}_X$. Zatem równanie (2.2) może zostać zapisane w postaci

$$\mathbf{\Xi}^{-1}\mathbf{P}_X\mathbf{y} = \mathbf{\Xi}^{-1}\mathbf{P}_X\mathbf{X}\boldsymbol{\beta}.$$

Mnożąc powyższe równanie obustronnie z lewej strony przez $\mathbf{\Xi}$ i z uwagi, że $\mathbf{P}_X\mathbf{X} = \mathbf{X}$, otrzymujemy

$$\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{P}_X\mathbf{y}. \quad (2.3)$$

Podsumowując, Twierdzenie 4 z pracy Szatrowski (1980) może zostać uogólnione na przypadek z nieznaną wartością oczekiwaną w następujący sposób.

Twierdzenie 2.1. Niech \mathbf{y} będzie wektorem obserwacji o rozkładzie normalnym z wartością oczekiwaną $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ i macierzą kowariancji $\boldsymbol{\Xi} = \boldsymbol{\Xi}(\boldsymbol{\sigma})$, należącą do przestrzeni liniowej \mathcal{V} , o nieznanym wektorze współczynników kombinacji liniowej $\boldsymbol{\sigma}$. Jeśli $\mathbf{P}_X\boldsymbol{\Xi} = \boldsymbol{\Xi}\mathbf{P}_X$, to następujące warunki są równoważne:

1. istnieje estymator największej wiarygodności macierzy $\boldsymbol{\Xi}$ w postaci jawnej;
2. \mathcal{V} jest podprzestrzenią kwadratową.

Z powyższego twierdzenia wiadomo, że jeżeli tylko przestrzeń macierzy o zadanej strukturze jest podprzestrzenią kwadratową i komutuje z przestrzenią wartości oczekiwanej, to estymator największej wiarygodności macierzy kowariancji ma postać jawną, która jednak nie jest sprecyzowana. Biorąc pod uwagę, że estymatorem największej wiarygodności funkcji parametrycznej jest funkcja estymatora największej wiarygodności parametru (por. Zehna, 1966) należy przypuszczać, że interesujący estymator jest pewną funkcją estymatora największej wiarygodności macierzy kowariancji bez struktury. Jaką funkcję należy wybrać, przedstawia następujące twierdzenie.

Twierdzenie 2.2. Niech \mathbf{y} będzie wektorem obserwacji o rozkładzie normalnym z wartością oczekiwaną $\mathbb{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ oraz macierzą kowariancji $\boldsymbol{\Xi} = \boldsymbol{\Xi}(\boldsymbol{\sigma})$ należącą do przestrzeni liniowej \mathcal{V} z nieznanym wektorem współczynników kombinacji liniowej $\boldsymbol{\sigma}$. Załóżmy, że $\mathbf{P}_X\boldsymbol{\Xi} = \boldsymbol{\Xi}\mathbf{P}_X$ oraz \mathcal{V} jest podprzestrzenią kwadratową. Wówczas estymatorem największej wiarygodności macierzy kowariancji $\boldsymbol{\Xi}$ jest rzut macierzy \mathbf{S}^* na przestrzeń \mathcal{V} , gdzie \mathbf{S}^* jest estymatorem największej wiarygodności macierzy kowariancji bez struktury, to znaczy $\mathbf{S}^* = \mathbf{Q}_X\mathbf{y}\mathbf{y}'\mathbf{Q}_X$.

Dowód. Ze względu na komutatywność \mathbf{P}_X i $\boldsymbol{\Xi}$ oraz fakt, iż \mathcal{V} jest podprzestrzenią kwadratową, z Twierdzenia 2.1 wiemy, że estymator największej wiarygodności $\boldsymbol{\sigma}$ ma jawną postać. Z definicji podprzestrzeni kwadratowej, jeśli $\boldsymbol{\Xi} \in \mathcal{V}$, to $\boldsymbol{\Xi}^{-1} \in \mathcal{V}$ (por. Seely, 1971), więc $\boldsymbol{\Xi}^{-1}$ może być zapisana jako kombinacja liniowa wektorów bazowych przestrzeni \mathcal{V} , to znaczy $\boldsymbol{\Xi}^{-1}(\boldsymbol{\alpha}) = \sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i$, gdzie $\nu = \dim(\mathcal{V})$, $\alpha_i > 0$, oraz \mathbf{L}_i , $i = 0, \dots, \nu - 1$, są elementami bazy ortonormalnej przestrzeni \mathcal{V} .

Rozważmy logarytm funkcji wiarygodności

$$\ln L(\boldsymbol{\alpha}) = -\frac{n}{2} \ln(2\pi) + \frac{1}{2} \ln \left| \sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right| - \frac{1}{2} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right) (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}).$$

Różniczkując powyższe ze względu na składowe wektora $\boldsymbol{\alpha}$, przy pomocy wzorów na różniczkowanie macierzy (zob. Lemat 1.3), otrzymujemy ν równań wiarygodności postaci

$$\frac{\partial \ln L(\boldsymbol{\alpha})}{\partial \alpha_j} = \frac{1}{2} \text{vec}' \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \right] \text{vec } \mathbf{L}_j - \frac{1}{2} [(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \otimes (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'] \text{vec } \mathbf{L}_j,$$

$j = 0, \dots, \nu - 1$. Korzystając z Własności 1.1(i), (ii), ze wzoru (2.3), a także wykorzystując definicje operatorów \mathbf{P}_X oraz \mathbf{Q}_X , powyższe wyrażenie można przekształcić do postaci

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\alpha})}{\partial \alpha_j} &= \text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] - \text{vec}' [(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'] \text{vec } \mathbf{L}_j \\ &= \text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] - \text{vec}' [(\mathbf{y} - \mathbf{P}_X \mathbf{y})(\mathbf{y} - \mathbf{P}_X \mathbf{y})] \text{vec } \mathbf{L}_j \\ &= \text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] - \text{vec}' [(\mathbf{I} - \mathbf{P}_X) \mathbf{y} \mathbf{y}' (\mathbf{I} - \mathbf{P}_X)] \text{vec } \mathbf{L}_j \\ &= \text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] - \text{vec}' [\mathbf{Q}_X \mathbf{y} \mathbf{y}' \mathbf{Q}_X] \text{vec } \mathbf{L}_j \\ &= \text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] - \text{Tr} (\mathbf{Q}_X \mathbf{y} \mathbf{y}' \mathbf{Q}_X \mathbf{L}_j) \\ &= \text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] - \text{Tr} (\mathbf{S}^* \mathbf{L}_j). \end{aligned}$$

Ponieważ $\left(\sum_{i=1}^{\nu} \alpha_i \mathbf{L}_i \right)^{-1} = \sum_{i=1}^{\nu} \sigma_i \mathbf{L}_i$ oraz $\text{Tr}(\mathbf{L}_i \mathbf{L}_j) = 0$ dla $i \neq j$, a także $\text{Tr}(\mathbf{L}_i^2) = 1$, przyrównując powyższe do zera, otrzymujemy

$$\text{Tr} \left[\left(\sum_{i=0}^{\nu-1} \alpha_i \mathbf{L}_i \right)^{-1} \mathbf{L}_j \right] = \text{Tr} (\mathbf{S} \mathbf{L}_j),$$

a stąd

$$\hat{\sigma}_j = \text{Tr}(\mathbf{S}\mathbf{L}_j),$$

$j = 0, \dots, \nu - 1$. Z drugiej strony, rzutując macierz \mathbf{S}^* na przestrzeń \mathcal{V} mamy

$$\hat{\mathbf{E}} = \sum_{i=0}^{\nu-1} \langle \mathbf{S}^*, \mathbf{L}_i \rangle \mathbf{L}_i = \sum_{i=0}^{\nu-1} \text{Tr}(\mathbf{S}^* \mathbf{L}_i) \mathbf{L}_i = \sum_{i=0}^{\nu-1} \hat{\sigma}_i^2 \mathbf{L}_i,$$

gdzie $\hat{\sigma}_i = \text{Tr}(\mathbf{S}\mathbf{L}_i)$, $i = 0, \dots, \nu - 1$. □

W kolejnym rozdziale Twierdzenie 2.2 zostanie zastosowane do estymacji szczególnych struktur kowariancyjnych należących do podprzestrzeni kwadratowych.

2.2 Estymacja szczególnych struktur macierzy kowariancji

W niniejszym rozdziale rozważymy trzy modele różniące się strukturą wartości oczekiwanej z uwzględnieniem szczególnych struktur kowariancyjnych należących do podprzestrzeni kwadratowej. Przypomnijmy, że przed zastosowaniem tezy Twierdzenia 2.2 należy sprawdzić jego założenia, to znaczy komutatywność przestrzeni wartości oczekiwanej i przestrzeni macierzy kowariancji.

2.2.1 Standardowy model wielowymiarowy

Niech $m = 1$. Dokonując wektoryzacji modelu (1.1) otrzymujemy

$$\text{vec } \mathbf{Y} \sim N_{qn}(\text{vec}(\mathbf{1}_n \boldsymbol{\mu}'), \boldsymbol{\Omega} \otimes \mathbf{I}_n).$$

Ponieważ $\text{vec}(\mathbf{1}_n \boldsymbol{\mu}') = (\mathbf{I}_q \otimes \mathbf{1}_n) \boldsymbol{\mu}$ oraz estymatorem największej wiarygodności macierzy $\boldsymbol{\Omega}$ jest $\mathbf{S}^* = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y} = \mathbf{S}$, a także

$$\mathbf{P}_{I_q \otimes \mathbf{1}_n} = \frac{1}{n} (\mathbf{I}_q \otimes \mathbf{1}_n) (\mathbf{I}_q \otimes \mathbf{1}_n') = \mathbf{I}_q \otimes \mathbf{P}_n$$

mamy, że $\mathbf{P}_{I_q \otimes \mathbf{1}_n} (\boldsymbol{\Omega} \otimes \mathbf{I}_n) = \boldsymbol{\Omega} \otimes \mathbf{P}_n = (\boldsymbol{\Omega} \otimes \mathbf{I}_n) \mathbf{P}_{I_q \otimes \mathbf{1}_n}$, co dowodzi komutatywności przestrzeni wartości oczekiwanej i macierzy kowariancji. Zwróćmy uwagę, że

powyższa własność nie zależy od potencjalnej struktury macierzy kowariancji, a więc zachodzi również dla $\Sigma = \sum_{i=1}^{\nu} \lambda_i \mathbf{V}_i$; por. (1.7).

Niech $v_i = \text{Tr } \mathbf{V}_i$. Wówczas macierze $v_i^{-1/2} \mathbf{V}_i$ tworzą bazę ortonormalną podprzestrzeni kwadratowej \mathcal{V} i, korzystając z Twierdzenia 2.2, estymator największej wiarygodności macierzy Σ o zadanej strukturze jest rzutem macierzy \mathbf{S} na przestrzeń \mathcal{V} , to znaczy

$$\hat{\Sigma} = \sum_{i=1}^{\nu} \langle \mathbf{S}, \frac{1}{\sqrt{v_i}} \mathbf{V}_i \rangle \frac{1}{\sqrt{v_i}} \mathbf{V}_i = \sum_{i=1}^{\nu} \frac{1}{v_i} \text{Tr}(\mathbf{S} \mathbf{V}_i) \mathbf{V}_i. \quad (2.4)$$

Zastosujmy teraz (2.4) do szczególnych struktur kowariancyjnych należących do podprzestrzeni kwadratowych odpowiedniego wymiaru, omówionych w Rozdziale 1.2.

Niech $\Sigma = \Sigma_{\text{CS}}$. Przyjmując parametryzację (1.5) łatwo zauważyć, że przestrzeń struktury CS jest podprzestrzenią kwadratową (co wynika z istnienia bazy złożonej z macierzy idempotentnych i ortogonalnych w ścisłym sensie; por. Seely, 1971). Wówczas, przechodząc do bazy ortonormalnej składającej się z macierzy \mathbf{P}_q i $\frac{1}{\sqrt{q-1}} \mathbf{Q}_q$, bezpośrednio z (2.4) otrzymujemy

$$\hat{\Sigma}_{\text{CS}} = \text{Tr}(\mathbf{S} \mathbf{P}_q) \mathbf{P}_q + \frac{1}{q-1} \text{Tr}(\mathbf{S} \mathbf{Q}_q) \mathbf{Q}_q,$$

a stąd $\hat{\gamma} = \text{Tr}(\mathbf{S} \mathbf{P}_q)$ i $\hat{\delta} = \frac{1}{q-1} \text{Tr}(\mathbf{S} \mathbf{Q}_q)$. Zauważmy, że zaletą powyższej parametryzacji jest jej liniowość ze względu na γ i δ , oraz liniowość estymatorów $\hat{\gamma}$ i $\hat{\delta}$ ze względu na estymator macierzy kowariancji bez struktury, \mathbf{S} .

Przyjmując parametryzację (1.4), nieliniową ze względu na σ^2 i ρ , w której bazę ortonormalną stanowią macierze $\frac{1}{\sqrt{q}} \mathbf{I}_q$ i $\frac{1}{\sqrt{q(q-1)}} (\mathbf{1}_q \mathbf{1}'_q - \mathbf{I}_q)$, estymator największej wiarygodności przyjmuje postać

$$\hat{\Sigma}_{\text{CS}} = \frac{1}{q} \text{Tr } \mathbf{S} \cdot \mathbf{I}_q + \frac{1}{q(q-1)} \text{Tr} \left[\mathbf{S} (\mathbf{1}_q \mathbf{1}'_q - \mathbf{I}_q) \right] (\mathbf{1}_q \mathbf{1}'_q - \mathbf{I}_q), \quad (2.5)$$

a stąd, $\hat{\sigma}^2 = \frac{1}{q} \text{Tr } \mathbf{S}$ oraz $\hat{\rho} = \frac{1}{q(q-1)} \text{Tr}[\mathbf{S}(\mathbf{1}_q \mathbf{1}'_q - \mathbf{I}_q)] / (\frac{1}{q} \text{Tr } \mathbf{S})$.

Zwróćmy uwagę, że obie parametryzacje macierzy Σ_{CS} prowadzą ostatecznie do takiej samej postaci estymatora największej wiarygodności; z zależności między parametryzacją (1.4) i (1.5) otrzymujemy $\hat{\gamma} = \hat{\sigma}^2 [1 + (q-1)\hat{\rho}] = \text{Tr}(\mathbf{S} \mathbf{P}_q)$ oraz $\hat{\delta} = \hat{\sigma}^2 (1 - \hat{\rho}) = \frac{1}{q-1} \text{Tr}(\mathbf{S} \mathbf{Q}_q)$.

Pokazaliśmy w ten sposób, że mimo że istnienie bazy ortogonalnej w ścisłym sensie, której elementami są macierze idempotentne, jest warunkiem koniecznym zastosowania Twierdzenia 2.2, gdyż determinuje kwadratowość podprzestrzeni, jej wykorzystanie do estymacji największej wiarygodności nie jest już konieczne – w procesie rzutowania można posłużyć się dowolną bazą ortonormalną.

Założmy teraz, że $\Sigma = \Sigma_{CT}$. Podobnie jak w poprzednim przypadku baza ortonormalna może być zdefiniowana przez znormalizowane wektory ℓ_i (por. Rozdział 1.2) lub znormalizowane macierze \mathbf{I}_q oraz $\mathbf{C}^i + \mathbf{C}^{i'}$; por. (1.6). Wówczas, bezpośrednio z (2.4) otrzymujemy

$$\widehat{\Sigma}_{CT} = \begin{cases} \frac{1}{q} \text{Tr}(\mathbf{S}\mathbf{1}_q\mathbf{1}'_q)\mathbf{1}_q\mathbf{1}'_q \\ + \frac{1}{2} \sum_{i=1}^{\omega-1} \text{Tr}[\mathbf{S}(\ell_i : \ell_{q-i+2})(\ell_i : \ell_{q-i+2})'](\ell_i : \ell_{q-i+2})(\ell_i : \ell_{q-i+2})' \\ + \frac{1}{2} \text{Tr}[\mathbf{S}(\ell_\omega : \ell_{\omega+1})(\ell_\omega : \ell_{\omega+1})'](\ell_\omega : \ell_{\omega+1})(\ell_\omega : \ell_{\omega+1})' & \text{dla } q \text{ nieparzystego} \\ \\ \frac{1}{q} \text{Tr}(\mathbf{S}\mathbf{1}_q\mathbf{1}'_q) \\ + \frac{1}{2} \sum_{i=1}^{\omega-1} \text{Tr}[\mathbf{S}(\ell_i : \ell_{q-i+2})(\ell_i : \ell_{q-i+2})'](\ell_i : \ell_{q-i+2})(\ell_i : \ell_{q-i+2})' \\ + \text{Tr}(\mathbf{S}\ell_\omega\ell'_\omega)\ell_\omega\ell'_\omega & \text{dla } q \text{ parzystego,} \end{cases}$$

gdzie j -ta składowa ℓ_i^j , $i = 2, \dots, \omega, j = 1, \dots, q$, jest postaci (1.8).

Analogicznie, estymator największej wiarygodności macierzy Σ_{CT} jest rzutem \mathbf{S} na przestrzeń \mathcal{V} generowaną przez macierze z parametryzacji (1.6) w następującej formie

$$\widehat{\Sigma}_{CT} = \begin{cases} \frac{1}{q} \text{Tr} \mathbf{S} \cdot \mathbf{I}_q + \frac{1}{2q} \sum_{i=1}^{\lfloor q/2 \rfloor} \left\{ \text{Tr} [\mathbf{S}(\mathbf{C}^i + \mathbf{C}^{i'})](\mathbf{C}^i + \mathbf{C}^{i'}) \right\} & \text{dla } q \text{ nieparzystego} \\ \\ \frac{1}{q} \text{Tr} \mathbf{S} \cdot \mathbf{I}_q + \frac{1}{2q} \sum_{i=1}^{q/2-1} \left\{ \text{Tr} [\mathbf{S}(\mathbf{C}^i + \mathbf{C}^{i'})](\mathbf{C}^i + \mathbf{C}^{i'}) + \text{Tr} (\mathbf{S}\mathbf{C}^{q/2})\mathbf{C}^{q/2} \right\} & \text{dla } q \text{ parzystego.} \end{cases}$$

2.2.2 Model podwójnie wielowymiarowy

Rozpatrzmy teraz model podwójnie wielowymiarowy, (1.1). Wektoryzacja tego modelu sprowadza go do postaci

$$\text{vec } \mathbf{Y} \sim N_{mqn}((\mathbf{I}_{mq} \otimes \mathbf{1}')\boldsymbol{\mu}, \boldsymbol{\Omega} \otimes \mathbf{I}_n),$$

przy czym estymatorem największej wiarygodności $\mathbf{\Omega}$ jest $\mathbf{S} = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$. Podobnie jak w przypadku modelu wielowymiarowego (dla $m = 1$)

$$\mathbf{P}_{\mathbf{I}_{mq} \otimes \mathbf{1}_n} = (\mathbf{I}_{mq} \otimes \mathbf{1}_n) [(\mathbf{I}_{mq} \otimes \mathbf{1}'_n)(\mathbf{I}_{mq} \otimes \mathbf{1}_n)]^{-1} (\mathbf{I}_{mq} \otimes \mathbf{1}'_n) = \mathbf{I}_{mq} \otimes \frac{1}{n} \mathbf{1}_n \mathbf{1}'_n = \mathbf{I}_{mq} \otimes \mathbf{P}_n,$$

a stąd $\mathbf{P}_{\mathbf{I}_{mq} \otimes \mathbf{1}_n}(\mathbf{\Omega} \otimes \mathbf{I}_n) = \mathbf{\Omega} \otimes \mathbf{P}_n = (\mathbf{\Omega} \otimes \mathbf{I}_n) \mathbf{P}_{\mathbf{I}_{mq} \otimes \mathbf{1}_n}$, co dowodzi komutatywności przestrzeni wartości oczekiwanej i macierzy kowariancji niezależnie od wyboru potencjalnej struktury kowariancyjnej, a więc zachodzi również dla $\mathbf{\Sigma} = \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \mathbf{\Delta}_i$; por. (1.11).

Niech $v_i = \text{Tr} \mathbf{V}_i$ oraz niech macierze \mathbf{A}_{ij} stanowią bazę przestrzeni macierzy symetrycznych, taką, że $\mathbf{A}_{jj} = \mathbf{e}_j \mathbf{e}'_j$, $\mathbf{A}_{jk} = \frac{1}{\sqrt{2}}(\mathbf{e}_j \mathbf{e}'_k + \mathbf{e}_k \mathbf{e}'_j)$, gdzie \mathbf{e}_j jest wektorem o j -tej składowej równej 1 i pozostałych składowych równych zero. Wówczas macierze $v_i^{-1/2} \mathbf{V}_i \otimes \mathbf{A}_{jk}$, $i = 1, \dots, \nu$, $j, k = 1, \dots, \omega$, tworzą bazę ortonormalną podprzestrzeni kwadratowej do której należy macierz $\mathbf{\Sigma}$. Korzystając z Twierdzenia 2.2 otrzymujemy

$$\begin{aligned} \hat{\mathbf{\Sigma}} &= \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \sum_{k=1}^j \langle \mathbf{S}^*, \frac{1}{\sqrt{v_i}} (\mathbf{V}_i \otimes \mathbf{A}_{jk}) \rangle \frac{1}{\sqrt{v_i}} (\mathbf{V}_i \otimes \mathbf{A}_{jk}) \\ &= \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \sum_{k=1}^j \frac{1}{v_i} \text{Tr}[\mathbf{S}^* (\mathbf{V}_i \otimes \mathbf{A}_{jk})] (\mathbf{V}_i \otimes \mathbf{A}_{jk}), \end{aligned} \quad (2.6)$$

gdzie $\mathbf{S}^* = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y} = \mathbf{S}$.

Przyjmijmy, że $\mathbf{\Sigma} = \mathbf{\Sigma}_{\text{BI}}$. Wówczas bezpośrednio z (2.6) otrzymujemy

$$\hat{\mathbf{\Sigma}}_{\text{BI}} = \sum_{j=1}^m \sum_{k=1}^j \frac{1}{q} \text{Tr}[\mathbf{S} (\mathbf{I}_q \otimes \mathbf{A}_{jk})] (\mathbf{I}_q \otimes \mathbf{A}_{jk}).$$

Ponieważ z powyższego wzoru niełatwo dostrzec postać estymatora $\mathbf{\Delta}$, wyznaczmy go korzystając z równoważnego podejścia wykorzystującego minimalizację normy Frobeniusa; por. Rozdział 1.4. Zauważmy, że

$$\|\mathbf{S} - \mathbf{\Sigma}_{\text{BI}}\|^2 = \|\mathbf{S} - \mathbf{I}_q \otimes \mathbf{\Delta}\|_F^2 = \text{Tr} [\mathbf{S}^2 - 2\mathbf{S}(\mathbf{I}_q \otimes \mathbf{\Delta}) + \mathbf{I}_q \otimes \mathbf{\Delta}^2].$$

Różniczkując powyższe równanie względem macierzy $\mathbf{\Delta}$ i korzystając z Lematu 1.3, Własności 1.1(i), definicji projektorów $\mathbf{P}_q, \mathbf{Q}_q$ i przyrównując otrzymaną pochodną

do zera, otrzymujemy

$$[-2(\text{vec}' \mathbf{I}_q \otimes \mathbf{I}_{m^2})(\mathbf{I}_q \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_m) \text{vec } \mathbf{S} + q(\mathbf{\Delta} \otimes \mathbf{I}_m + \mathbf{I}_m \otimes \mathbf{\Delta}) \text{vec } \mathbf{I}_m]' \mathbf{D}_m = \mathbf{0},$$

co, na mocy Własności 1.2(i) i ponownie Własności 1.1(i), jest równoważne równaniu

$$(-\text{vec}\{\text{BTr}_m[(\mathbf{I}_q \otimes \mathbf{I}_m)\mathbf{S}]\} + q \text{vec } \mathbf{\Delta})' \mathbf{D}_m = \mathbf{0}$$

i ostatecznie

$$\widehat{\mathbf{\Delta}} = \frac{1}{q} \text{BTr}_m \mathbf{S}. \quad (2.7)$$

Założmy teraz, że $\mathbf{\Sigma} = \mathbf{\Sigma}_{\text{BCS}}$. Bazę ortonormalną dla takiej struktury stanowią macierze $\mathbf{P}_q \otimes \mathbf{A}_{jk}$ oraz $\frac{1}{\sqrt{q-1}} \mathbf{Q}_q \otimes \mathbf{A}_{jk}$, $j = 1, 2, \dots, m$, $k = 1, 2, \dots, j$. Bezpośrednio ze wzoru (2.6) otrzymujemy estymator największej wiarygodności dla BCS jako rzut ortogonalny \mathbf{S} na przestrzeń macierzy $\mathbf{\Sigma}_{\text{BCS}}$ postaci

$$\widehat{\mathbf{\Sigma}}_{\text{BCS}} = \sum_{j=1}^m \sum_{k=1}^j \left\{ \text{Tr} [\mathbf{S} (\mathbf{P}_q \otimes \mathbf{A}_{jk})] (\mathbf{P}_q \otimes \mathbf{A}_{jk}) + \frac{1}{q-1} \text{Tr} [\mathbf{S} (\mathbf{Q}_q \otimes \mathbf{A}_{jk})] (\mathbf{Q}_q \otimes \mathbf{A}_{jk}) \right\}.$$

Podobnie jak w poprzednim przykładzie, ponieważ postać estymatorów największej wiarygodności macierzy $\mathbf{\Delta}_1$ i $\mathbf{\Delta}_2$ nie wynika bezpośrednio z powyższego wzoru, wykorzystamy ponownie definicję rzutu ortogonalnego za pomocą minimum odpowiedniej normy Frobeniusa. Uwzględniając tę samą parametryzację dostajemy

$$\begin{aligned} \|\mathbf{S} - \mathbf{\Sigma}_{\text{BCS}}\|^2 &= \|\mathbf{S} - \mathbf{Q}_q \otimes \mathbf{\Delta}_1 - \mathbf{P}_q \otimes \mathbf{\Delta}_2\|_F^2 \\ &= \text{Tr} \left[\mathbf{S}^2 - 2\mathbf{S}(\mathbf{Q}_q \otimes \mathbf{\Delta}_1) - 2\mathbf{S}(\mathbf{P}_q \otimes \mathbf{\Delta}_2) + \mathbf{Q}_q \otimes \mathbf{\Delta}_1^2 + \mathbf{P}_q \otimes \mathbf{\Delta}_2^2 \right]. \end{aligned}$$

Różniczkując powyższe odpowiednio ze względu na $\mathbf{\Delta}_1$ i $\mathbf{\Delta}_2$ oraz mając na uwadze Lemat 1.3, Własność 1.1(i), definicje projektorów $\mathbf{P}_q, \mathbf{Q}_q$ i przyrównując otrzymane pochodne do zera, otrzymujemy równania

$$\begin{aligned} &[-2(\text{vec}' \mathbf{Q}_q \otimes \mathbf{I}_{m^2})(\mathbf{I}_q \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_m) \text{vec } \mathbf{S} \\ &\quad + (q-1)(\mathbf{\Delta}_1 \otimes \mathbf{I}_m + \mathbf{I}_m \otimes \mathbf{\Delta}_1) \text{vec } \mathbf{I}_m]' \mathbf{D}_m = \mathbf{0} \\ &[-2(\text{vec}' \mathbf{P}_q \otimes \mathbf{I}_{m^2})(\mathbf{I}_q \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_m) \text{vec } \mathbf{S} \\ &\quad + (\mathbf{\Delta}_2 \otimes \mathbf{I}_m + \mathbf{I}_m \otimes \mathbf{\Delta}_2) \text{vec } \mathbf{I}_m]' \mathbf{D}_m = \mathbf{0}, \end{aligned}$$

które, na mocy Własności 1.2(i) i ponownie Własności 1.1(i), są równoważne

$$\begin{aligned} [-\text{vec}\{\text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{S}]\} + (q-1)\text{vec}\mathbf{\Delta}_1]'\mathbf{D}_m &= \mathbf{0} \\ [-\text{vec}\{\text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{S}]\} + \text{vec}\mathbf{\Delta}_2]'\mathbf{D}_m &= \mathbf{0} \end{aligned}$$

i ostatecznie

$$\begin{cases} \widehat{\mathbf{\Delta}}_1 &= \frac{1}{q-1} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{S}] \\ \widehat{\mathbf{\Delta}}_2 &= \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{S}]. \end{cases} \quad (2.8)$$

Warto zauważyć, że otrzymane estymatory są dokładnie takie same, jak te uzyskane przez różniczkowanie logarytmu funkcji wiarygodności, przedstawione m.in. w pracy Roy i in. (2018).

W przypadku parametryzacji za pomocą macierzy $\mathbf{\Gamma}_0$ i $\mathbf{\Gamma}_1$, por. (1.9a), po prostych przekształceniach estymatory największej wiarygodności przyjmują postać

$$\begin{cases} \widehat{\mathbf{\Gamma}}_0 &= \frac{1}{q}[(q-1)\widehat{\mathbf{\Delta}}_1 + \widehat{\mathbf{\Delta}}_2] = \frac{1}{q} \text{BTr}_m(\mathbf{S}) \\ \widehat{\mathbf{\Gamma}}_1 &= \frac{1}{q}(\widehat{\mathbf{\Delta}}_2 - \widehat{\mathbf{\Delta}}_1) = \frac{1}{q(q-1)}[\text{BSum}_m(\mathbf{S}) - \text{BTr}_m(\mathbf{S})]. \end{cases} \quad (2.9)$$

Z postaci (2.9) łatwo zauważyć, że aby otrzymać estymatory największej wiarygodności bloków macierzy kowariancji o strukturze BCS wystarczy uśrednić wszystkie bloki znajdujące się na przekątnej głównej macierzy \mathbf{S} oraz wszystkie te, które znajdują się poza przekątną główną macierzy \mathbf{S} . Podobną tezę można znaleźć również w pracy Szatrowski (1980).

2.2.3 Model krzywych wzrostu

Na koniec, rozważmy model (1.3), w którym $\mathbb{E}(\mathbf{Y}) = \mathbf{X}\mathbf{B}\mathbf{Z}$ i $\text{Cov}(\mathbf{Y}) = \mathbf{\Xi}$, gdzie \mathbf{X} i \mathbf{Z} są znanymi macierzami układu, a \mathbf{B} jest macierzą nieznanymi parametrów. Wektoryzacja tego modelu sprowadza go do postaci

$$\text{vec}\mathbf{Y} \sim N_{mn}((\mathbf{Z}' \otimes \mathbf{X}) \text{vec}\mathbf{B}, \mathbf{\Xi}).$$

Jeśli założymy, że macierz kowariancji nie posiada struktury, to estymatorem największej wiarygodności $\mathbf{\Xi}$ jest

$$\mathbf{S}^* = \mathbf{Q}_{\mathbf{Z}' \otimes \mathbf{X}} \mathbf{y} \mathbf{y}' \mathbf{Q}_{\mathbf{Z}' \otimes \mathbf{X}}, \quad (2.10)$$

gdzie $\mathbf{y} = \text{vec } \mathbf{Y}$.

Założmy teraz, że macierz kowariancji posiada strukturę $\mathbf{\Omega} \otimes \mathbf{I}_n$, wynikającą z niezależności obserwacji. Wówczas powyższy model sprowadza się do modelu krzywych wzrostu postaci (1.3), w którym estymatorem największej wiarygodności macierzy $\mathbf{\Omega}$ jest \mathbf{S}^{**} postaci

$$\mathbf{S}^{**} = \frac{1}{n} \left(\mathbf{Y}' \mathbf{Q}_X \mathbf{Y} + \mathbf{Q}_{Z'; \tilde{\mathbf{S}}^{-1}} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} \mathbf{Q}_{Z'; \tilde{\mathbf{S}}^{-1}} \right), \quad (2.11)$$

gdzie $\tilde{\mathbf{S}} = \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$ oraz $\mathbf{Q}_{Z'; \tilde{\mathbf{S}}^{-1}} = \mathbf{I}_m - \mathbf{Z}' (\mathbf{Z} \tilde{\mathbf{S}}^{-1} \mathbf{Z}')^{-1} \mathbf{Z} \tilde{\mathbf{S}}^{-1}$; por. Kollo, von Rosen (2005), Ohlson, von Rosen (2010), Kopčová, Žežula (2020).

Aby zastosować Twierdzenie 2.2, musi zachodzić komutatywność przestrzeni wartości oczekiwanej, $\mathbf{P}_{Z' \otimes X}$, z przestrzenią macierzy kowariancji. Zauważmy, że w rozważanym przypadku $\mathbf{P}_{Z' \otimes X} (\mathbf{\Omega} \otimes \mathbf{I}_n) = \mathbf{P}_{Z'} \mathbf{\Omega} \otimes \mathbf{P}_X \neq \mathbf{\Omega} \mathbf{P}_{Z'} \otimes \mathbf{P}_X = (\mathbf{\Omega} \otimes \mathbf{I}_n) \mathbf{P}_{Z' \otimes X}$. Założmy więc dodatkowo, że macierz kowariancji $\mathbf{\Omega}$ posiada określoną strukturę, tj. $\mathbf{\Omega} = \mathbf{\Sigma}_{CS}$ oraz $\mathbf{1}_m \in \mathcal{C}(\mathbf{Z}')$. Wówczas

$$\begin{aligned} \mathbf{P}_{Z' \otimes X} (\mathbf{\Sigma}_{CS} \otimes \mathbf{I}_n) &= \mathbf{P}_{Z'} (\lambda_1 \mathbf{P}_m + \lambda_2 \mathbf{Q}_m) \otimes \mathbf{P}_X = (\lambda_1 \mathbf{P}_{Z'} \mathbf{P}_m + \lambda_2 \mathbf{P}_{Z'} \mathbf{Q}_m) \otimes \mathbf{P}_X \\ &= (\lambda_1 \mathbf{P}_{Z'} + \mathbf{0}) \otimes \mathbf{P}_X = (\lambda_1 \mathbf{P}_m \mathbf{P}_{Z'} + \lambda_2 \mathbf{Q}_m \mathbf{P}_{Z'}) \otimes \mathbf{P}_X \\ &= (\mathbf{\Sigma}_{CS} \otimes \mathbf{I}_n) \mathbf{P}_{Z' \otimes X}. \end{aligned}$$

Ponieważ $\mathbf{\Sigma}_{CS} \otimes \mathbf{I}_n$ jest podprzestrzenią kwadratową możemy zastosować Twierdzenie 2.2.

Ze względu na skomplikowaną procedurę, rzutowanie odbędzie się w dwóch krokach. Na początku zrzutujemy estymator największej wiarygodności macierzy kowariancji bez struktury, \mathbf{S}^* , dany wzorem (2.10) na przestrzeń macierzy o strukturze iloczynu Kroneckera, $\mathbf{\Omega} \otimes \mathbf{I}_n$. Różniczkując

$$\|\mathbf{S}^* - \mathbf{\Omega} \otimes \mathbf{I}_n\|_F^2 = \text{Tr}(\mathbf{S}^{*2}) - 2 \text{Tr}[\mathbf{S}^* (\mathbf{\Omega} \otimes \mathbf{I}_n)] + \text{Tr}(\mathbf{\Omega}^2) \cdot \text{Tr } \mathbf{I}_n$$

ze względu na $\mathbf{\Omega}$, korzystając z Lematu 1.3, otrzymujemy

$$-2 [(\mathbf{I}_{m^2} \otimes \text{vec}' \mathbf{I}_n) (\mathbf{I}_m \otimes \mathbf{K}_{n,m} \otimes \mathbf{I}_n) \text{vec } \mathbf{S}^* + n (\mathbf{\Omega} \otimes \mathbf{I}_m + \mathbf{I}_m \otimes \mathbf{\Omega}) \text{vec } \mathbf{I}_m] \mathbf{D}_m = \mathbf{0}$$

i na podstawie Własności 1.1(*iii*)

$$n \operatorname{vec} \boldsymbol{\Omega} = (\mathbf{I}_{m^2} \otimes \operatorname{vec}' \mathbf{I}_n)(\mathbf{I}_m \otimes \mathbf{K}_{n,m} \otimes \mathbf{I}_n) \operatorname{vec} \mathbf{S}^*.$$

Korzystając z Własności 1.3(*i*), otrzymujemy

$$n\boldsymbol{\Omega} = \operatorname{PTr}_n \mathbf{S}^*.$$

Wstawiając za \mathbf{S}^* postać (2.10) oraz korzystając z Własności 1.1(*iii*), Własności 1.3(*ii*),

$$\begin{aligned} \operatorname{PTr}_n \mathbf{S}^* &= \operatorname{PTr}_n \left(\mathbf{Q}_{Z' \otimes X} \operatorname{vec} \mathbf{Y} \operatorname{vec}' \mathbf{Y} \mathbf{Q}_{Z' \otimes X} \right) \\ &= \operatorname{PTr}_n [(\mathbf{I}_{mn} - \mathbf{P}_{Z'} \otimes \mathbf{P}_X) \operatorname{vec} \mathbf{Y} \operatorname{vec}' \mathbf{Y} (\mathbf{I}_{mn} - \mathbf{P}_{Z'} \otimes \mathbf{P}_X)] \\ &= \operatorname{PTr}_n \{[\operatorname{vec} \mathbf{Y} - \operatorname{vec}(\mathbf{P}_X \mathbf{Y} \mathbf{P}_{Z'})][\operatorname{vec}' \mathbf{Y} - \operatorname{vec}'(\mathbf{P}_X \mathbf{Y} \mathbf{P}_{Z'})]\} \\ &= \mathbf{Y}' \mathbf{Y} - \mathbf{Y}' \mathbf{P}_X \mathbf{Y} \mathbf{P}_{Z'} - \mathbf{P}_{Z'} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} + \mathbf{P}_{Z'} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} \mathbf{P}_{Z'} \\ &= \mathbf{Y}' \mathbf{Q}_X \mathbf{Y} + \mathbf{Q}_{Z'} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} - \mathbf{Q}_{Z'} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} \mathbf{P}_{Z'} \\ &= \mathbf{Y}' \mathbf{Q}_X \mathbf{Y} + \mathbf{Q}_{Z'} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} \mathbf{Q}_{Z'}, \end{aligned}$$

a stąd $\tilde{\boldsymbol{\Omega}} = \frac{1}{n}(\mathbf{Y}' \mathbf{Q}_X \mathbf{Y} + \mathbf{Q}_{Z'} \mathbf{Y}' \mathbf{P}_X \mathbf{Y} \mathbf{Q}_{Z'}) =: \mathbf{S}^{***}$.

W drugim kroku zrzutujemy $\tilde{\boldsymbol{\Omega}}$ na przestrzeń $\Sigma_{\text{CS}} \otimes \mathbf{I}_n$. Przyjmując za bazę ortonormalną przestrzeni CS macierze $\frac{1}{\sqrt{m}} \mathbf{I}_m$ i $\frac{1}{\sqrt{m(m-1)}} (\mathbf{1}_m \mathbf{1}'_m - \mathbf{I}_m)$, na podstawie Twierdzenia 2.2, estymator największej wiarygodności macierzy Σ_{CS} ma postać

$$\hat{\Sigma}_{\text{CS}} = \frac{1}{m} \operatorname{Tr} \mathbf{S}^{***} \cdot \mathbf{I}_m + \frac{1}{m(m-1)} \operatorname{Tr} [\mathbf{S}^{***} (\mathbf{1}_m \mathbf{1}'_m - \mathbf{I}_m)] (\mathbf{1}_m \mathbf{1}'_m - \mathbf{I}_m).$$

Zauważmy, że powyższy wzór jest analogiczny do (2.5) z tą różnicą, że rolę \mathbf{S} pełni \mathbf{S}^{***} . Ponadto, otrzymany wynik jest taki sam jak w Žežula (2006). Warto również zwrócić uwagę, że estymatorem największej wiarygodności macierzy $\boldsymbol{\Omega}$ jest \mathbf{S}^{**} zdefiniowany w (2.11), natomiast \mathbf{S}^{***} jest rzutem macierzy kowariancji bez żadnej struktury na przestrzeń $\boldsymbol{\Omega} \otimes \mathbf{I}_n$.

3 Testowanie hipotez dotyczących struktur kowariancji należących do podprzestrzeni kwadratowych w modelu podwójnie wielowymiarowym

Celem niniejszego rozdziału jest zaproponowanie testów statystycznych dla hipotezy (3.1) uwzględniającej zależności wewnątrz- i międzyklasowe należące do pewnych podprzestrzeni kwadratowych. Wyprowadzimy statystykę LRT oraz jej rozkład prawdopodobieństwa, jak również wyznaczymy statystykę RST. Dodatkowo, przy pomocy badań symulacyjnych porównamy asymptotyczną zbieżność rozkładów wyżej wymienionych statystyk testowych do granicznego rozkładu chi-kwadrat z odpowiednią liczbą stopni swobody. Prezentowane testy zostaną również porównane między sobą pod kątem ich mocy, przy czym do analizy rozbieżności pomiędzy rozważanymi hipotezami wykorzystamy entropijną funkcję straty; por. (1.16). Na koniec zaproponowane testy statystyczne zastosujemy do danych rzeczywistych; por. Rozdział 1.5.

W niniejszym rozdziale nie został uwzględniony test Walda ze względu na fakt, że niezbędne byłoby tu precyzyjne określenie hipotezy złożonej (zdefiniowanie funkcji $\mathbf{w}(\boldsymbol{\theta})$ we wzorze (1.15)), co w przypadku hipotez dotyczących ogólnych struktur kowariancyjnych, jest problematyczne.

Zaprezentowane w tym rozdziale wyniki znaleźć można w zaakceptowanej do publikacji pracy Filipiak, John, Liang (2024).

3.1 Model i hipoteza

Rozważmy standardowy model podwójnie wielowymiarowy (1.1) z symetryczną, określoną dodatnio macierzą kowariancji $\boldsymbol{\Sigma}$, należącą do podprzestrzeni kwadratowej o strukturze blokowej reprezentującej zależności międzyklasowe.

Przedmiotem zainteresowania będzie hipoteza (1.12), w której zależność wewnątrzklasowa posiada również pewien wzorzec (strukturę kowariancyjną) należący do komutatywnej podprzestrzeni kwadratowej. Zatem na podstawie notacji wprowadzonej

w Rozdziale 1.2 rozważać będziemy hipotezę postaci

$$H_0 : \Sigma_0 = \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \lambda_{ij} (\mathbf{V}_i \otimes \mathbf{U}_j) \quad \text{vs} \quad H_1 : \Sigma_1 = \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \Delta_i, \quad (3.1)$$

gdzie $\{\mathbf{U}_1, \dots, \mathbf{U}_\omega\}$ jest bazą komutatywnej podprzestrzeni kwadratowej \mathcal{U} macierzy stopnia q , a $\{\mathbf{V}_1, \dots, \mathbf{V}_\nu\}$ jest bazą komutatywnej podprzestrzeni kwadratowej \mathcal{V} macierzy stopnia m .

Należy zwrócić uwagę, że struktura międzyklasowa jest tutaj taka sama zarówno w hipotezie zerowej jak i alternatywnej. Ze względu na szczególny charakter zależności występujących w macierzy kowariancji, tj. zależności międzyklasowych oraz wewnątrzklasowych wyrażonych za pomocą iloczynów Kroneckera, por. (3.1), przestrzenie liniowe \mathcal{W}_0 oraz \mathcal{W}_1 występujące w hipotezie (1.12) oznaczać będziemy jako

$$\begin{aligned} \mathcal{W}_0 &= \mathcal{V} \boxtimes \mathcal{U}, \\ \mathcal{W}_1 &= \mathcal{V} \boxtimes (\mathcal{A} \setminus \mathcal{U}), \end{aligned}$$

gdzie \mathcal{V} i \mathcal{U} są komutatywnymi podprzestrzeniami kwadratowymi macierzy stopnia q i m odpowiednio, takimi, że $\dim(\mathcal{V}) = \nu$ i $\dim(\mathcal{U}) = \omega$, natomiast \mathcal{A} jest dowolną podprzestrzenią $m(m+1)/2$ -wymiarowej przestrzeni macierzy symetrycznych stopnia m , gdzie $\mathcal{U} \subset \mathcal{A}$. Symbol \boxtimes oznacza iloczyn przestrzeni liniowych w sensie iloczynu Kroneckera, to znaczy $\mathcal{V} \boxtimes \mathcal{U} = \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \lambda_{ij} \mathbf{V}_i \otimes \mathbf{U}_j$; por. (3.1). Zatem hipotezę (1.12) możemy również zapisać jako

$$H_0 : \Sigma_0 \in \mathcal{V} \boxtimes \mathcal{U} \quad \text{vs} \quad H_1 : \Sigma_1 \in \mathcal{V} \boxtimes (\mathcal{A} \setminus \mathcal{U}). \quad (3.2)$$

Można zauważyć, że $\mathcal{V} \boxtimes (\mathcal{A} \setminus \mathcal{U})$ jest nadal kwadratową przestrzenią, ale niekoniecznie komutatywną.

Zwróćmy uwagę, że do przestrzeni zdefiniowanej w hipotezie H_0 należy wiele różnych struktur kowariancyjnych, przykładowo:

- macierze o strukturze BD z blokami posiadającymi strukturę D, CS lub CT (oznaczone odpowiednio jako $\Sigma_{\text{BD-D}}$, $\Sigma_{\text{BD-CS}}$, oraz $\Sigma_{\text{BD-CT}}$);

- macierze o strukturze BCS z tymi samymi blokami na przekątnej głównej oraz tymi samymi blokami poza przekątną główną, posiadającymi strukturę D, CS czy też CT (oznaczone odpowiednio jako $\Sigma_{\text{BCS_D}}$, $\Sigma_{\text{BCS_CS}}$, oraz $\Sigma_{\text{BCS_CT}}$);
- macierze o strukturze BCT z tymi samymi blokami na przekątnej głównej oraz z tymi samymi blokami na i -tych przekątnych blokowych, $i = 1, \dots, \frac{q}{2} - 1$ dla q parzystego i $i = 1, \dots, \lfloor \frac{q}{2} \rfloor$ dla q nieparzystego, posiadającymi strukturę D, CS lub CT (oznaczone odpowiednio jako $\Sigma_{\text{BCT_D}}$, $\Sigma_{\text{BCT_CS}}$, oraz $\Sigma_{\text{BCT_CT}}$);
- iloczyn Kroneckera dowolnej kombinacji struktur D, CS, czy CT, przykładowo $\text{CT} \otimes \text{CS}$, $\text{CT} \otimes \text{CT}$.

Należy również wspomnieć, że hipoteza (3.2) nie obejmuje struktury rozdzielnej (tzn. iloczynu Kroneckera dwóch dowolnych określonych dodatnio macierzy symetrycznych), powszechnie stosowanej do analizy modeli podwójnie wielowymiarowych, jako że nie jest to struktura liniowa.

3.2 Statystyki testowe

W tym rozdziale wyprowadzimy statystykę ilorazu wiarygodności i jej dokładny rozkład, a następnie wyznaczmy postać statystyki wynikowej Rao. Rozważania związane z testowaniem struktury kowariancyjnej rozpoczniemy od estymacji wartości oczekiwanej, mimo że nie nakładamy na nią żadnych restrykcji, oraz macierzy kowariancji bez struktury, jako że estymatory te będą niezbędne do wyprowadzenia statystyk testowych.

Przypomnijmy, że w modelu (1.1) przestrzeń wartości oczekiwanej komutuje z macierzą kowariancji, niezależnie od struktury (por. Rozdział 2.2.2). Zatem z równań normalnych otrzymujemy, że estymatory największej wiarygodności $\boldsymbol{\mu}$ i $\boldsymbol{\Omega}$ mają postać

$$\hat{\boldsymbol{\mu}} = \frac{1}{n} \mathbf{Y}' \mathbf{1}_n, \quad \hat{\boldsymbol{\Omega}} = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y} =: \mathbf{S}.$$

3.2.1 Test ilorazu wiarygodności

Niech $\mathbf{H} = (\mathbf{H}_1, \dots, \mathbf{H}_\nu)$ oraz $\mathbf{H}_i \mathbf{H}_i' = \mathbf{V}_i$ stanowią bazę przestrzeni \mathcal{V} , gdzie $\mathbf{H}_i' \mathbf{H}_i = \mathbf{I}_{v_i}$ oraz $\mathbf{H}_i' \mathbf{H}_{i'} = \mathbf{0}_{v_i \times v_{i'}}$ dla $i \neq i'$. Rozważmy macierze $(\mathbf{H}_i' \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H}_i \otimes \mathbf{I}_m)$ stopnia $v_i m$, które można wyrazić jako macierze blokowe o blokach stopnia m . Na potrzeby dalszych rozważań zauważmy, że

$$|\mathbf{S}| = |(\mathbf{H}' \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H} \otimes \mathbf{I}_m)|.$$

Niech

$$\mathbf{S}_i = \text{BTr}_m [(\mathbf{H}_i' \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H}_i \otimes \mathbf{I}_m)]. \quad (3.3)$$

Zauważmy, że macierze \mathbf{S}_i mogą być równoważnie zapisane jako $\sum_{j=1}^{v_i} (\mathbf{h}'_{ij} \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{h}_{ij} \otimes \mathbf{I}_m)$, gdzie \mathbf{h}_{ij} jest j -tą kolumną macierzy \mathbf{H}_i lub $\sum_{j=1}^{v_i} \mathbf{W}'_{ij} \mathbf{Q}_n \mathbf{W}_{ij}$, gdzie $\mathbf{W}_{ij} = \mathbf{Y}(\mathbf{h}_{ij} \otimes \mathbf{I}_m)$. Taką definicję macierzy \mathbf{S}_i można znaleźć w pracy Liang i in. (2021). Wówczas udowodnić możemy następujące twierdzenie.

Twierdzenie 3.1. *Estymator największej wiarygodności macierzy Σ dla modelu (1.1) przy prawdziwości hipotezy zerowej i alternatywnej (3.1) przyjmuje postać, odpowiednio,*

$$\begin{aligned} \hat{\Sigma}_0 &= \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \hat{\lambda}_{ij} (\mathbf{V}_i \otimes \mathbf{U}_j), & \text{gdzie} & \quad \hat{\lambda}_{ij} = \frac{1}{v_i \cdot u_j} \text{Tr} (\mathbf{S}_i \mathbf{U}_j), \\ \hat{\Sigma}_1 &= \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \hat{\Delta}_i, & \text{gdzie} & \quad \hat{\Delta}_i = \frac{1}{v_i} \mathbf{S}_i, \end{aligned}$$

przy czym $i = 1, \dots, \nu$, $j = 1, \dots, \omega$.

Dowód. Przy prawdziwości hipotezy zerowej logarytm funkcji wiarygodności można zapisać jako

$$\ln L(\lambda) = -\frac{nqp}{2} \ln(2\pi) - \frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\{ v_i \cdot u_j \ln \lambda_{ij} + \lambda_{ij}^{-1} \text{Tr} [\mathbf{S}(\mathbf{V}_i \otimes \mathbf{U}_j)] \right\},$$

gdzie $\lambda = (\lambda_{ij})$, $i = 1, \dots, \nu$, $j = 1, \dots, \omega$. Na podstawie własności śladu, śladu

blokowego i definicji macierzy \mathbf{S}_i , powyższą funkcję można przekształcić następująco:

$$\begin{aligned}
\ln L(\lambda) &= -\frac{nmq}{2} \ln(2\pi) - \frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\{ v_i \cdot u_j \ln \lambda_{ij} + \lambda_{ij}^{-1} \text{Tr} [\mathbf{S}(\mathbf{V}_i \otimes \mathbf{U}_j)] \right\} \\
&= -\frac{nmq}{2} \ln(2\pi) \\
&\quad - \frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\{ v_i \cdot u_j \ln \lambda_{ij} + \lambda_{ij}^{-1} \text{Tr} [\mathbf{S}(\mathbf{H}_i \otimes \mathbf{I}_m)(\mathbf{I}_q \otimes \mathbf{U}_j)(\mathbf{H}'_i \otimes \mathbf{I}_m)] \right\} \\
&= -\frac{nmq}{2} \ln(2\pi) \\
&\quad - \frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\{ v_i \cdot u_j \ln \lambda_{ij} + \lambda_{ij}^{-1} \text{Tr} [(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{S}(\mathbf{H}_i \otimes \mathbf{I}_m)(\mathbf{I}_q \otimes \mathbf{U}_j)] \right\} \\
&= -\frac{nmq}{2} \ln(2\pi) \\
&\quad - \frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\{ v_i \cdot u_j \ln \lambda_{ij} + \lambda_{ij}^{-1} \text{Tr} [\text{BTr}_m [(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{S}(\mathbf{H}_i \otimes \mathbf{I}_m)] \mathbf{U}_j] \right\} \\
&= -\frac{nmq}{2} \ln(2\pi) - \frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\{ v_i \cdot u_j \ln \lambda_{ij} + \lambda_{ij}^{-1} \text{Tr} (\mathbf{S}_i \mathbf{U}_j) \right\}.
\end{aligned}$$

Różniczkując powyższą funkcję względem λ_{ij} , $i = 1, \dots, \nu$, $j = 1, \dots, \omega$, otrzymamy

$$v_i u_j \frac{1}{\lambda_{ij}} = \frac{1}{\lambda_{ij}^2} \text{Tr} (\mathbf{S}_i \mathbf{U}_j), \text{ a stąd}$$

$$\hat{\lambda}_{ij} = \frac{1}{v_i \cdot u_j} \text{Tr} (\mathbf{S}_i \mathbf{U}_j). \quad (3.4)$$

Założmy teraz, że H_1 jest hipotezą prawdziwą. Z uwagi na to, że macierz Σ_1 można przedstawić jak w hipotezie alternatywnej (3.1), mamy

$$(\mathbf{H}' \otimes \mathbf{I}_m) \Sigma_1 (\mathbf{H} \otimes \mathbf{I}_m) = \sum_{i=1}^{\nu} \mathbf{H}'_i \mathbf{V}_i \mathbf{H}_i \otimes \Delta_i = \sum_{i=1}^{\nu} \mathbf{H}'_i \mathbf{H}_i \mathbf{H}'_i \mathbf{H}_i \otimes \Delta_i = \sum_{i=1}^{\nu} \mathbf{I}_{v_i} \otimes \Delta_i,$$

a w konsekwencji, korzystając z Własności 1.1(v) i faktu, że wyznacznik macierzy blokowo-diagonalnej jest równy iloczynowi wyznaczników jej bloków diagonalnych (por. Kollo, von Rosen, 2005), otrzymujemy

$$|\Sigma_1| = \prod_{i=1}^{\nu} |\Delta_i|^{v_i}. \quad (3.5)$$

Ponadto, ponieważ $\Sigma_1^{-1} = \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \Delta_i^{-1}$, to funkcję $\ln L$ można sprowadzić do

postaci

$$\ln L(\boldsymbol{\mu}, \boldsymbol{\Delta}_1, \dots, \boldsymbol{\Delta}_\nu) = -\frac{nmq}{2} \ln(2\pi) - \frac{n}{2} \sum_{i=1}^{\nu} (v_i \cdot \ln |\boldsymbol{\Delta}_i|) - \frac{n}{2} \sum_{i=1}^{\nu} \text{Tr} [\mathbf{S}(\mathbf{V}_i \otimes \boldsymbol{\Delta}_i^{-1})]. \quad (3.6)$$

Po zróżniczkowaniu (3.6) ze względu na macierze $\boldsymbol{\Delta}_i$, $i = 1, \dots, \nu$, otrzymamy ν równań postaci

$$v_i \cdot \text{vec } \boldsymbol{\Delta}_i = (\text{vec}' \mathbf{V}_i \otimes \mathbf{I}_{m^2})(\mathbf{I}_q \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_m) \text{vec } \mathbf{S}.$$

Wykorzystując Własność 1.2(i) oraz postać macierzy \mathbf{S}_i podaną w (3.3), z powyższego równania otrzymujemy estymator największej wiarygodności bloków macierzy kowariancji postaci

$$\widehat{\boldsymbol{\Delta}}_i = \frac{1}{v_i} \text{BTr}_m [(\mathbf{V}_i \otimes \mathbf{I}_m) \mathbf{S}] = \frac{1}{v_i} \mathbf{S}_i. \quad (3.7)$$

□

Wykorzystując estymatory podane w Twierdzeniu 3.1, wyznaczmy postać statystyki testowej ilorazu wiarygodności oraz wyznaczmy jej rozkład.

Twierdzenie 3.2. *Statystyka ilorazu wiarygodności Λ dla hipotezy (3.1) przyjmuje postać*

$$\Lambda = \left\{ \prod_{i=1}^{\nu} \left[|\mathbf{S}_i|^{v_i} \prod_{j=1}^{\omega} \left(\frac{u_j}{t_{ij}} \right)^{v_i u_j} \right] \right\}^{n/2},$$

gdzie $t_{ij} = \text{Tr}(\mathbf{S}_i \mathbf{U}_j)$ oraz \mathbf{S}_i zdefiniowane jest w (3.3).

Dowód. Niech prawdziwa będzie hipoteza alternatywna. Wówczas,

$$\sup_{H_1} L(\boldsymbol{\mu}, \boldsymbol{\Delta}_1, \dots, \boldsymbol{\Delta}_\nu) = (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \left| \frac{1}{v_i} \mathbf{S}_i \right|^{-nv_i/2} \exp \left\{ -\frac{n}{2} \sum_{i=1}^{\nu} v_i \text{Tr} [\mathbf{S}(\mathbf{V}_i \otimes \mathbf{S}_i^{-1})] \right\}.$$

Korzystając z własności wyznacznika $|a\mathbf{A}| = a^n |\mathbf{A}|$ dla dowolnej, kwadratowej macierzy \mathbf{A} stopnia n , z dekompozycji macierzy $\mathbf{V}_i = \mathbf{H}_i \mathbf{H}_i'$, z Własności 1.2(iii), (iv)

oraz z postaci macierzy \mathbf{S}_i podanej w (3.3), powyższe supremum można przekształcić następująco

$$\begin{aligned}
\sup_{H_1} L(\boldsymbol{\mu}, \boldsymbol{\Delta}_1, \dots, \boldsymbol{\Delta}_\nu) &= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \left(\frac{1}{v_i^m} |\mathbf{S}_i|^{-nv_i/2} \right) \\
&\quad \cdot \exp \left\{ -\frac{n}{2} \sum_{i=1}^{\nu} v_i \operatorname{Tr} \left[(\mathbf{H}'_i \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H}_i \otimes \mathbf{I}_m) (\mathbf{I}_q \otimes \mathbf{S}_i^{-1}) \right] \right\} \\
&= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \left(\frac{1}{v_i^m} |\mathbf{S}_i|^{-nv_i/2} \right) \\
&\quad \cdot \exp \left(-\frac{n}{2} \sum_{i=1}^{\nu} v_i \operatorname{Tr} \left\{ \operatorname{BTr}_m \left[(\mathbf{H}'_i \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H}_i \otimes \mathbf{I}_m) (\mathbf{I}_q \otimes \mathbf{S}_i^{-1}) \right] \right\} \right) \\
&= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \left(\frac{1}{v_i^m} |\mathbf{S}_i|^{-nv_i/2} \right) \\
&\quad \cdot \exp \left(-\frac{n}{2} \sum_{i=1}^{\nu} v_i \operatorname{Tr} \left\{ \operatorname{BTr}_m \left[(\mathbf{H}'_i \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H}_i \otimes \mathbf{I}_m) \right] \mathbf{S}_i^{-1} \right\} \right) \\
&= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \left(\frac{1}{v_i^m} |\mathbf{S}_i|^{-nv_i/2} \right) \cdot \exp \left\{ -\frac{n}{2} \sum_{i=1}^{\nu} v_i \operatorname{Tr} (\mathbf{I}_m) \right\} \\
&= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \left(\frac{1}{v_i^m} |\mathbf{S}_i|^{-nv_i/2} \right) \cdot \exp \left(-\frac{nmq}{2} \right).
\end{aligned}$$

Analogicznie, przy prawdziwości hipotezy zerowej, ponieważ estymator największej wiarygodności wyraża się wzorem (3.4),

$$\begin{aligned}
\sup_{H_0} L(\boldsymbol{\mu}, \lambda) &= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \prod_{j=1}^{\omega} \left(\frac{t_{ij}}{v_i u_j} \right)^{-nv_i u_j / 2} \exp \left(-\frac{n}{2} \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} v_i u_j \right) \\
&= (2\pi)^{-nmq/2} \prod_{i=1}^{\nu} \prod_{j=1}^{\omega} \left(\frac{t_{ij}}{v_i u_j} \right)^{-nv_i u_j / 2} \exp \left(-\frac{nmq}{2} \right).
\end{aligned}$$

Stąd

$$\Lambda = \frac{\sup_{H_0} L(\boldsymbol{\mu}, \lambda)}{\sup_{H_1} L(\boldsymbol{\mu}, \boldsymbol{\Delta}_1, \dots, \boldsymbol{\Delta}_\nu)}$$

daje tezę twierdzenia. \square

Niech $\mathcal{W}_p(\boldsymbol{\Omega}, \text{df})$ oznacza rozkład Wisharta z określoną dodatnio macierzą $\boldsymbol{\Omega}$ stopnia p i liczbą stopni swobody równą df . Ponadto, niech symbol $\stackrel{d}{\sim}$ oznacza, że zmienna losowa "ma taki sam rozkład, jak".

Do wyprowadzenia rozkładu statystyki ilorazu wiarygodności będą pomocne poniższe lematy zaczerpnięte odpowiednio z prac Muirhead (1982) i Olkin, Press (1969).

Lemat 3.1. *Jeśli macierz \mathbf{A} ma rozkład $\mathcal{W}_m(\boldsymbol{\Omega}, n)$, gdzie $n \geq m$, to $|\mathbf{A}| / |\boldsymbol{\Omega}| \stackrel{d}{\sim} \prod_{i=1}^m X_i$, gdzie $X_i \sim \chi_{n-i+1}^2$, $i = 1, \dots, m$.*

Lemat 3.2. *Niech W_0, W_1, \dots, W_m będą niezależnymi zmiennymi losowymi o rozkładach $W_j \sim \chi_{2a_j}^2$, $j = 0, 1, \dots, m$. Jeśli*

$$L = m^m \frac{\prod_{j=1}^m W_j}{(W_0 + \sum_{j=1}^m W_j)^m},$$

to $L \stackrel{d}{\sim} \prod_{j=1}^m X_j$, gdzie X_1, \dots, X_m są niezależnymi zmiennymi losowymi takimi, że

$$X_j \sim \beta(a_j, b_j), \quad b_j = \frac{a + j - 1}{m} - a_j, \quad a = \sum_{j=0}^m a_j.$$

Wykorzystamy również następujący lemat.

Lemat 3.3. *Jeśli macierz \mathbf{A} ma rozkład $\mathcal{W}_{mq}(\boldsymbol{\Sigma}_1, n)$, gdzie $\boldsymbol{\Sigma}_1 = \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \boldsymbol{\Delta}_i$ z $\{\mathbf{V}_1, \dots, \mathbf{V}_{\nu}\}$ jako bazą komutatywnej podprzestrzeni kwadratowej \mathcal{V} macierzy stopnia m , to*

$$\text{BTr}_m[(\mathbf{V}_i \otimes \mathbf{I}_m)\mathbf{A}] \sim \mathcal{W}_{mq}(\boldsymbol{\Delta}_i, nv_i).$$

Dowód. Niech $\mathbf{A} \sim \mathcal{W}_{mq}(\boldsymbol{\Sigma}_1, n)$, gdzie $\boldsymbol{\Sigma}_1$ zdefiniowana jest w hipotezie alternatywnej (3.1). Na podstawie Kollo, von Rosen (2005), oznaczając $\mathbf{H} = (\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_{\nu})$, można zapisać, że

$$(\mathbf{H}' \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H} \otimes \mathbf{I}_m) \sim \mathcal{W}_{mq}(\text{BDiag}(\mathbf{I}_{\nu_i} \otimes \boldsymbol{\Delta}_i), n),$$

gdzie BDiag oznacza macierz blokowo diagonalną. Implikuje to, że diagonalne bloki macierzy $(\mathbf{H}' \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H} \otimes \mathbf{I}_m)$, czyli $(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H}_i \otimes \mathbf{I}_m)$ oraz $(\mathbf{H}'_j \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H}_j \otimes \mathbf{I}_m)$, $i, j = 1, \dots, \nu$, $i \neq j$, są niezależne oraz

$$(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H}_i \otimes \mathbf{I}_m) \sim \mathcal{W}_{\nu_i m}(\mathbf{I}_{\nu_i} \otimes \boldsymbol{\Delta}_i, n), \quad i = 1, \dots, \nu;$$

por. Anderson (2003). Co więcej, ponieważ każda macierz \mathbf{H}_i składa się z kolumn ortogonalnych, $\mathbf{H}_i = (\mathbf{h}_{i1}, \dots, \mathbf{h}_{iv_i})$, bloki diagonalne macierzy $(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H}_i \otimes \mathbf{I}_m)$, dla $i = 1, \dots, \nu$, są niezależne oraz

$$(\mathbf{h}'_{ik} \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{h}_{ik} \otimes \mathbf{I}_m) \sim \mathcal{W}_m(\Delta_i, n), \quad i = 1, \dots, \nu, \quad k = 1, \dots, v_i.$$

Dalej, na podstawie dekompozycji macierzy $\mathbf{V}_i = \mathbf{H}_i\mathbf{H}'_i$ oraz Własności 1.2(v) można zapisać, że

$$\begin{aligned} \text{BTr}_m[(\mathbf{V}_i \otimes \mathbf{I}_m)\mathbf{A}] &= \text{BTr}_m[(\mathbf{H}_i \otimes \mathbf{I}_m)(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{A}] = \text{BTr}_m[(\mathbf{H}'_i \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{H}_i \otimes \mathbf{I}_m)] \\ &= \sum_{k=1}^{v_i} (\mathbf{h}'_{ik} \otimes \mathbf{I}_m)\mathbf{A}(\mathbf{h}_{ik} \otimes \mathbf{I}_m). \end{aligned}$$

Ponieważ suma niezależnych zmiennych losowych o rozkładach Wisharta ma również rozkład Wisharta z odpowiednimi parametrami (por. Anderson, 2003), otrzymujemy tezę lematu. \square

Sformułujemy teraz jeden z najważniejszych wyników niniejszego rozdziału, mianowicie twierdzenie o rozkładzie statystyki ilorazu wiarogodności Λ . Z uwagi na fakt, że w definicji Λ w Twierdzeniu 3.2 występują wyznaczniki macierzy \mathbf{S}_i stopnia m , w celu zapewnienia ich dodatniej określoności zakładamy, że $n > m$. Ponadto niech $\beta(\alpha_1, \alpha_2)$ oznacza rozkład beta z parametrami skali α_1 i α_2 .

Twierdzenie 3.3. *Jeśli hipoteza zerowa w (3.1) jest prawdziwa i $n > m$, to*

$$\Lambda^{2/n} \stackrel{d}{\sim} \begin{cases} \prod_{i=1}^{\nu} \prod_{j=2}^{\omega} \prod_{k=1}^{u_j} B_{ijk}^{v_i} & \text{jeśli } u_1 = 1 \\ \prod_{i=1}^{\nu} \prod_{j=1}^{\omega} \prod_{k=1}^{u_j} B_{ijk}^{v_i} & \text{jeśli } u_1 \geq 2, \end{cases}$$

gdzie

$$B_{ijk} \sim \beta\left(\frac{(n-1)v_i}{2} - \frac{s_{jk} - 1}{2}, \frac{k-1}{u_j} + \frac{s_{jk} - 1}{2}\right)$$

oraz $s_{jk} = k + \sum_{c=0}^{j-1} u_c$, $u_0 = 0$.

Dowód. Przy prawdziwości hipotezy H_0 z (3.2) każdą macierz \mathbf{U}_j generującą bazę podprzestrzeni kwadratowej \mathcal{U} można przedstawić w postaci $\mathbf{G}_j\mathbf{G}'_j$, $j = 1, \dots, \omega$,

gdzie $\mathbf{G}'_j \mathbf{G}_j = \mathbf{I}_{u_j}$ i $\mathbf{G}'_j \mathbf{G}_{j'} = \mathbf{0}_{u_j \times u_{j'}}$. Z ortogonalności macierzy $\mathbf{G} = (\mathbf{G}_1, \dots, \mathbf{G}_\omega)$ wynika, że $|\mathbf{G}' \mathbf{S}_i \mathbf{G}| = |\mathbf{S}_i|$, dla $i = 1, \dots, \nu$. Co więcej, $t_{ij} = \text{Tr}(\mathbf{S}_i \mathbf{U}_j)$ w Twierdzeniu 3.2 są niczym innym jak sumą odpowiednich elementów znajdujących się na przekątnej głównej macierzy $\mathbf{G}' \mathbf{S}_i \mathbf{G}$. Zatem $\Lambda^{2/n}$ można zapisać jako

$$\Lambda^{2/n} = \prod_{i=1}^{\nu} \left\{ \frac{\prod_{j=1}^{\omega} u_j^{u_j} \cdot |\mathbf{G}' \mathbf{S}_i \mathbf{G}|}{\prod_{j=1}^{\omega} \left[\sum_{k=1}^{u_j} (\mathbf{G}' \mathbf{S}_i \mathbf{G})_{s_{jk}, s_{jk}} \right]^{u_j}} \right\}^{v_i}, \quad (3.8)$$

gdzie $s_{jk} = k + \sum_{c=0}^{j-1} u_c$ i $u_0 = 0$. Na podstawie Lematu 3.3 można zauważyć, że

$$\mathbf{S}_i \sim \mathcal{W}_m(\Delta_i, (n-1)v_i),$$

dla $i = 1, \dots, \nu$ są wzajemnie niezależne, a zatem wszystkie czynniki w (3.8) są również niezależne. Ponadto, przy prawdziwości hipotezy zerowej, oznaczając na potrzeby tego dowodu przez \mathbf{D}_i macierz blokowo-diagonalną z macierzami $\delta_{ij} \mathbf{I}_{u_j}$, $j = 1, \dots, \omega$, na przekątnej głównej, dla każdego $i = 1, \dots, \nu$

$$\mathbf{D}_i^{-1/2} \mathbf{G}' \mathbf{S}_i \mathbf{G} \mathbf{D}_i^{-1/2} \sim \mathcal{W}_m(\mathbf{I}_m, (n-1)v_i).$$

Wykorzystując rozkład Choleskiego powyższej macierzy, $\mathbf{W}_i \mathbf{W}'_i$, gdzie \mathbf{W}_i jest macierzą dolnotrójkątną z dodatnimi elementami na przekątnej głównej mamy, że wszystkie elementy macierzy \mathbf{W}_i są niezależne, a niezerowe elementy znajdujące się poza przekątną główną mają rozkład normalny, tj. $W_{i,kk'} \sim N(0, 1)$, $m \geq k > k' \geq 1$ i $W_{i,kk}^2 \sim \chi_{(n-1)v_i - k + 1}^2$; por. Kollo, von Rosen (2005). Zatem,

$$|\mathbf{S}_i| = |\mathbf{G}' \mathbf{S}_i \mathbf{G}| = |\mathbf{D}_i^{1/2} \mathbf{W}_i \mathbf{W}'_i \mathbf{D}_i^{1/2}| = |\mathbf{D}_i| \cdot \prod_{k=1}^m W_{i,kk}^2 = \left(\prod_{j=1}^{\omega} \delta_{ij}^{u_j} \right) \prod_{k=1}^m W_{i,kk}^2.$$

Oznaczając przez $Z_{is_{jk}}^{(1)}$ zmienną losową o rozkładzie $\chi_{(n-1)v_i - s_{jk} + 1}^2$ mamy, że

$$\prod_{k=1}^m W_{i,kk}^2 = \prod_{j=1}^{\omega} \prod_{k=1}^{u_j} Z_{is_{jk}}^{(1)}. \quad (3.9)$$

Rozważmy teraz mianownik (3.8). Korzystając z powyższej notacji możemy zapisać

$$\begin{aligned} \sum_{k=1}^{u_j} (\mathbf{G}'\mathbf{S}_i\mathbf{G})_{s_{jk},s_{jk}} &= \delta_{ij} \sum_{k=1}^{u_j} (\mathbf{W}_i\mathbf{W}'_i)_{s_{jk},s_{jk}} \\ &= \begin{cases} \delta_{ij} W_{i,11}^2 & \text{dla } u_1 = 1 \\ \delta_{ij} \left[\sum_{k=1}^{u_1} W_{i,kk}^2 + \sum_{k=2}^{u_1} \sum_{k'=1}^{k-1} W_{i,kk'}^2 \right] & \text{dla } j = 1 \text{ oraz } u_1 \geq 2 \\ \delta_{ij} \left[\sum_{k=1}^{u_j} W_{i,s_{jk}s_{jk}}^2 + \sum_{k=1}^{u_j} \sum_{k'=1}^{s_{jk}-1} W_{i,s_{jk}k'}^2 \right] & \text{dla } j = 2, \dots, \omega. \end{cases} \end{aligned}$$

Zauważmy, że kwadrat każdego elementu znajdującego się na przekątnej głównej macierzy \mathbf{W}_i ma taki sam rozkład jak $Z_{is_{jk}}^{(1)}$, a kwadrat każdego elementu poza przekątną główną ma rozkład χ_1^2 jako kwadrat zmiennej losowej o standardowym rozkładzie normalnym. Ponadto, wszystkie elementy poza przekątną są niezależne. Zatem, przyjmując że $Z_{ij}^{(2)} \sim \chi_{d_j}^2$, gdzie

$$d_j = \frac{(2s_{j0} + u_j - 1)u_j}{2}, \quad (3.10)$$

otrzymujemy

$$\sum_{k=1}^{u_j} (\mathbf{G}'\mathbf{S}_i\mathbf{G})_{s_{jk},s_{jk}} \stackrel{d}{\sim} \begin{cases} \delta_{ij} Z_{i1}^{(1)} & \text{dla } u_1 = 1 \\ \delta_{ij} \left[\sum_{k=1}^{u_j} Z_{is_{jk}}^{(1)} + Z_{ij}^{(2)} \right] & \text{dla } j = 1, \dots, \omega \text{ oraz } u_1 \geq 2. \end{cases} \quad (3.11)$$

Podstawiając (3.9) i (3.11) do (3.8), otrzymamy

$$\Lambda^{2/n} \stackrel{d}{\sim} \begin{cases} \prod_{i=1}^{\nu} \prod_{j=2}^{\alpha} \left\{ \frac{u_j \cdot \prod_{k=1}^{u_j} Z_{is_{jk}}^{(1)}}{\left[\sum_{k=1}^{u_j} Z_{is_{jk}}^{(1)} + Z_{ij}^{(2)} \right]^{u_j}} \right\}^{v_i} & \text{dla } u_1 = 1 \\ \prod_{i=1}^{\nu} \prod_{j=1}^{\alpha} \left\{ \frac{u_j \cdot \prod_{k=1}^{u_j} Z_{is_{jk}}^{(1)}}{\left[\sum_{k=1}^{u_j} Z_{is_{jk}}^{(1)} + Z_{ij}^{(2)} \right]^{u_j}} \right\}^{v_i} & \text{dla } u_1 \geq 2. \end{cases}$$

Na podstawie Lematu 3.2 możemy zapisać teraz rozkład $\Lambda^{2/n}$ jako

$$\Lambda^{2/n} \stackrel{d}{\sim} \begin{cases} \prod_{i=1}^{\nu} \prod_{j=2}^{\alpha} \prod_{k=1}^{u_j} B_{ijk}^{v_i} & \text{dla } u_1 = 1 \\ \prod_{i=1}^{\nu} \prod_{j=1}^{\alpha} \prod_{k=1}^{u_j} B_{ijk}^{v_i} & \text{dla } u_1 \geq 2, \end{cases}$$

gdzie

$$B_{ijk} \sim \beta \left(a_{is_{jk}}^{(1)}, \frac{a_{ij} + k - 1}{u_j} - a_{is_{jk}}^{(1)} \right),$$

$$a_{is_{jk}}^{(1)} = \frac{(n-1)v_i - s_{jk} + 1}{2}, \quad a_{ij} = \sum_{k=1}^{u_j} a_{is_{jk}}^{(1)} + \frac{d_j}{2}.$$

Mając na uwadze (3.10) i definicję s_{jk} , wyrażenie a_{ij} można przekształcić do postaci

$$\begin{aligned} a_{ij} &= \sum_{k=1}^{u_j} \frac{(n-1)v_i - s_{jk} + 1}{2} + \frac{(2s_{j0} + u_j - 1)u_j}{4} \\ &= u_j \frac{(n-1)v_i + 1}{2} - \frac{1}{2} \sum_{k=1}^{u_j} s_{jk} + \frac{(2 \sum_{c=0}^{j-1} u_c + u_j - 1)u_j}{4} \\ &= u_j \frac{(n-1)v_i + 1}{2} - \frac{1}{2} \sum_{k=1}^{u_j} \left(\sum_{c=0}^{j-1} u_c + k \right) + \frac{(2 \sum_{c=0}^j u_c - u_j - 1)u_j}{4} \\ &= u_j \frac{(n-1)v_i + 1}{2} - \frac{u_j \cdot \sum_{c=0}^{j-1} u_c}{2} - \frac{u_j(u_j + 1)}{4} + \frac{(2 \sum_{c=0}^j u_c - u_j - 1)u_j}{4} \\ &= \frac{u_j}{2} \left[(n-1)v_i + 1 - \sum_{c=0}^{j-1} u_c - \frac{u_j}{2} - \frac{1}{2} + \sum_{c=0}^j u_c - \frac{u_j}{2} - \frac{1}{2} \right] \\ &= \frac{(n-1)u_j v_i}{2} \end{aligned}$$

skąd otrzymujemy drugi parametr rozkładu beta zgodny z tezą twierdzenia. Istotnie,

$$\begin{aligned} \frac{a_{ij} + k - 1}{u_j} - a_{is_{jk}}^{(1)} &= \frac{(n-1)u_j v_i / 2 + k - 1}{u_j} - \frac{(n-1)v_i - s_{jk} + 1}{2} \\ &= \frac{(n-1)v_i}{2} + \frac{k-1}{u_j} - \frac{(n-1)v_i - s_{jk} + 1}{2} \\ &= \frac{k-1}{u_j} + \frac{s_{jk} - 1}{2}. \end{aligned}$$

□

Do zilustrowania Twierdzenia 3.3 pokażemy teraz rozkłady statystyki $\Lambda^{2/n}$ dla trzech szczególnych hipotez.

Przykład. Załóżmy, że w eksperymencie obserwowanych jest 7 cech w 5 punktach czasowych, tzn. $m = 7$, $q = 5$. Przyjmijmy, że interesują nas następujące hipotezy:

$$\begin{aligned} H_0^{(a)} : \Sigma_0 &= \Sigma_{\text{BCT_CT}} & \text{vs} & & H_1^{(a)} : \Sigma_1 &= \Sigma_{\text{BCT}} \\ H_0^{(b)} : \Sigma_0 &= \Sigma_{\text{BCS_CS}} & \text{vs} & & H_1^{(b)} : \Sigma_1 &= \Sigma_{\text{BCS}} \\ H_0^{(c)} : \Sigma_0 &= \mathbf{I}_q \otimes \Sigma_{\text{CT}} & \text{vs} & & H_1^{(c)} : \Sigma_1 &= \mathbf{I}_q \otimes \Omega_{\text{UN}}. \end{aligned} \tag{3.12}$$

Dla kolejnych struktur kowariancyjnych wymienionych hipotez zerowych mamy odpowiednio:

- dla hipotezy (a): $\nu = 3, \omega = 4$,
- dla hipotezy (b): $\nu = 2, \omega = 2$,
- dla hipotezy (c): $\nu = 1, \omega = 4$.

Ponadto, dla każdej hipotezy $u_1 = 1$, zatem parametry rozkładu statystyki ilorazu wiarygodności przedstawione w Tabeli 1, zgodnie z Twierdzeniem 3.3, rozpoczynają się będą od $j = 2$.

3.2.2 Test wynikowy Rao

W tej części pracy sformułujemy postać statystyki testowej Rao. Podobnie jak w przypadku statystyki ilorazu wiarygodności niezbędne jest użycie estymatorów największej wiarygodności macierzy kowariancji przy prawdziwości hipotezy zerowej i alternatywnej zdefiniowanej w (3.1). Jednakże ze względu na możliwość prostszego wyprowadzenia statystyki testowej, posłużymy się inną niż w Twierdzeniu 3.1 postacią odpowiednich estymatorów, podaną w następującym twierdzeniu.

Twierdzenie 3.4. *Estymator największej wiarygodności macierzy Σ dla modelu (1.1) przy prawdziwości hipotezy zerowej i alternatywnej w (3.1) przyjmuje odpowiednio postać*

$$\begin{aligned}\widehat{\Sigma}_0 &= \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \widehat{\lambda}_{ij} (\mathbf{V}_i \otimes \mathbf{U}_j), & \text{gdzie} & \quad \widehat{\lambda}_{ij} = \frac{1}{v_i \cdot u_j} \text{Tr} [\mathbf{S} (\mathbf{V}_i \otimes \mathbf{U}_j)], \\ \widehat{\Sigma}_1 &= \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Delta}_i, & \text{gdzie} & \quad \widehat{\Delta}_i = \frac{1}{v_i} \text{BTr}_m [(\mathbf{V}_i \otimes \mathbf{I}_m) \mathbf{S}],\end{aligned}$$

oraz $i = 1, \dots, \nu, j = 1, \dots, \omega$.

Dowód. Niech prawdziwa będzie hipoteza H_0 . Z uwagi na fakt, że $\mathcal{V} \boxtimes \mathcal{U}$ jest komutatywną podprzestrzenią kwadratową, macierz Σ_0 może być zapisana jak w hipotezie

Hipoteza (a): $v_1 = 1, v_2 = v_3 = 2;$
 $u_1 = 1, u_2 = u_3 = u_4 = 2$

j	k	$s_{jk} - 1$	parametry rozkładu beta	
2	1	1	$[(n-1)v_i - 1]/2$	1/2
	2	2	$[(n-1)v_i - 2]/2$	3/2
3	1	3	$[(n-1)v_i - 3]/2$	3/2
	2	4	$[(n-1)v_i - 4]/2$	5/2
4	1	5	$[(n-1)v_i - 5]/2$	5/2
	2	6	$[(n-1)v_i - 6]/2$	7/2

Hipoteza (b): $v_1 = 1, v_2 = 4; \quad u_1 = 1, u_2 = 6$

j	k	$s_{jk} - 1$	parametry rozkładu beta	
2	1	1	$[(n-1)v_i - 1]/2$	1/2
	2	2	$[(n-1)v_i - 2]/2$	3/2
	3	3	$[(n-1)v_i - 3]/2$	5/2
	4	4	$[(n-1)v_i - 4]/2$	7/2
	5	5	$[(n-1)v_i - 5]/2$	9/2
	6	6	$[(n-1)v_i - 6]/2$	11/2

Hipoteza (c): $v_1 = q; \quad u_1 = 1, u_2 = u_3 = u_4 = 2$

j	k	$s_{jk} - 1$	parametry rozkładu beta	
2	1	1	$[(n-1)q - 1]/2$	1/2
	2	2	$[(n-1)q - 2]/2$	3/2
3	1	3	$[(n-1)q - 3]/2$	3/2
	2	4	$[(n-1)q - 4]/2$	5/2
4	1	5	$[(n-1)q - 5]/2$	5/2
	2	6	$[(n-1)q - 6]/2$	7/2

Tabela 1: Parametry rozkładu beta dla rozważanych hipotez zerowych zdefiniowanych w (3.12)

zerowej (3.1). Wykorzystując Twierdzenie 2.2, estymator największej wiarygodności macierzy Σ_0 można uzyskać poprzez rzut \mathbf{S} na przestrzeń $\mathcal{V} \boxtimes \mathcal{U}$. Wówczas,

$$\widehat{\Sigma}_0 = \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\langle \mathbf{S}, \left(\frac{1}{\sqrt{v_i}} \mathbf{V}_i \otimes \frac{1}{\sqrt{u_j}} \mathbf{U}_j \right) \right\rangle \left(\frac{1}{\sqrt{v_i}} \mathbf{V}_i \otimes \frac{1}{\sqrt{u_j}} \mathbf{U}_j \right) = \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \widehat{\lambda}_{ij} (\mathbf{V}_i \otimes \mathbf{U}_j),$$

gdzie dla $i = 1, \dots, \nu, j = 1, \dots, \omega$

$$\widehat{\lambda}_{ij} = \frac{1}{v_i \cdot u_j} \text{Tr}(\mathbf{S}(\mathbf{V}_i \otimes \mathbf{U}_j)).$$

Założmy teraz, że H_1 jest hipotezą prawdziwą. Na podstawie dowodu Twierdzenia 3.1, przekształcając postać estymatora Δ_i podanego w (3.7) mamy, że

$$\widehat{\Delta}_i = \frac{1}{v_i} \mathbf{S}_i = \frac{1}{v_i} \text{BTr}_m [(\mathbf{H}'_i \otimes \mathbf{I}_m) \mathbf{S} (\mathbf{H}_i \otimes \mathbf{I}_m)] = \frac{1}{v_i} \text{BTr}_m [(\mathbf{V}_i \otimes \mathbf{I}_m) \mathbf{S}].$$

□

Wyznaczona zostanie teraz statystyka testu wynikowego Rao dla hipotezy (3.1). Mając na uwadze Definicję 1.3, wektor parametrów przy prawdziwości hipotezy alternatywnej z (3.1) ma postać $\boldsymbol{\theta} = (\boldsymbol{\mu}', \text{vech}' \Delta_1, \text{vech}' \Delta_2, \dots, \text{vech}' \Delta_\nu)'$, podczas gdy przy prawdziwości hipotezy zerowej $\boldsymbol{\theta}_0 = (\boldsymbol{\mu}', \lambda_{11}, \lambda_{12}, \dots, \lambda_{1\omega}, \dots, \lambda_{\nu 2}, \dots, \lambda_{\nu \omega})'$; por. (3.4). Oznaczmy $\widehat{\boldsymbol{\theta}}_0 = \text{MLE}(\boldsymbol{\theta}_0)$. Ponieważ w hipotezie (3.1) nie nakładamy żadnych restrykcji na $\boldsymbol{\mu}$, dlatego też pierwsza składowa wektora wynikowego, $\mathbf{s}(\widehat{\boldsymbol{\theta}}_0)$, jest równa 0. Powyższe spostrzeżenie, bez utraty ogólności, pozwala zredukować wymiar wektora \mathbf{s} o mq . Co więcej, z tego samego powodu elementy poza główną przekątną w pierwszym wierszu i kolumnie macierzy informacji Fishera również zredukują się do 0. Z uwagi na ten fakt, $\boldsymbol{\theta}$ oraz $\boldsymbol{\theta}_0$ są rozważane bez pierwszego składnika, a macierz informacji Fishera jest stopnia $\nu\omega$ zamiast $\nu\omega + mq$.

Twierdzenie 3.5. *Dla testu hipotezy (3.1) statystyka testowa Rao ma postać*

$$\text{RST} = \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - \widehat{\Sigma}_1 \widehat{\Sigma}_0^{-1} \right]^2 \right\},$$

gdzie $\widehat{\Sigma}_0$ i $\widehat{\Sigma}_1$ są estymatorami największej wiarygodności macierzy kowariancji przy prawdziwości, odpowiednio, H_0 i H_1 , których postać przedstawiona została w Twierdzeniu 3.4.

Dowód. Przy prawdziwości hipotezy H_1 logarytm funkcji wiarygodności jest postaci

$$\ln L(\boldsymbol{\mu}, \boldsymbol{\Sigma}_1) = -\frac{nmq}{2} \ln(2\pi) - \frac{n}{2} \sum_{i=1}^{\nu} (v_i \ln |\boldsymbol{\Delta}_i|) - \frac{1}{2} \sum_{i=1}^{\nu} \text{Tr} [\mathbf{Z}'\mathbf{Z}(\mathbf{V}_i \otimes \boldsymbol{\Delta}_i^{-1})], \quad (3.13)$$

gdzie $\mathbf{Z} = \mathbf{Z}(\boldsymbol{\mu}) = \mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu}' \sim N(\mathbf{0}, \mathbf{I}_n, \boldsymbol{\Sigma}_1)$.

Aby wyznaczyć wektor wynikowy, należy zróżniczkować (3.13) ze względu na składowe wektora $\boldsymbol{\theta}$. Stosując regułę łańcuchową opisaną w Magnus, Neudecker (1986), wzory na pochodne z Lematu 1.3, a także Własność 1.2(i), dla każdego $i = 1, \dots, \nu$ mamy, że

$$\begin{aligned} \frac{\partial \ln L}{\partial \boldsymbol{\Delta}_i} &= \left\{ -\frac{nv_i}{2} \text{vec}' \boldsymbol{\Delta}_i^{-1} \right. \\ &\quad \left. + \frac{1}{2} \text{vec}' (\mathbf{Z}'\mathbf{Z})(\mathbf{I}_q \otimes \mathbf{K}_{m,q} \otimes \mathbf{I}_m) (\text{vec } \mathbf{V}_i \otimes \mathbf{I}_{m^2}) (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \right\} \mathbf{D}_m \\ &= \frac{n}{2} \left(-v_i \text{vec}' \boldsymbol{\Delta}_i^{-1} + \text{vec}' \left\{ \text{BTr}_m[(\mathbf{V}_i \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}'\mathbf{Z}] \right\} (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \right) \mathbf{D}_m. \end{aligned}$$

Mając na uwadze, że

$$\frac{1}{n} \mathbf{Z}'\mathbf{Z} = \frac{1}{n} \mathbf{Z}'(\hat{\boldsymbol{\mu}}) \mathbf{Z}(\hat{\boldsymbol{\mu}}) = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y} = \mathbf{S},$$

wstawiając do wektora wynikowego $\mathbf{s}(\boldsymbol{\theta}) = \left(\frac{\partial \ln L}{\partial \boldsymbol{\Delta}_i} \right)'_{i=1, \dots, \nu}$ estymator największej wiarygodności $\boldsymbol{\theta}_0$, otrzymujemy

$$\mathbf{s}(\hat{\boldsymbol{\theta}}_0) = -\frac{n}{2} (\mathbf{I}_\nu \otimes \mathbf{D}'_m) \hat{\mathbf{S}},$$

gdzie

$$\hat{\mathbf{S}} = \left(v_i \text{vec } \hat{\boldsymbol{\Lambda}}_i^{-1} - \left(\hat{\boldsymbol{\Lambda}}_i^{-1} \otimes \hat{\boldsymbol{\Lambda}}_i^{-1} \right) \text{vec} \{ \text{BTr}_m[(\mathbf{V}_i \otimes \mathbf{I}_m) \mathbf{S}] \} \right)'_{i=1, \dots, \nu}$$

oraz

$$\hat{\boldsymbol{\Lambda}}_i = \sum_{j=1}^{\omega} \hat{\delta}_{ij} \mathbf{U}_j.$$

Uwzględniając postać $\hat{\boldsymbol{\Delta}}_i$ z Twierdzenia (3.4), każdą składową wektora wynikowego, $\hat{\mathbf{s}}_i$, $i = 1, \dots, \nu$, można przedstawić jako

$$\hat{\mathbf{s}}_i = v_i \left(\left(\hat{\boldsymbol{\Lambda}}_i^{-1} \otimes \hat{\boldsymbol{\Lambda}}_i^{-1} \right) \text{vec} \left(\hat{\boldsymbol{\Lambda}}_i - \hat{\boldsymbol{\Delta}}_i \right) \right)'_{i=1, \dots, \nu}. \quad (3.14)$$

Do wyznaczenia macierzy informacji Fishera skorzystamy ze wzoru

$$\mathbf{F}_{ij}(\boldsymbol{\theta}) = -\mathbb{E}\left(\frac{\partial^2 \ln L}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j}\right),$$

dla $i, j = 1, \dots, \nu$. Ze względu na fakt, że dla $i \neq j, i, j = 1, \dots, \nu$, pochodna $\frac{\partial \ln L}{\partial \boldsymbol{\Delta}_i}$ nie zależy od $\boldsymbol{\Delta}_j$, bloki pozadiagonalne macierzy informacji Fishera są równe zero. Różniczkując dwukrotnie funkcję $\ln L$ ze względu na i -tą składową wektora $\boldsymbol{\theta}$, dla $i = 1, \dots, \nu$, otrzymujemy

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_i^2} &= \mathbf{D}'_m \left(\frac{nv_i}{2} (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \right. \\ &\quad \left. - \frac{n}{2} \left[(\text{vec}'\{\text{BTr}_m[(\mathbf{V}_i \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}\mathbf{Z}']\} \otimes \mathbf{I}_{m^2}) \cdot (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) \right. \right. \\ &\quad \left. \left. \cdot (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \text{vec } \boldsymbol{\Delta}_i^{-1} + \text{vec } \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \right] \right) \mathbf{D}_m. \end{aligned}$$

Korzystając z Twierdzenia 2.2.9 (i) z Kollo, von Rosen (2005) wiadomo, że $\mathbb{E}(\mathbf{Z}'\mathbf{Z}) = \text{Tr } \mathbf{I}_n \cdot \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \boldsymbol{\Delta}_i = n \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \boldsymbol{\Delta}_i$. Mając na uwadze ortogonalność oraz idempotentność macierzy $\mathbf{V}_i, i = 1, 2, \dots, \nu$, otrzymujemy

$$\begin{aligned} \mathbf{F}_{ii}(\boldsymbol{\theta}) &= -\mathbb{E}\left(\frac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_i^2}\right) \\ &= -\frac{n}{2} \mathbf{D}'_m \left[v_i (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \right. \\ &\quad \left. - \{\text{vec}'[\text{BTr}_m(\mathbf{V}_i \otimes \boldsymbol{\Delta}_i)] \otimes \mathbf{I}_{m^2}\} (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) \right. \\ &\quad \left. (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \text{vec } \boldsymbol{\Delta}_i^{-1} + \text{vec } \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \right] \mathbf{D}_m. \\ &= \frac{nv_i}{2} \left[-\mathbf{D}'_m (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \mathbf{D}_m \right. \\ &\quad \left. + (\text{vec}' \boldsymbol{\Delta}_i \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \text{vec } \boldsymbol{\Delta}_i^{-1}) \mathbf{D}_m \right. \\ &\quad \left. + (\text{vec}' \boldsymbol{\Delta}_i \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\text{vec } \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \mathbf{D}_m \right] \\ &= \frac{nv_i}{2} \left[-\mathbf{D}'_m (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \mathbf{D}_m \right. \\ &\quad \left. + (\text{vec}' \boldsymbol{\Delta}_i \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\mathbf{I}_{m^2} \otimes \text{vec } \boldsymbol{\Delta}_i^{-1}) \right. \\ &\quad \left. \cdot (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \mathbf{D}_m \right. \\ &\quad \left. + (\text{vec}' \boldsymbol{\Delta}_i \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\text{vec } \boldsymbol{\Delta}_i^{-1} \otimes \mathbf{I}_{m^2}) \right. \\ &\quad \left. \cdot (\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \mathbf{D}_m \right], \end{aligned}$$

ponieważ $\text{BTr}_m(\mathbf{V}_i \otimes \mathbf{\Delta}_i) = v_i \mathbf{\Delta}_i$ (por. Filipiak i in., 2018a, Lemat 2.11, Lemat 2.13).
Dodatkowo, z Lematu 1.1, dla każdego $i = 1, 2, \dots, \nu$,

$$\begin{aligned} \mathbf{F}_{ii}(\boldsymbol{\theta}) &= -\frac{nv_i}{2} \mathbf{D}'_m (\mathbf{\Delta}_i^{-1} \otimes \mathbf{\Delta}_i^{-1}) \mathbf{D}_m + \frac{nv_i}{2} \mathbf{D}'_m (\mathbf{\Delta}_i^{-1} \otimes \mathbf{\Delta}_i^{-1}) \mathbf{D}_m \\ &\quad + \frac{nv_i}{2} \mathbf{D}'_m (\mathbf{\Delta}_i^{-1} \otimes \mathbf{\Delta}_i^{-1}) \mathbf{D}_m \\ &= \frac{nv_i}{2} \mathbf{D}'_m (\mathbf{\Delta}_i^{-1} \otimes \mathbf{\Delta}_i^{-1}) \mathbf{D}_m. \end{aligned}$$

Podstawiając za $\boldsymbol{\theta}$ estymatory największej wiarygodności nieznanych parametrów przy prawdziwości hipotezy zerowej, $\hat{\boldsymbol{\theta}}_0$, otrzymujemy

$$\mathbf{F}(\hat{\boldsymbol{\theta}}_0) = \frac{n}{2} \text{BDiag} \left(v_1 \hat{\mathbf{F}}_{11}, v_2 \hat{\mathbf{F}}_{22}, \dots, v_\nu \hat{\mathbf{F}}_{\nu\nu} \right),$$

gdzie $\hat{\mathbf{F}}_{ii} = \mathbf{D}'_m (\hat{\mathbf{\Lambda}}_i^{-1} \otimes \hat{\mathbf{\Lambda}}_i^{-1}) \mathbf{D}_m$, $i = 1, 2, \dots, \nu$. Korzystając z (1.19) otrzymujemy

$$\mathbf{F}^{-1}(\hat{\boldsymbol{\theta}}_0) = \frac{2}{n} \text{BDiag} \left(\frac{1}{v_1} \hat{\mathbf{F}}_{11}^{-1}, \dots, \frac{1}{v_\nu} \hat{\mathbf{F}}_{\nu\nu}^{-1} \right),$$

gdzie $\hat{\mathbf{F}}_{ii}^{-1} = \mathbf{D}_m^+ (\hat{\mathbf{\Lambda}}_i \otimes \hat{\mathbf{\Lambda}}_i) \mathbf{D}_m^{+'}$, lub równoważnie

$$\mathbf{F}^{-1}(\hat{\boldsymbol{\theta}}) = \frac{2}{n} (\mathbf{I}_\nu \otimes \mathbf{D}_m^+) \text{BDiag} \left(\frac{1}{v_1} \hat{\mathbf{\Lambda}}_1 \otimes \hat{\mathbf{\Lambda}}_1, \dots, \frac{1}{v_\nu} \hat{\mathbf{\Lambda}}_\nu \otimes \hat{\mathbf{\Lambda}}_\nu \right) (\mathbf{I}_\nu \otimes \mathbf{D}_m^{+'}). \quad (3.15)$$

Wstawiając do wzoru (1.14) wyrażenia (3.14) oraz (3.15), mamy, że

$$\begin{aligned} \text{RST} &= \frac{n}{2} \mathbf{s}' (\mathbf{I}_\nu \otimes \mathbf{D}_m \mathbf{D}_m^+) \text{BDiag} \left(\frac{1}{v_1} \hat{\mathbf{\Lambda}}_1 \otimes \hat{\mathbf{\Lambda}}_1, \dots, \frac{1}{v_\nu} \hat{\mathbf{\Lambda}}_\nu \otimes \hat{\mathbf{\Lambda}}_\nu \right) (\mathbf{I}_\nu \otimes \mathbf{D}_m^{+'} \mathbf{D}_m') \\ &= \frac{n}{2} \left[\sum_{i=1}^{\nu} v_i \text{vec}' (\hat{\mathbf{\Lambda}}_i - \widehat{\mathbf{\Delta}}_i) (\hat{\mathbf{\Lambda}}_i^{-1} \otimes \hat{\mathbf{\Lambda}}_i^{-1}) \mathbf{D}_m \mathbf{D}_m^+ (\hat{\mathbf{\Lambda}}_i \otimes \hat{\mathbf{\Lambda}}_i) \mathbf{D}_m^{+'} \mathbf{D}_m' \right. \\ &\quad \left. \cdot (\hat{\mathbf{\Lambda}}_i^{-1} \otimes \hat{\mathbf{\Lambda}}_i^{-1}) \text{vec} (\hat{\mathbf{\Lambda}}_i - \widehat{\mathbf{\Delta}}_i) \right]. \end{aligned}$$

Wykorzystując Definicje 1.1 oraz 1.3, a także równania (1.17) i (1.18) otrzymujemy

$$\begin{aligned}
\text{RST} &= \frac{n}{2} \left[\sum_{i=1}^{\nu} v_i \text{vec}'(\widehat{\Lambda}_i - \widehat{\Delta}_i)(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1}) \mathbf{N}_m(\widehat{\Lambda}_i \otimes \widehat{\Lambda}_i) \mathbf{N}_m(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1}) \right. \\
&\quad \left. \cdot \text{vec}(\widehat{\Lambda}_i - \widehat{\Delta}_i) \right] \\
&= \frac{n}{8} \left[\sum_{i=1}^{\nu} v_i \text{vec}'(\widehat{\Lambda}_i - \widehat{\Delta}_i)(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1})(\mathbf{I}_{m^2} + \mathbf{K}_{m,m})(\widehat{\Lambda}_i \otimes \widehat{\Lambda}_i) \right. \\
&\quad \left. \cdot (\mathbf{I}_{m^2} + \mathbf{K}_{m,m})(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1}) \text{vec}(\widehat{\Lambda}_i - \widehat{\Delta}_i) \right] \\
&= \frac{n}{2} \left[\sum_{i=1}^{\nu} v_i \text{vec}'(\widehat{\Lambda}_i - \widehat{\Delta}_i)(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1})(\widehat{\Lambda}_i \otimes \widehat{\Lambda}_i)(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1}) \right. \\
&\quad \left. \cdot \text{vec}(\widehat{\Lambda}_i - \widehat{\Delta}_i) \right] \\
&= \frac{n}{2} \left[\sum_{i=1}^{\nu} v_i \text{vec}'(\widehat{\Lambda}_i - \widehat{\Delta}_i)(\widehat{\Lambda}_i^{-1} \otimes \widehat{\Lambda}_i^{-1}) \text{vec}(\widehat{\Lambda}_i - \widehat{\Delta}_i) \right].
\end{aligned}$$

Z Własności 1.1(ii) wynika dalej, że

$$\begin{aligned}
\text{RST} &= \frac{n}{2} \left\{ \sum_{i=1}^{\nu} v_i \text{vec}'(\widehat{\Lambda}_i - \widehat{\Delta}_i) \text{vec}[\widehat{\Lambda}_i^{-1}(\widehat{\Lambda}_i - \widehat{\Delta}_i)\widehat{\Lambda}_i^{-1}] \right\} \\
&= \frac{n}{2} \left\{ \sum_{i=1}^{\nu} v_i \text{Tr} \left[(\widehat{\Lambda}_i - \widehat{\Delta}_i)\widehat{\Lambda}_i^{-1}(\widehat{\Lambda}_i - \widehat{\Delta}_i)\widehat{\Lambda}_i^{-1} \right] \right\} \\
&= \frac{n}{2} \left[\sum_{i=1}^{\nu} v_i \text{Tr}(\mathbf{I}_m - 2\widehat{\Delta}_i\widehat{\Lambda}_i^{-1} + \widehat{\Delta}_i\widehat{\Lambda}_i^{-1}\widehat{\Delta}_i\widehat{\Lambda}_i^{-1}) \right].
\end{aligned}$$

Z idempotentności i ortogonalności macierzy bazowych \mathbf{V}_i , $i = 1, \dots, \nu$, równości $v_i \text{Tr} \widehat{\Delta}_i\widehat{\Lambda}_i^{-1} = \text{Tr}(\mathbf{V}_i \otimes \widehat{\Delta}_i\widehat{\Lambda}_i^{-1})$, otrzymujemy ostatecznie

$$\begin{aligned}
\text{RST} &= \frac{n}{2} \left\{ \sum_{i=1}^{\nu} \left[\text{Tr}(\mathbf{V}_i \otimes \mathbf{I}_m) - 2 \text{Tr}(\mathbf{V}_i \otimes \widehat{\Delta}_1\widehat{\Lambda}_i^{-1}) + \text{Tr}(\mathbf{V}_i \otimes \widehat{\Delta}_1\widehat{\Lambda}_i^{-1}\widehat{\Delta}_1\widehat{\Lambda}_i^{-1}) \right] \right\} \\
&= \frac{n}{2} \left\{ \text{Tr} \mathbf{I}_{mq} - \sum_{i=1}^{\nu} \text{Tr} \left[2(\mathbf{V}_i \otimes \widehat{\Delta}_1)(\mathbf{I}_q \otimes \widehat{\Lambda}_i^{-1}) + (\mathbf{V}_i \otimes \widehat{\Delta}_1)(\mathbf{I}_q \otimes \widehat{\Lambda}_i^{-1}) \right. \right. \\
&\quad \left. \left. \cdot (\mathbf{V}_i \otimes \widehat{\Delta}_1)(\mathbf{I}_q \otimes \widehat{\Lambda}_i^{-1}) \right] \right\} \\
&= \frac{n}{2} \text{Tr} \left[\mathbf{I}_{mq} - 2 \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Delta}_i \right) \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Lambda}_i^{-1} \right) \right. \\
&\quad \left. + \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Delta}_i \right) \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Lambda}_i^{-1} \right) \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Delta}_i \right) \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \widehat{\Lambda}_i^{-1} \right) \right] \\
&= \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - \widehat{\Sigma}_1\widehat{\Sigma}_0^{-1} \right]^2 \right\}.
\end{aligned}$$

□

3.3 Porównanie LRT i RST – badania symulacyjne

W celu porównania proponowanych testów przeprowadzimy badania symulacyjne. Przedmiotem zainteresowania w ogólności jest hipoteza (3.1), jednakże w badaniach symulacyjnych rozważymy hipotezę (3.12)(a) jako dotyczącą jednej z najszerszych klas zdefiniowanych struktur kowariancyjnych przedstawianych w niniejszej pracy. Aby można było przeprowadzić test ilorazu wiarygodności, w symulacjach zakładamy, że $n > m$.

Na początku porównamy zbieżność rozkładu dokładnego LRT oraz rozkładu empirycznego RST przy prawdziwości hipotezy zerowej do granicznego rozkładu $\chi^2_{(m(m+1)/2-\omega)\nu}$. Dalej, dokonamy analizy mocy wspomnianych testów, gdzie dane generowane będą z rozkładu $N_{n,mq}(\mathbf{0}, \mathbf{I}_n, \Sigma_{\text{BCT}})$.

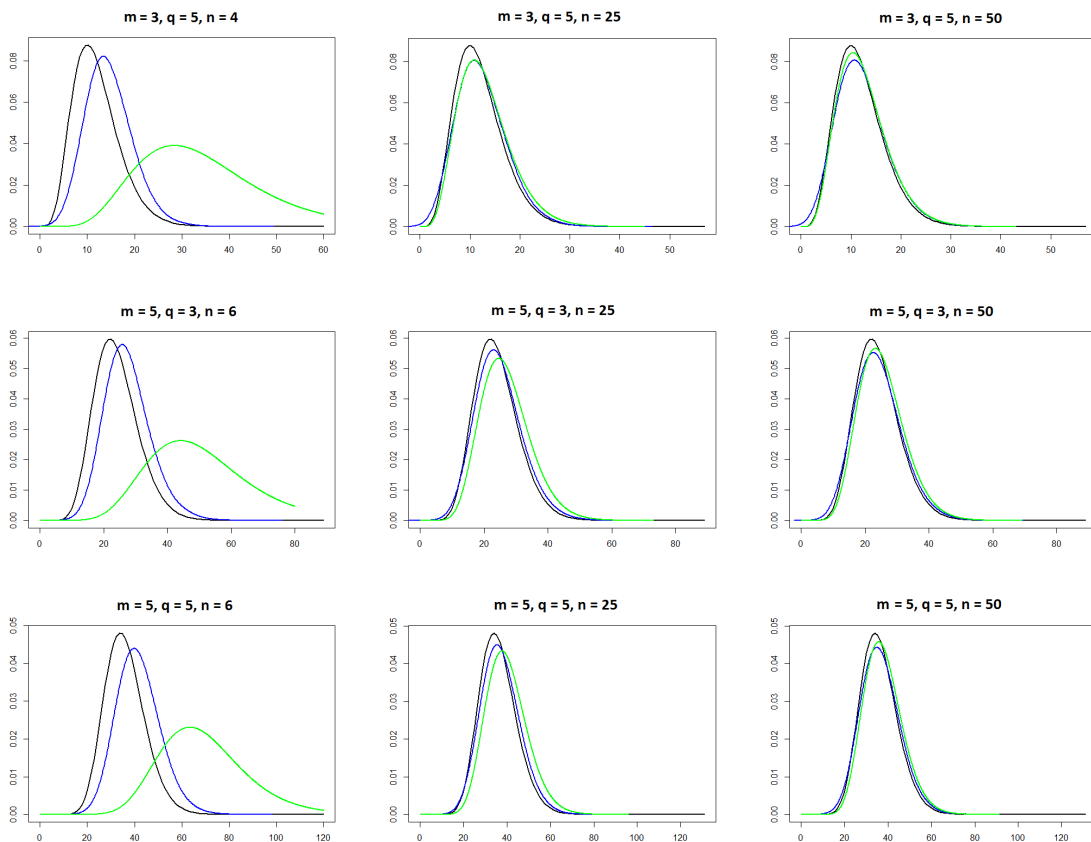
Kolejne sekcje niniejszego rozdziału rozprawy pokazują, że RST przewyższa LRT pod względem szybkości zbieżności rozkładu przy prawdziwości hipotezy zerowej do granicznego rozkładu $\chi^2_{(m(m+1)/2-\omega)\nu}$, podczas gdy nie ma istotnej różnicy w zachowaniu mocy testów: oba są konkurencyjne.

3.3.1 Zbieżność do rozkładu granicznego

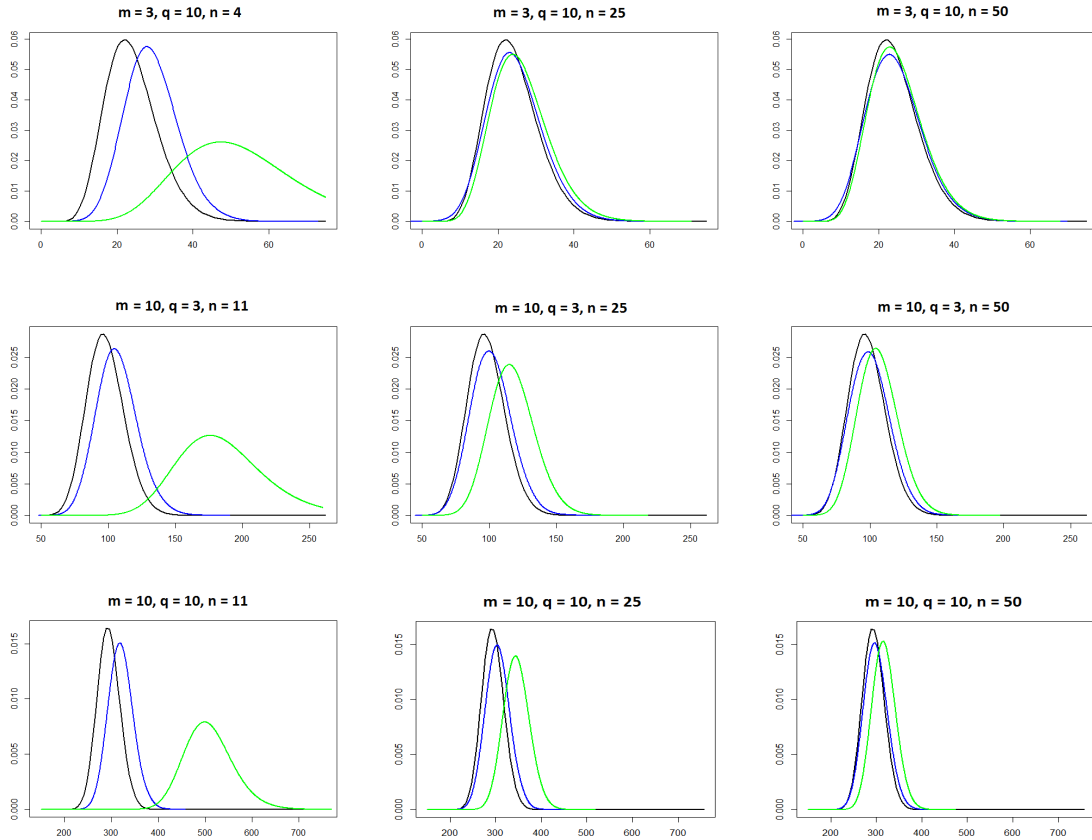
W celu wyznaczenia rozkładu dokładnego statystyki LRT (w szczególności kwantyli) przy prawdziwości hipotezy zerowej został wykorzystany pakiet R `CharFunToolR` opracowany przez Gajdoš (2018). W przypadku RST posłużyliśmy się rozkładem empirycznym wyznaczonym jako rozkład 10 000 statystyk RST obliczonych na podstawie danych wygenerowanych z rozkładu $N_{n,mq}(\mathbf{0}, \mathbf{I}_n, \Sigma_{\text{BCT_CT}})$, gdzie parametry modelu zostały wybrane jako $(m, q) \in \{(3, 5), (5, 3), (5, 5), (3, 10), (10, 3), (10, 10)\}$ oraz $n \in \{m + 1, 25, 50\}$.

Rozkłady statystyk LRT i RST wraz z rozkładem granicznym dla wyżej podanych kombinacji parametrów m i q przedstawiono na Rysunkach 3, 4. Zgodnie z teorią podaną przez Rao można zauważyć, że wraz ze wzrostem n oba rozkłady dążą do

rozkładu granicznego, jednak zbieżność RST jest szybsza niż LRT we wszystkich rozpatrywanych przypadkach. Co więcej, dla relatywnie małych parametrów m i q , empiryczny rozkład RST jest zbliżony do rozkładu granicznego nawet dla małej próby. Należy jednak zauważyć, że wzrost m i q (zwłaszcza m) zmniejsza szybkość zbieżności obu rozkładów.



Rysunek 3: Rozkład dokładny LRT (zielony) oraz rozkład empiryczny RST (niebieski) przy prawdziwości hipotezy zerowej wraz z rozkładem χ^2_ν (czarny) dla $m = 3, q = 5, \nu = 12$ (pierwszy wiersz), $m = 5, q = 3, \nu = 24$ (drugi wiersz), $m = 5, q = 5, \nu = 36$ (trzeci wiersz).



Rysunek 4: Rozkład dokładny LRT (zielony) oraz rozkład empiryczny RST (niebieski) przy prawdziwości hipotezy zerowej wraz z rozkładem χ^2_ν (czarny) dla $m = 3, q = 10, \nu = 24$ (pierwszy wiersz), $m = 10, q = 3, \nu = 98$ (drugi wiersz), $m = 10, q = 10, \nu = 294$ (trzeci wiersz).

Z uwagi na to, iż dla małych wielkości próby kwantyle rozkładu LRT różnią się istotnie od kwantyli rozkładu granicznego, do testowania hipotez dotyczących danych rzeczywistych sugerowalibyśmy wykorzystanie rozkładu dokładnego, ponieważ rozkład przybliżony może prowadzić do podjęcia błędnej decyzji (zob. Rozdział 3.4).

3.3.2 Analiza mocy testów

Aby dokonać analizy mocy testów LRT i RST w zależności od rozbieżności pomiędzy hipotezą zerową a alternatywną, podanych w (3.1), jak również pod względem wielkości próby, wykorzystamy rozbieżność Kullbacka-Leiblera zdefiniowaną w (1.16).

Niech $\lambda = \{\lambda_{ij} : i = 1, \dots, \nu, j = 1, \dots, \omega\}$ oraz $\mathcal{D} = \{\Delta_i : i = 1, \dots, \nu\}$. Wówczas funkcja rozbieżności dla hipotezy (3.1) ma postać

$$f_E(\lambda, \mathcal{D}) = \text{Tr} \left[\left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \Delta_i^{-1} \right) \left(\sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \lambda_{ij} (\mathbf{V}_i \otimes \mathbf{U}_j) \right) \right] - \ln \left| \left(\sum_{i=1}^{\nu} \mathbf{V}_i \otimes \Delta_i^{-1} \right) \left(\sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \lambda_{ij} (\mathbf{V}_i \otimes \mathbf{U}_j) \right) \right| - mq.$$

Zatem dla zadanej hipotezy alternatywnej (zbiór \mathcal{D}), aby wyznaczyć strukturę Σ_0 o najmniejszej rozbieżności (zbiór $\tilde{\lambda}$) należy wyznaczyć następujące minimum

$$\xi(\tilde{\lambda}) = \min_{\lambda_{ij}} f_E(\lambda | \mathcal{D}),$$

gdzie $i = 1, \dots, \nu, j = 1, \dots, \omega$. Pamiętając, że macierze $\mathbf{V}_i, i = 1, \dots, \nu$, są idempotentne i ortogonalne oraz uwzględniając (3.5), problem sprowadza się do wyznaczenia

$$\xi(\tilde{\lambda}) = \min_{\lambda_{ij}} \left\{ \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} v_i \lambda_{ij} \text{Tr}(\Delta_i^{-1} \mathbf{U}_j) + \ln \left| \sum_{i=1}^{\nu} \mathbf{V}_i \otimes \Delta_i \right| - \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} v_i u_j \ln(\lambda_{ij}) - mq \right\}.$$

Różniczkując powyższe ze względu na λ_{ij} oraz przyrównując pochodną do zera mamy, że

$$\tilde{\lambda} = \left\{ \tilde{\lambda}_{ij} = \frac{u_j}{\text{Tr}(\Delta_i^{-1} \mathbf{U}_j)} : i = 1, \dots, \nu, j = 1, \dots, \omega \right\}.$$

Ponieważ funkcja rozbieżności Kullbacka-Leiblera nie jest ograniczona z góry, do analizy mocy testów została wykorzystana skorygowana rozbieżność $\eta(\tilde{\lambda}) = 1 - 1/(1 - \xi(\tilde{\lambda}))$; por. Rozdział 1.3.

W pierwszym kroku sprawdzimy zachowanie mocy testów LRT i RST w zależności od skorygowanej rozbieżności η . Dla $(m, q) \in \{(3, 5), (5, 3), (5, 5)\}$ oraz $n = 25$ wygenerujemy 100 określonych dodatnio macierzy Σ_{BCT} , dla których obliczymy wartość

rozbieżności η . Następnie dla każdej z wyznaczonych 100 macierzy wygenerujemy zostało 10 000 macierzy obserwacji z rozkładu $N_{n,mq}(\mathbf{0}, \mathbf{I}_n, \Sigma_{\text{BCT}})$, na podstawie których obliczymy moc empiryczną jako stosunek liczby odrzuconych hipotez zerowych do wszystkich 10 000 powtórzeń. Do odrzucenia H_0 wykorzystamy 95-ty kwantyl rozkładu dokładnego (przy prawdziwości H_0) statystyki LRT oraz rozkładu empirycznego (przy prawdziwości H_0) statystyki RST.

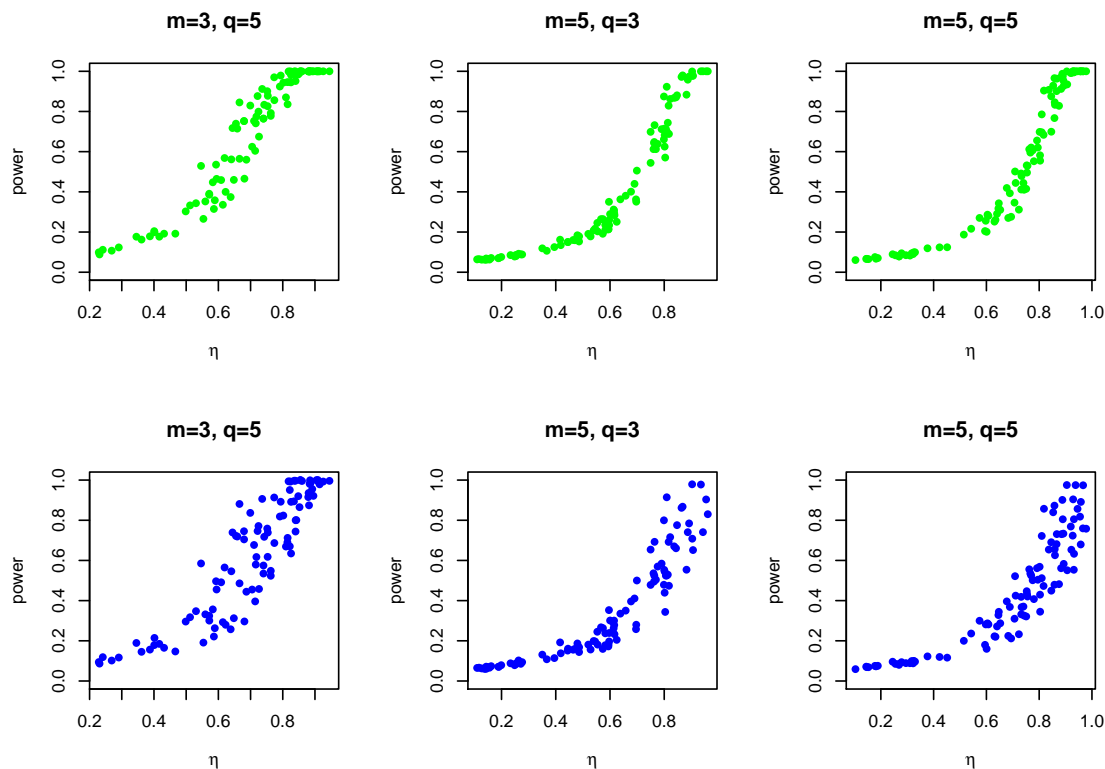
Moc testów LRT i RST w zależności od skorygowanej rozbieżności przedstawiono na Rysunku 5. Na wszystkich wykresach można zaobserwować wzrastający trend mocy testów w zależności od wielkości rozbieżności przy nieco większych odchyleniach między mocami dla RST w porównaniu do LRT. Podobne zjawiska zaobserwowano w pracy Filipiak i in. (2024), gdzie badano wpływ różnych miar rozbieżności na moc testów LRT i RST w kontekście testowania struktury rozdzielnej.

Na koniec porównamy moc testów LRT i RST w zależności od wielkości próby. W tym celu dla $m = 3$, $q = 5$ oraz $n \in \{4, 15, 25, 35, 50, 75, 100\}$ wybierzemy dwie macierze Σ_{BCT} takie, że $\eta = 0.4$ i dwie macierze Σ_{BCT} takie, że $\eta = 0.6$. Każdy wykres na Rysunku 6 reprezentuje moc empiryczną LRT (zielona linia) i RST (niebieska linia) w zależności od wielkości próby dla 10 000 macierzy obserwacji wygenerowanych z rozkładu $N_{n,mq}(\boldsymbol{\mu}, \mathbf{I}_n, \Sigma_{\text{BCT}})$, gdzie w kolumnach przedstawiono wyniki w zależności od wartości skorygowanej rozbieżności (pierwsza kolumna: $\eta = 0.4$, druga kolumna; $\eta = 0.6$). Zgodnie z oczekiwaniami moc obu testów rośnie wraz ze wzrostem wielkości próby oraz, zgodnie z wynikami prezentowanymi na Rysunku 5, moc rośnie wraz z wartością skorygowanej rozbieżności. Można zauważyć, że oba testy są porównywalne: zazwyczaj moce obu testów nie różnią się istotnie, przy czym w niektórych przypadkach moc RST przewyższa moc LRT, a w innych zależność jest odwrotna.

3.4 Przykład rzeczywisty

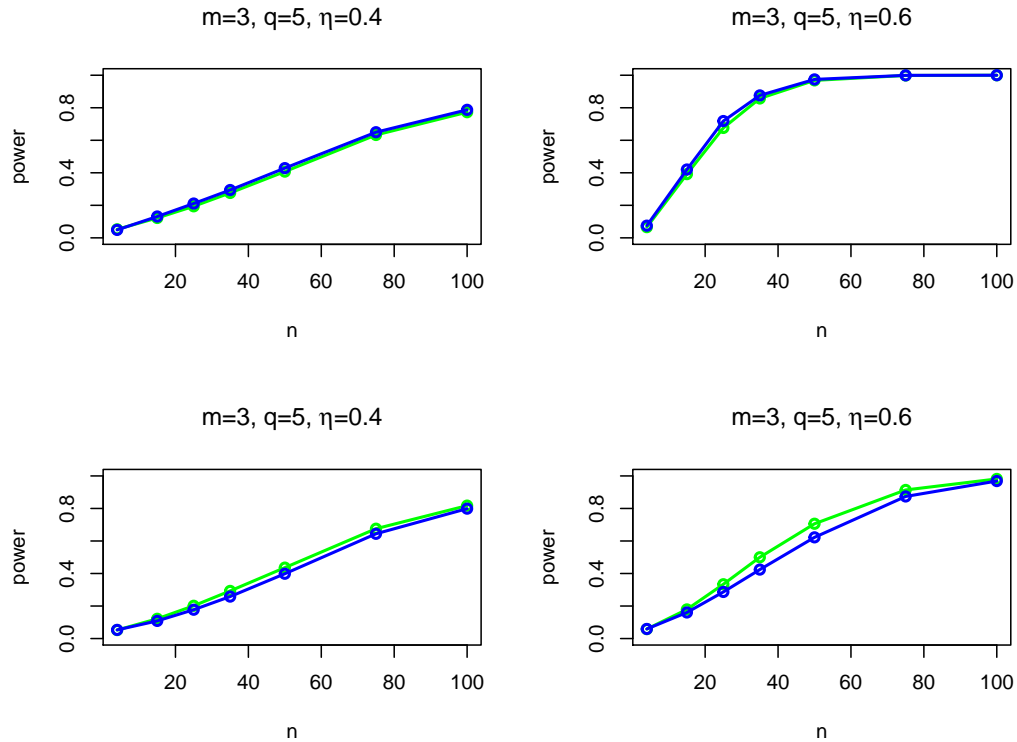
Rozważmy rzeczywiste dane zaprezentowane w Rozdziale 1.5, dla których $n = 11$, $q = 3$ oraz $m = 4$.

Jak zaobserwowano w pracy Liang i in. (2015), układ płatków na każdym kwie-



Rysunek 5: Moc empiryczna LRT (pierwszy wiersz) oraz RST (drugi wiersz) w zależności od wartości skorygowanej rozbieżności η , dla różnych kombinacji m, q , oraz dla $n = 25$.

cie jest kołowy. Ponadto można przyjąć, że korelacja między długością dowolnych dwóch płatków na jednym kwiecie zależy od liczby płatków pomiędzy nimi. Innymi słowy, zakładamy równe korelacje pomiędzy obserwacjami z sąsiednich płatków i równe korelacje pomiędzy obserwacjami z przeciwległych płatków. Dlatego rozsądnym wydaje się przeprowadzenie testu o strukturze kołowej macierzy Toeplitza zależności wewnątrzklasowej między płatkami na pojedynczym kwiecie. Ponadto, ponieważ kolejność kwiatów na pojedynczym kwiatostanie nie jest istotna, do opisu korelacji między kwiatostanami (zależność międzyklasowa) można zastosować również strukturę kompletnej symetrii. Zauważmy jednocześnie, że ponieważ wybrane są tylko



Rysunek 6: Moc dokładna LRT (linia zielona) oraz moc empiryczna RST (niebieska linia) w zależności od wielkości próby dla $m = 3$, $q = 5$ oraz dla dwóch hipotez alternatywnych, dla których $\eta = 0.4$ (pierwsza kolumna) oraz $\eta = 0.6$ (druga kolumna).

3 kwiatostany, struktura kompletnej symetrii jest równoważna strukturze kołowej macierzy Toeplitza.

W Tabeli 2 zostały podane wymiary odpowiednich przestrzeni rozważanych w hipotezie zerowej (3.12) wraz z macierzami bazowymi. Warto przypomnieć, że z uwagi na to, że dla każdej rozważanej hipotezy struktury reprezentujące zależność międzyklasową w hipotezie zerowej i alternatywnej są dokładnie takie same, natomiast w przypadku hipotezy alternatywnej nie zakładamy żadnej zależności wewnątrzklasowej, nie ma potrzeby przedstawiania bazy dla hipotez alternatywnych.

Dla każdej hipotezy, korzystając z Twierdzeń 3.2 oraz 3.5, zostały obliczone staty-

styki LRT i RST oraz wyznaczone p -wartości za pomocą rozkładu dokładnego (przedstawionego w Twierdzeniu 3.3) dla statystyki LRT, rozkładu empirycznego dla statystyki RST oraz granicznego rozkładu chi-kwadrat. Dokładny rozkład LRT został wyznaczony przy użyciu pakietu `CharFunToolR`; por. Gajdoš (2018). Liczba stopni swobody granicznego rozkładu chi-kwadrat wynosi odpowiednio $(m(m+1)/2 - \alpha)\nu = 14$ dla hipotezy (a), $m(m+1) - 4 = 16$ dla hipotezy (b) i $m(m+1)/2 - \omega = 7$ dla hipotezy (c). Wyniki przedstawione zostały w Tabeli 3.

Hipoteza	zależność międzyklasowa			zależność wewnątrzklasowa		
	ν	v_i	\mathbf{V}_i	ω	u_j	\mathbf{U}_j
(a)	2	$v_1 = 1$ $v_2 = 2$	$\mathbf{V}_1 = \mathbf{P}_3$ $\mathbf{V}_2 = \mathbf{Q}_3$	3	$u_1 = 1$ $u_2 = 2$ $u_3 = 1$	$\mathbf{U}_1 = \mathbf{P}_4$ $\mathbf{U}_2 = \ell_2 \ell_2' + \ell_4 \ell_4'$ $\mathbf{U}_3 = \ell_3 \ell_3'$
(b)	2	$v_1 = 1$ $v_2 = 2$	$\mathbf{V}_1 = \mathbf{P}_3$ $\mathbf{V}_2 = \mathbf{Q}_3$	2	$u_1 = 1$ $u_2 = 3$	$\mathbf{U}_1 = \mathbf{P}_4$ $\mathbf{U}_2 = \mathbf{Q}_4$
(c)	1	$v_1 = 3$	$\mathbf{V}_1 = \mathbf{I}_3$	3	$u_1 = 1$ $u_2 = 2$ $u_3 = 1$	$\mathbf{U}_1 = \mathbf{P}_4$ $\mathbf{U}_2 = \ell_2 \ell_2' + \ell_4 \ell_4'$ $\mathbf{U}_3 = \ell_3 \ell_3'$

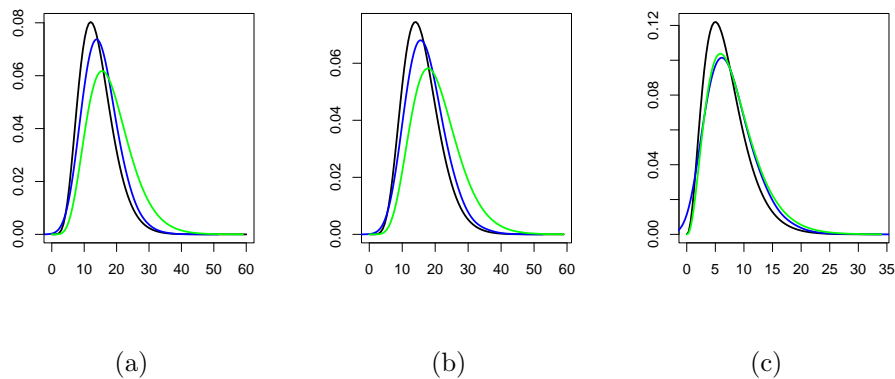
Tabela 2: Wymiary przestrzeni i bazy struktur kowariancyjnych rozważanych w hipotezie zerowej (3.12).

Zakładając poziom istotności $\alpha = 0,05$, przy testowaniu hipotez (a) i (b) decyzje podjęte w oparciu o rozkład dokładny statystyki LRT oraz rozkład graniczny chi-kwadrat są różne (stąd odpowiednie p -wartości rozkładu chi-kwadrat zostały wyróżnione w Tabeli 3 pogrubioną czcionką). Wynika to z faktu, że dla małej próby rozkład dokładny statystyki LRT różni się istotnie od rozkładu granicznego chi-kwadrat; por. Rysunek 7(a), (b). W przypadku hipotezy (c) ze względu na specyfikę rozważanej struktury, rozkłady statystyk LRT i RST nie różnią się istotnie i są relatywnie bliskie rozkładowi granicznemu nawet dla tak małej próby jak $n = 11$; por. Rysunek 7(c).

Zauważmy również, że ze względu na szybką zbieżność rozkładu RST do rozkładu granicznego (por. Rozdział 3.3.1 oraz Rysunek 7), decyzje podjęte na podstawie obu rozkładów są takie same.

	(a)		(b)		(c)	
	LRT	RST	LRT	RST	LRT	RST
statystyki testowe	24.89	20.82	27.56	23.98	12.01	12.36
p -wartość – dokładny rozkład	0.165	0.137	0.163	0.118	0.182	0.122
p -wartość – graniczny rozkład χ^2	0.036	0.106	0.036	0.090	0.100	0.089

Tabela 3: Statystyki testowe wraz z odpowiednimi p -wartościami dla hipotez (3.12).



Rysunek 7: Rozkład dokładny LRT (zielony) i rozkład empiryczny RST (niebieski) wraz z rozkładem χ^2 z odpowiednią liczbą stopni swobody (czarny) dla $m = 4$, $q = 3$, $n = 11$ dla hipotez (3.12).

Jak zaobserwowaliśmy, przy dużej liczebności próby, zarówno rozkład dokładny statystyki LRT jak i rozkład empiryczny statystyki RST są bliskie rozkładowi granicznemu chi-kwadrat, stąd nie ma istotnych różnic w podjętej decyzji. Dla małej liczebności próby rozkład dokładny LRT różni się istotnie od rozkładu chi-kwadrat, zatem wnioskowanie na podstawie wartości statystyki LRT i rozkładu granicznego może prowadzić do podjęcia błędnej decyzji.

4 Testowanie niezależności międzyklasowej w modelu z macierzą kowariancji o blokowej strukturze kompletnej symetrii

Jedną z powszechnych struktur dla modeli podwójnie wielowymiarowych jest struktura BCS, wprowadzona w Rao (1945, 1953), gdzie badano problem dyskryminacji grup odmiennych genetycznie. Celem niniejszego rozdziału rozprawy jest wyznaczenie testu RST oraz WT dla testowania niezależności międzyklasowej w ramach modelu normalnego ze strukturą kowariancyjną BCS oraz porównanie ich z istniejącymi w literaturze testami: LRT, testem zaproponowanym przez Fonseca i in. (2018), oznaczanym w tym rozdziale przez FT, oraz testem Roya (por. Mardia i in., 1979), oznaczanym w tym rozdziale przez RLRT. Warto wspomnieć, że Tsukada (2018) podał m.in. postać statystyki WT, jednakże nie pokazał jej wyprowadzenia. W pracy Filipiak, John, Klein (2023) udowodniliśmy, że wspomniana statystyka WT nie jest zgodna z definicją testu Walda podaną przez Rao (2005). W związku z tym w niniejszym rozdziale, obok postaci statystyki RST, wyprowadzona zostanie również postać statystyki WT. Ponadto, zostanie podany dokładny rozkład statystyki LRT i porównamy za pomocą badań symulacyjnych wspomniane testy między sobą pod kątem szybkości zbieżności do granicznego rozkładu chi-kwadrat, jak również pod kątem mocy z testem FT o rozkładzie F przedstawionym m. in. w pracy Fonseca i in. (2018) oraz testem RLRT o rozkładzie największego pierwiastka Roya (ang. *Roy's largest root*, por. Mardia i in., 1979). Do badania rozbieżności pomiędzy hipotezami posłużymy nam entropijna funkcja straty. Pokazano, że takie rozwiązanie pozwala porównać moc testu dla rozmaitych hipotez alternatywnych, w przeciwieństwie do podejścia zwykle rozważanego w literaturze, gdzie badane są poszczególne struktury hipotez alternatywnych; por. Fonseca i in. (2018), Tsukada (2018). Za pomocą badań symulacyjnych zostanie zbadana również odporność rozkładu rozważanych statystyk testowych na zaburzenie założenia o normalności rozkładu obserwacji. Na koniec proponowane metody zilustrowane zostaną przykładem ogrodniczym, omówionym w

Rozdziale 1.

Wszystkie wyniki zaprezentowane w tym rozdziale znaleźć można w pracy Filipiak, John, Klein (2023).

4.1 Model i hipotezy

Rozważmy model (1.1), w którym macierz kowariancji ma strukturę BCS, por. (1.9). Ponieważ przestrzeń struktur BCS jest podprzestrzenią kwadratową, MLE macierzy Σ_{BCS} jest rzutem $\mathbf{S} = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$ na przestrzeń struktur BCS, czyli

$$\widehat{\Sigma}_{\text{BCS}} = \mathbf{Q}_q \otimes \widehat{\Delta}_1 + \mathbf{P}_q \otimes \widehat{\Delta}_2 = \mathbf{I}_q \otimes \widehat{\Gamma}_0 + (\mathbf{J}_q - \mathbf{I}_q) \otimes \widehat{\Gamma}_1, \quad (4.1)$$

gdzie $\widehat{\Delta}_1$ oraz $\widehat{\Delta}_2$ są zdefiniowane w (2.8), natomiast $\widehat{\Gamma}_0$ i $\widehat{\Gamma}_1$ są określone w (2.9).

Przedmiotem zainteresowania jest zaproponowanie statystyk testowych dla hipotezy związanej z niezależnością międzyklasową, tzn., dla parametryzacji (1.9a)

$$H_0 : \Sigma_0 = \mathbf{I}_q \otimes \Gamma_0 \quad \text{vs} \quad H_1 : \Sigma_1 = \mathbf{I}_q \otimes \Gamma_0 + (\mathbf{J}_q - \mathbf{I}_q) \otimes \Gamma_1 \quad (4.2)$$

lub dla parametryzacji (1.9b),

$$H_0 : \Delta_1 = \Delta_2 \quad \text{vs} \quad H_1 : \Sigma_1 = \mathbf{Q}_q \otimes \Delta_1 + \mathbf{P}_q \otimes \Delta_2. \quad (4.3)$$

Oznaczając w powyższej hipotezie zerowej $\Delta_1 = \Delta_2$ przez Δ , hipotezę (4.3) możemy zapisać również jako

$$H_0 : \Sigma_0 = \mathbf{I}_q \otimes \Delta \quad \text{vs} \quad H_1 : \Sigma_1 = \mathbf{Q}_q \otimes \Delta_1 + \mathbf{P}_q \otimes \Delta_2. \quad (4.4)$$

Z uwagi na to, że przestrzeń macierzy blokowo-diagonalnych jest również podprzestrzenią kwadratową, estymator największej wiarygodności macierzy Σ_0 jest rzutem \mathbf{S} na przestrzeń macierzy blokowo-diagonalnych tzn. $\widehat{\Sigma}_0 = \widehat{\Sigma}_{\text{BI}} = \mathbf{I}_q \otimes \widehat{\Delta}$, gdzie $\widehat{\Delta}$ zdefiniowane jest w (2.7).

4.2 Statystyki testowe

W tej części pracy podamy przegląd testów dla rozważanej hipotezy oraz wyprowadzimy postać testu RST i WT. Testy, FT oraz RLRT, zostały także zaprezentowane m.in. w pracy Kozioł i in. (2021).

4.2.1 Test F

Dla testu hipotezy (4.4) statystyka FT zaproponowana w pracy Fonseca i in. (2018) ma postać

$$FT = \frac{\mathbf{v}' \widetilde{\Delta}_2 \mathbf{v}}{\mathbf{v}' \widetilde{\Delta}_1 \mathbf{v}},$$

gdzie $\widetilde{\Delta}_i$ są estymatorami nieobciążonymi macierzy Δ_i , $i = 1, 2$, natomiast \mathbf{v} jest dowolnym wektorem m -wymiarowym. Przy prawdziwości hipotezy zerowej powyższa statystyka ma rozkład F z $(n - 1)$ i $(n - 1)(m - 1)$ stopniami swobody. W pracy Roy i in. (2016) estymatory te przedstawione zostały jako wielokrotne sumy iloczynów wektorowych. Ponieważ przestrzeń struktur BCS jest podprzestrzenią kwadratową, można je również uzyskać poprzez rzutowanie macierzy kowariancji z próby, $\mathbf{S}_1 = \frac{1}{n-1} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$, na przestrzeń macierzy BCS, otrzymując w rezultacie

$$\widetilde{\Delta}_1 = \frac{1}{u-1} \text{BTr}_m [(\mathbf{Q}_u \otimes \mathbf{I}_m) \mathbf{S}_1], \quad \widetilde{\Delta}_2 = \text{BTr}_m [(\mathbf{P}_u \otimes \mathbf{I}_m) \mathbf{S}_1];$$

por. Filipiak i in. (2020). Zauważmy, że estymatory nieobciążone Δ_1 i Δ_2 różnią się od swoich odpowiedników największej wiarygodności tylko stałą, stąd statystykę F można również przedstawić w postaci

$$F = \frac{\mathbf{v}' \widehat{\Delta}_2 \mathbf{v}}{\mathbf{v}' \widehat{\Delta}_1 \mathbf{v}},$$

gdzie $\widehat{\Delta}_i$ są estymatorami największej wiarygodności Δ_i , $i = 1, 2$, natomiast rozkład F tej statystyki testowej przy prawdziwości hipotezy zerowej wynika z rozkładu Wisharta estymatorów największej wiarygodności oraz ich niezależności; por. Roy i in. (2015). Zatem nawet jeśli przy prawdziwości H_0 rozkład FT nie zależy od prawdziwej wartości Δ , wektor \mathbf{v} powinien być dobrany tak, aby jak najlepiej oddać sens

testowanej hipotezy. Przykładowo, jeśli $\mathbf{v} = \mathbf{e}_i$, $i = 1, \dots, m$, testowana hipoteza dotyczy zerowości i -tego elementu macierzy kowariancji (brak korelacji między sprecyzowanymi cechami w określonym punkcie czasowym), co jest tym samym co (4.4) tylko wtedy, gdy $m = 1$. Jeśli natomiast $\mathbf{v} = \mathbf{1}_m$ (jak założono np. w Fonseca i in., 2018) to hipoteza dotyczy zerowej sumy wszystkich elementów macierzy $\mathbf{\Gamma}_1$ danej w (1.9a). Wówczas proponowana statystyka testowa jest odpowiednia do testowania (4.4) (lub równoważnie (4.2)) jeśli wszystkie elementy macierzy $\mathbf{\Gamma}_1$ są tego samego znaku (lub niektóre z nich, ale nie wszystkie, są zerami). Podobnie, wybór wektora \mathbf{v} jako wektora o składowych nieujemnych (niedodatnich) odpowiada testowaniu wartości sumy ważonej elementów $\mathbf{\Gamma}_1$. Podsumowując, do testowania (4.4) bardziej naturalne wydaje się nie ustalanie pojedynczego \mathbf{v} , ale wybór wektora optymalnego.

4.2.2 Test Roya

Aby uniezależnić się od szczególnego typu hipotezy, czy też znaków korelacji, naturalnym wydaje się przyjęcie za \mathbf{v} wektora maksymalizującego wartość statystyki testowej FT (przypomnijmy, że test odrzuca hipotezę zerową dla dużych wartości statystyki). Zatem maksymalizując FT dla wszystkich $\mathbf{v} \in \mathbb{R}^m$ otrzymujemy statystykę testową $F_\lambda = \lambda_{max}(\widehat{\mathbf{\Delta}}_2 \widehat{\mathbf{\Delta}}_1^{-1})$, gdzie $\lambda_{max}(\cdot)$ oznacza największą wartość własną macierzy $\widehat{\mathbf{\Delta}}_2 \widehat{\mathbf{\Delta}}_1^{-1}$; por. Kozioł i in. (2021). Statystyka postaci

$$\text{RLRT} = \frac{\frac{1}{u-1} F_\lambda}{1 + \frac{1}{u-1} F_\lambda},$$

nazywana statystyką Roya, ma rozkład RLR z parametrami m , $(n-1)(q-1)$ i $(n-1)$; por. Mardia i in. (1979).

Należy zauważyć, że ponieważ wektor \mathbf{v} w F_λ jest wektorem własnym odpowiadającym największej wartości własnej iloczynu $\widehat{\mathbf{\Delta}}_2 \widehat{\mathbf{\Delta}}_1^{-1}$, wartość statystyki zależy od danych. W rezultacie, jak wspomniano w pracy Kozioł i in. (2021), test Roya niekoniecznie ma wyższą moc niż FT. Jest mocniejszy tylko wtedy, gdy największa wartość własna jest znacznie większa od pozostałych wartości własnych.

4.2.3 Test ilorazu wiarygodności

Dla hipotezy (4.4) iloraz wiarygodności Λ ma postać

$$\Lambda = \left(\frac{|\widehat{\Delta}_1|^{q-1} |\widehat{\Delta}_2|}{|\widehat{\Delta}|^q} \right)^{n/2}, \quad (4.5)$$

gdzie $\widehat{\Delta}$, $\widehat{\Delta}_1$, $\widehat{\Delta}_2$ są zdefiniowane odpowiednio w (2.7) i (2.8).

Wiadomo, że przy prawdziwości hipotezy zerowej statystyka $LRT = -2 \ln \Lambda$ ma w przybliżeniu rozkład χ^2 z $m(m+1)/2$ stopniami swobody; por. Rao (2005). Uwzględniając (4.5) statystyka LRT ma postać

$$LRT = -n \left[(q-1) \ln |\widehat{\Delta}_1| + \ln |\widehat{\Delta}_2| - q \ln |\widehat{\Delta}| \right]. \quad (4.6)$$

Jak zauważyli Self, Liang (1987), jeśli parametry kowariancji leżą na granicy ich przestrzeni parametrów, to rozkład asymptotyczny LRT staje się mieszaniną rozkładów χ^2 . Ponadto, ponieważ prędkość zbieżności do rozkładu granicznego zależy od wielkości parametrów układu eksperymentalnego, m, q (por. Rozdział 5.3), dla małej liczebności próby rozkład dokładny może znacznie odbiegać od rozkładu granicznego. W takich przypadkach warto posłużyć się rozkładem dokładnym, którego funkcję charakterystyczną podajemy w następującym twierdzeniu.

Twierdzenie 4.1. *Przy prawdziwości hipotezy zerowej w (4.4) funkcja charakterystyczna statystyki $LRT = -2 \ln \Lambda$, danej w (4.6), ma postać*

$$\varphi(t) = \frac{q^{-nmqit}}{(q-1)^{-n(q-1)mit}} \cdot \prod_{j=1}^m \left[\frac{\Gamma\left(\frac{(n-1)(q-1)+1-j}{2} - itn(q-1)\right)}{\Gamma\left(\frac{(n-1)(q-1)+1-j}{2}\right)} \cdot \frac{\Gamma\left(\frac{n-j}{2} - itn\right)}{\Gamma\left(\frac{n-j}{2}\right)} \cdot \frac{\Gamma\left(\frac{(n-1)q+1-j}{2}\right)}{\Gamma\left(\frac{(n-1)q+1-j}{2} - itnq\right)} \right].$$

Dowód. Niech $\mathbf{A}_1 = n(q-1)\widehat{\Delta}_1$ oraz $\mathbf{A}_2 = n\widehat{\Delta}_2$, gdzie $\widehat{\Delta}_1$ oraz $\widehat{\Delta}_2$ są zdefiniowane w (2.8). Zauważmy ponadto, że estymator $\widehat{\Delta}$ zdefiniowany w (2.7) możemy zapisać jako

$$\begin{aligned} nq\widehat{\Delta} &= n \text{BTr}_m \mathbf{S} = n \left\{ \text{BTr}_m \left[\left(\mathbf{Q}_q \otimes \mathbf{I}_m \right) \mathbf{S} \right] + \text{BTr}_m \left[\left(\mathbf{P}_q \otimes \mathbf{I}_m \right) \mathbf{S} \right] \right\} \\ &= nq(q-1)\widehat{\Delta}_1 + nq\widehat{\Delta}_2 = \mathbf{A}_1 + \mathbf{A}_2. \end{aligned}$$

Wówczas potęga $2/n$ ilorazu (4.5) przyjmuje postać

$$\Lambda_* = \Lambda^{2/n} = \frac{(nq)^q}{n[n(q-1)]^{q-1}} \frac{|\mathbf{A}_1|^{q-1} |\mathbf{A}_2|}{|\mathbf{A}_1 + \mathbf{A}_2|^q} = \frac{q^q}{(q-1)^{q-1}} \frac{|\mathbf{A}_1|^{q-1} |\mathbf{A}_2|}{|\mathbf{A}_1 + \mathbf{A}_2|^q}$$

przy czym $\mathbf{A}_1 \sim W_m(\mathbf{\Delta}_1, (n-1)(q-1))$, $\mathbf{A}_2 \sim W_m(\mathbf{\Delta}_2, n-1)$ oraz \mathbf{A}_1 i \mathbf{A}_2 są niezależne; por. Roy i in. (2015); Filipiak, Klein (2021).

Niech \mathcal{W} oznacza zbiór wartości zmiennej losowej $\mathbf{W} \sim W_m(\mathbf{\Sigma}, \nu)$. Korzystając z funkcji gęstości rozkładu Wisharta, postaci

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{e^{-\frac{1}{2} \text{Tr}(\mathbf{\Sigma}^{-1} \mathbf{w})} |\mathbf{w}|^{(\nu-m-1)/2}}{2^{m\nu/2} \Gamma_m(\frac{\nu}{2}) |\mathbf{\Sigma}|^{\nu/2}},$$

z wielowymiarową funkcją Γ stopnia m postaci

$$\Gamma_m(\frac{\nu}{2}) = \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma(\frac{\nu-i+1}{2}),$$

(por. Kollo, von Rosen, 2005), możemy wyprowadzić wzór na h -ty moment $|\mathbf{W}|$ następująco:

$$\begin{aligned} \mathbb{E}(|\mathbf{W}|^h) &= \int_{\mathcal{W}} |\mathbf{W}|^h \cdot f(\mathbf{W}) d\mathbf{W} \\ &= \int_{\mathcal{W}} |\mathbf{W}|^h \frac{e^{-\frac{1}{2} \text{Tr}(\mathbf{\Sigma}^{-1} \mathbf{w})} |\mathbf{W}|^{(\nu-m-1)/2}}{2^{m\nu/2} \Gamma_m(\frac{\nu}{2}) |\mathbf{\Sigma}|^{\nu/2}} d\mathbf{W} \\ &= \frac{\Gamma_m(\frac{\nu}{2} + h)}{2^{-mh} \Gamma_m(\frac{\nu}{2}) |\mathbf{\Sigma}|^{-h}} \int_{\mathcal{W}} \frac{e^{-\frac{1}{2} \text{Tr}(\mathbf{\Sigma}^{-1} \mathbf{w})} |\mathbf{W}|^{(\nu-m-1+2h)/2}}{2^{m(\nu+2h)/2} \Gamma_m(\frac{\nu}{2} + h) |\mathbf{\Sigma}|^{(\nu+2h)/2}} d\mathbf{W} \\ &= 2^{mh} \cdot |\mathbf{\Sigma}|^h \cdot \frac{\Gamma_m(\frac{\nu}{2} + h)}{\Gamma_m(\frac{\nu}{2})}, \end{aligned} \tag{4.7}$$

gdzie ostatnia równość wynika z faktu, iż funkcja podcałkowa jest gęstością prawdopodobieństwa zmiennej losowej o rozkładzie $W_m(\mathbf{\Sigma}, \nu + 2h)$.

Stąd h -ty moment Λ_* można przedstawić w postaci

$$\begin{aligned}
\mathbb{E}(\Lambda_*^h) &= \frac{q^{mqh}}{(q-1)^{(q-1)mh}} \\
&\cdot \int_{\mathcal{A}_1} \int_{\mathcal{A}_2} \frac{|\mathbf{A}_1|^{(q-1)h} |\mathbf{A}_2|^h}{|\mathbf{A}_1 + \mathbf{A}_2|^{qh}} \\
&\cdot \frac{e^{-\text{Tr}(\Delta^{-1}\mathbf{A}_1)/2} |\mathbf{A}_1|^{[(n-1)(q-1)-m-1]/2}}{2^{m(n-1)(q-1)/2} \Gamma_m\left(\frac{(n-1)(q-1)}{2}\right) |\Delta|^{(n-1)(q-1)/2}} \\
&\cdot \frac{e^{-\text{Tr}(\Delta^{-1}\mathbf{A}_2)/2} |\mathbf{A}_2|^{(n-2-m)/2}}{2^{m(n-1)/2} \Gamma_m\left(\frac{n-1}{2}\right) |\Delta|^{(n-1)/2}} d\mathbf{A}_1 d\mathbf{A}_2 \\
&= \frac{q^{mqh}}{(q-1)^{(q-1)mh}} \frac{|\Delta|^{(n-1+2h)(q-1)/2}}{|\Delta|^{(n-1)(q-1)/2}} \cdot \frac{|\Delta|^{(n-1+2h)/2}}{|\Delta|^{(n-1)/2}} \\
&\cdot 2^{mqh} \cdot \frac{\Gamma_m\left(\frac{(n-1+2h)(q-1)}{2}\right)}{\Gamma_m\left(\frac{(n-1)(q-1)}{2}\right)} \cdot \frac{\Gamma_m\left(\frac{n-1+2h}{2}\right)}{\Gamma_m\left(\frac{n-1}{2}\right)} \\
&\cdot \int_{\mathcal{A}_1} \int_{\mathcal{A}_2} |\mathbf{A}_1 + \mathbf{A}_2|^{-qh} \\
&\cdot \frac{e^{-\text{Tr}(\Delta^{-1}\mathbf{A}_1)/2} |\mathbf{A}_1|^{[(q-1)(n-1+2h)-m-1]/2}}{2^{m(n-1+2h)(q-1)/2} \Gamma_m\left(\frac{(n-1+2h)(q-1)}{2}\right) |\Delta|^{(n-1+2h)(q-1)/2}} \\
&\cdot \frac{e^{-\text{Tr}(\Delta^{-1}\mathbf{A}_2)/2} |\mathbf{A}_2|^{(n-1+2h-m-1)/2}}{2^{m(n-1+2h)/2} \Gamma_m\left(\frac{n-1+2h}{2}\right) |\Delta|^{(n-1+2h)/2}} d\mathbf{A}_1 d\mathbf{A}_2 \\
&= \frac{(2q)^{mqh}}{(q-1)^{(q-1)mh}} \cdot \frac{\Gamma_m\left(\frac{(n-1+2h)(q-1)}{2}\right) \Gamma_m\left(\frac{n-1+2h}{2}\right)}{\Gamma_m\left(\frac{(n-1)(q-1)}{2}\right) \Gamma_m\left(\frac{n-1}{2}\right)} \cdot |\Delta|^{qh} \cdot \mathbb{E}(|\mathbf{A}|^{-qh}).
\end{aligned}$$

Ponieważ $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 \sim W_m(\Delta, (n-1+2h)u)$, korzystając z (4.7) otrzymujemy

$$\mathbb{E}(\Lambda_*^h) = \frac{q^{mqh}}{(q-1)^{(q-1)mh}} \frac{\Gamma_m\left(\frac{(n-1+2h)(q-1)}{2}\right) \Gamma_m\left(\frac{n-1+2h}{2}\right) \Gamma_m\left(\frac{(n-1+2h)q-2qh}{2}\right)}{\Gamma_m\left(\frac{(n-1)(q-1)}{2}\right) \Gamma_m\left(\frac{n-1}{2}\right) \Gamma_m\left(\frac{q(n-1+2h)}{2}\right)}$$

$$\begin{aligned}
&= \frac{q^{mqh}}{(q-1)^{(q-1)mh}} \frac{\Gamma_m\left(\frac{(n-1+2h)(q-1)}{2}\right) \Gamma_m\left(\frac{n-1+2h}{2}\right) \Gamma_m\left(\frac{(n-1)q}{2}\right)}{\Gamma_m\left(\frac{(n-1)(q-1)}{2}\right) \Gamma_m\left(\frac{n-1}{2}\right) \Gamma_m\left(\frac{q(n-1+2h)}{2}\right)} \\
&= \frac{q^{mqh}}{(q-1)^{(q-1)mh}} \cdot \prod_{j=1}^m \left[\frac{\Gamma\left(\frac{(n-1)(q-1)+1-j}{2} + h(q-1)\right)}{\Gamma\left(\frac{(n-1)(q-1)+1-j}{2}\right)} \cdot \frac{\Gamma\left(\frac{n-j}{2} + h\right)}{\Gamma\left(\frac{n-j}{2}\right)} \cdot \frac{\Gamma\left(\frac{(n-1)q+1-j}{2}\right)}{\Gamma\left(\frac{(n-1)q+1-j}{2} + hq\right)} \right],
\end{aligned}$$

co oznacza, że jest to h -ty moment zmiennej losowej o rozkładzie z funkcją charakterystyczną podaną w twierdzeniu. \square

Ponieważ znajomość wszystkich momentów zmiennej losowej jednoznacznie definiuje jej rozkład prawdopodobieństwa (por. Billingsley, 1995), podana funkcja charakterystyczna identyfikuje rozkład dokładny statystyki LRT.

Do zilustrowania funkcji gęstości rozkładu statystyki LRT można posłużyć się pakietem Matlab `CharFunTool` (Witkovský, 2018), czy też pakietem R `CharFunToolR` opracowanym przez Gajdoš (2018).

4.2.4 Test wynikowy Rao

W tej części pracy sformułujemy postać statystyki testowej Rao.

Uwzględniając Definicję 1.3, wektor parametrów przy prawdziwości hipotezy alternatywnej z (4.4) jest postaci $\boldsymbol{\theta} = (\boldsymbol{\mu}', \text{vech}' \boldsymbol{\Delta}_1, \text{vech}' \boldsymbol{\Delta}_2)'$, natomiast przy prawdziwości hipotezy zerowej ma on postać $\boldsymbol{\theta}_0 = (\boldsymbol{\mu}', \text{vech}' \boldsymbol{\Delta}, \text{vech}' \boldsymbol{\Delta})'$. Oznaczmy $\hat{\boldsymbol{\theta}}_0 = \text{MLE}(\boldsymbol{\theta}_0)$, gdzie odpowiedni estymator największej wiarygodności macierzy $\boldsymbol{\Delta}$ zdefiniowany jest w (2.7). Z uwagi na to, że $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_0$, to pierwsza składowa wektora wynikowego $\mathbf{s}(\hat{\boldsymbol{\theta}}_0)$ (pochodna pierwszego rzędu względem $\boldsymbol{\mu}$) jest równa 0. Dlatego bez straty ogólności możemy zredukować wymiar wektora wynikowego o mq . Dodatkowo, elementy poza główną przekątną w pierwszym wierszu i kolumnie macierzy informacji Fishera również zredukują się do $\mathbf{0}$. Z uwagi na ten fakt, $\boldsymbol{\theta}$ oraz $\boldsymbol{\theta}_0$ w tym rozdziale są rozważane bez pierwszego składnika, a macierz informacji Fishera jest stopnia $m(m+1)$.

Twierdzenie 4.2. Statystyka wynikowa Rao dla testu hipotezy (4.4) ma postać

$$\text{RST} = \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - \widehat{\Sigma}_{\text{BCS}} \left(\mathbf{I}_q \otimes \widehat{\Delta}^{-1} \right) \right]^2 \right\},$$

gdzie $\widehat{\Sigma}_{\text{BCS}}$ oraz $\widehat{\Delta}$ są estymatorami największej wiarygodności Σ_{BCS} i Δ wyrażonymi odpowiednio wzorami (4.1) oraz (2.7).

Dowód. Rozważmy hipotezę (4.4). Przy prawdziwości hipotezy alternatywnej logarytm funkcji wiarygodności przyjmuje postać

$$\ln L = -\frac{nmq}{2} \ln(2\pi) - \frac{n(q-1)}{2} \ln |\Delta_1| - \frac{n}{2} \ln |\Delta_2| - \frac{1}{2} \text{Tr}(\mathbf{Z}\Sigma_{\text{BCS}}^{-1}\mathbf{Z}'), \quad (4.8)$$

gdzie $\mathbf{Z} := \mathbf{Z}(\boldsymbol{\mu}) = \mathbf{Y} - \mathbf{1}_n\boldsymbol{\mu}'$.

Różniczkując (4.8) względem $\boldsymbol{\theta}$, przy zastosowaniu reguły łańcuchowej (zob. (1.20)), wzorów na pochodne z Lematu 1.3, a także korzystając z Własności 1.2(i), oraz mając na uwadze, że $\text{Tr} \mathbf{Q}_q = q - 1$ oraz $\text{Tr} \mathbf{P}_q = 1$, otrzymujemy

$$\begin{aligned} \frac{\partial \ln L}{\partial \Delta_1} &= -\frac{n(q-1)}{2} \text{vec}' \Delta_1^{-1} \mathbf{D}_m \\ &\quad + \frac{n}{2} \text{vec}' \left(\frac{1}{n} \mathbf{Z}' \mathbf{Z} \right) (\mathbf{I}_q \otimes \mathbf{K}_{m,q} \otimes \mathbf{I}_m) (\text{vec} \mathbf{Q}_q \otimes \mathbf{I}_{m^2}) (\Delta_1^{-1} \otimes \Delta_1^{-1}) \mathbf{D}_m \\ &= \left(-\frac{n(q-1)}{2} \text{vec}' \Delta_1^{-1} + \frac{n}{2} \text{vec}' \{ \text{BTr}_m [(\mathbf{Q}_q \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}' \mathbf{Z}] \} (\Delta_1^{-1} \otimes \Delta_1^{-1}) \right) \mathbf{D}_m, \\ \frac{\partial \ln L}{\partial \Delta_2} &= -\frac{n}{2} \text{vec}' \Delta_2^{-1} \mathbf{D}_m \\ &\quad + \frac{n}{2} \text{vec}' \left(\frac{1}{n} \mathbf{Z} \mathbf{Z}' \right) (\mathbf{I}_q \otimes \mathbf{K}_{m,q} \otimes \mathbf{I}_m) (\text{vec} \mathbf{P}_q \otimes \mathbf{I}_{m^2}) (\Delta_2^{-1} \otimes \Delta_2^{-1}) \mathbf{D}_m \\ &= \left(-\frac{n}{2} \text{vec}' \Delta_2^{-1} + \frac{n}{2} \text{vec}' \{ \text{BTr}_m [(\mathbf{P}_q \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z} \mathbf{Z}'] \} (\Delta_2^{-1} \otimes \Delta_2^{-1}) \right) \mathbf{D}_m. \end{aligned}$$

Wstawiając do powyższych formuł $\widehat{\boldsymbol{\theta}}_0$ z estymatorem Δ danym wzorem (2.7) oraz uwzględniając

$$\frac{1}{n} \mathbf{Z}' \mathbf{Z} = \frac{1}{n} \mathbf{Z}'(\widehat{\boldsymbol{\mu}}) \mathbf{Z}(\widehat{\boldsymbol{\mu}}) = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y} = \mathbf{S},$$

uzyskujemy wektor wynikowy postaci

$$\mathbf{s}(\widehat{\boldsymbol{\theta}}_0) = -\frac{n}{2} (\mathbf{I}_2 \otimes \mathbf{D}'_m) \begin{pmatrix} (q-1) \text{vec} \widehat{\Delta}^{-1} - (\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \text{vec} \{ \text{BTr}_m [(\mathbf{Q}_q \otimes \mathbf{I}_m) \mathbf{S}] \} \\ \text{vec} \widehat{\Delta}^{-1} - (\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \text{vec} \{ \text{BTr}_m [(\mathbf{P}_q \otimes \mathbf{I}_m) \mathbf{S}] \} \end{pmatrix}.$$

Zauważmy, że korzystając ze wzorów (2.8) oraz Własności 1.1(iii), możemy powyższy wektor wynikowy zapisać jako

$$\mathbf{s}(\hat{\boldsymbol{\theta}}_0) = -\frac{n}{2}(\mathbf{I}_2 \otimes \mathbf{D}'_m) \begin{pmatrix} (q-1) \left[\text{vec } \widehat{\boldsymbol{\Delta}}^{-1} - (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \text{vec } \widehat{\boldsymbol{\Delta}}_1 \right] \\ \text{vec } \widehat{\boldsymbol{\Delta}}^{-1} - (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \text{vec } \widehat{\boldsymbol{\Delta}}_2 \end{pmatrix},$$

lub równoważnie

$$\mathbf{s}(\hat{\boldsymbol{\theta}}_0) = -\frac{n}{2}(\mathbf{I}_2 \otimes \mathbf{D}'_m) \begin{pmatrix} (q-1)(\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \text{vec}(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_1) \\ (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \text{vec}(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_2) \end{pmatrix}. \quad (4.9)$$

Korzystając ze wzorów na różniczkowanie macierzy otrzymujemy dalej

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_1^2} &= \mathbf{D}'_m \left\{ \frac{n(q-1)}{2} (\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}) \right. \\ &\quad \left. - \frac{n}{2} \left[(\text{vec}' \{ \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}'\mathbf{Z}] \} \otimes \mathbf{I}_{m^2}) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) \right. \right. \\ &\quad \left. \left. \cdot (\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1} \otimes \text{vec}' \boldsymbol{\Delta}_1^{-1} + \text{vec } \boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}) \right] \right\} \mathbf{D}_m, \\ \frac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_2^2} &= \mathbf{D}'_m \left(\frac{n}{2} (\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}) \right. \\ &\quad \left. - \frac{n}{2} \left[(\text{vec}' \{ \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}'\mathbf{Z}] \} \otimes \mathbf{I}_{m^2}) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) \right. \right. \\ &\quad \left. \left. \cdot (\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1} \otimes \text{vec}' \boldsymbol{\Delta}_2^{-1} + \text{vec } \boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}) \right] \right\} \mathbf{D}_m. \end{aligned}$$

Uwzględniając operatory śladów blokowych w powyższych wyrażeniach i wykorzystując (1.9b), ortogonalność projektorów \mathbf{Q}_q i \mathbf{P}_q oraz Twierdzenie 2.2.9(i) podane przez Kollo, von Rosen (2005) mamy, że $\mathbb{E}(\mathbf{Z}'\mathbf{Z}) = n\boldsymbol{\Sigma}_{\text{BCS}}$. Możemy zatem zapisać, że $\mathbb{E}\{\text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}'\mathbf{Z}]\} = \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m) \boldsymbol{\Sigma}_{\text{BCS}}] = \text{BTr}_m[(\mathbf{Q}_q \otimes \boldsymbol{\Delta}_1)]$ oraz analogicznie $\mathbb{E}\{\text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m) \frac{1}{n} \mathbf{Z}'\mathbf{Z}]\} = \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m) \boldsymbol{\Sigma}_{\text{BCS}}] = \text{BTr}_m[(\mathbf{P}_q \otimes \boldsymbol{\Delta}_2)]$. Dalej, z definicji \mathbf{Q}_q i \mathbf{P}_q możemy zauważyć, że bloki na przekątnej głównej macierzy $\mathbf{Q}_q \otimes \boldsymbol{\Delta}_1$ i $\mathbf{P}_q \otimes \boldsymbol{\Delta}_2$ są odpowiednio równe $\frac{q-1}{q} \boldsymbol{\Delta}_1$ i $\frac{1}{q} \boldsymbol{\Delta}_2$ i jest ich q , co implikuje, że

$\text{BTr}_m(\mathbf{Q}_q \otimes \mathbf{\Delta}_1) = (q-1)\mathbf{\Delta}_1$ i $\text{BTr}_m(\mathbf{P}_q \otimes \mathbf{\Delta}_2) = \mathbf{\Delta}_2$. Stąd,

$$\begin{aligned}
-\mathbb{E} \left(\frac{\partial^2 \ln L}{\partial \mathbf{\Delta}_1^2} \right) &= -\frac{n(q-1)}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&\quad + \frac{n(q-1)}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_1 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) \\
&\quad \quad \cdot (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1} \otimes \text{vec} \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&\quad + \frac{n(q-1)}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_1 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) \\
&\quad \quad \cdot (\text{vec} \widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m, \\
-\mathbb{E} \left(\frac{\partial^2 \ln L}{\partial \mathbf{\Delta}_2^2} \right) &= -\frac{n}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&\quad + \frac{n}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_2 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1} \otimes \text{vec} \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&\quad + \frac{n}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_2 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\text{vec} \widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m.
\end{aligned}$$

Korzystając z Lematu 1.1 otrzymujemy

$$\begin{aligned}
\mathbf{F}_{11}(\widehat{\boldsymbol{\theta}}) &= -\frac{n(q-1)}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&\quad + \frac{n(q-1)}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_1 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\mathbf{I}_{m^2} \otimes \text{vec} \widehat{\mathbf{\Delta}}_1^{-1}) (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&\quad + \frac{n(q-1)}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_1 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\text{vec} \widehat{\mathbf{\Delta}}_1^{-1} \otimes \mathbf{I}_{m^2}) (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&= -\frac{n(q-1)}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m + \frac{n(q-1)}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&\quad + \frac{n(q-1)}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m \\
&= \frac{n(q-1)}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_1^{-1} \otimes \widehat{\mathbf{\Delta}}_1^{-1}) \mathbf{D}_m, \tag{4.10}
\end{aligned}$$

$$\begin{aligned}
\mathbf{F}_{22}(\widehat{\boldsymbol{\theta}}) &= -\frac{n}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&\quad + \frac{n}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_2 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\mathbf{I} \otimes \text{vec} \widehat{\mathbf{\Delta}}_2^{-1}) (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&\quad + \frac{n}{2} (\text{vec}' \widehat{\mathbf{\Delta}}_2 \otimes \mathbf{D}'_m) (\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m) (\text{vec} \widehat{\mathbf{\Delta}}_2^{-1} \otimes \mathbf{I}) (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&= -\frac{n}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m + \frac{n}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&\quad + \frac{n}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m \\
&= \frac{n}{2} \mathbf{D}'_m (\widehat{\mathbf{\Delta}}_2^{-1} \otimes \widehat{\mathbf{\Delta}}_2^{-1}) \mathbf{D}_m. \tag{4.11}
\end{aligned}$$

Z uwagi na to, że dla $i \neq j, i, j = 1, 2$, pochodna $\frac{\partial \ln L}{\partial \mathbf{\Delta}_i}$ nie zależy od $\mathbf{\Delta}_j$, bloki pozadiagonalne macierzy informacji Fishera są równe zero. Wstawiając estymatory

największej wiarygodności nieznanymi parametrami przy prawdziwości H_0 , macierz informacji Fishera przyjmuje zatem postać

$$\mathbf{F}(\hat{\boldsymbol{\theta}}_0) = \frac{n}{2} \begin{pmatrix} (q-1)\mathbf{D}'_m(\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1})\mathbf{D}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{D}'_m(\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1})\mathbf{D}_m \end{pmatrix}.$$

Korzystając z (1.19) mamy, że

$$\begin{aligned} \mathbf{F}^{-1}(\hat{\boldsymbol{\theta}}_0) &= \frac{2}{n} \begin{pmatrix} \frac{1}{(q-1)}\mathbf{D}_m^+(\widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}})\mathbf{D}_m^{+'} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_m^+(\widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}})\mathbf{D}_m^{+'} \end{pmatrix} \\ &= \frac{2}{n}(\mathbf{I}_2 \otimes \mathbf{D}_m^+) \begin{pmatrix} \frac{1}{(q-1)}(\widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}}) & \mathbf{0} \\ \mathbf{0} & \widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}} \end{pmatrix} (\mathbf{I}_2 \otimes \mathbf{D}_m^{+'}). \end{aligned}$$

Zapisując wektor wynikowy (4.9) w postaci

$$\mathbf{s}(\hat{\boldsymbol{\theta}}_0) = -\frac{n}{2}(\mathbf{I}_2 \otimes \mathbf{D}'_m) \begin{pmatrix} \hat{\mathbf{s}}_1 \\ \hat{\mathbf{s}}_2 \end{pmatrix},$$

gdzie $\hat{\mathbf{s}}_1 = (q-1)(\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \text{vec}(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_1)$ oraz $\hat{\mathbf{s}}_2 = (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \text{vec}(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_2)$, statystykę RST wyznaczmy jako

$$\text{RST} = \frac{n}{2} (\hat{\mathbf{s}}_1', \hat{\mathbf{s}}_2') (\mathbf{I}_2 \otimes \mathbf{D}_m \mathbf{D}_m^+) \begin{pmatrix} \frac{1}{(q-1)}(\widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}}) & \mathbf{0} \\ \mathbf{0} & \widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}} \end{pmatrix} (\mathbf{I}_2 \otimes \mathbf{D}_m^{+'} \mathbf{D}_m') \begin{pmatrix} \hat{\mathbf{s}}_1 \\ \hat{\mathbf{s}}_2 \end{pmatrix}.$$

Otrzymujemy zatem

$$\begin{aligned} \text{RST} &= \frac{n(q-1)}{2} \text{vec}'(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_1) (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \mathbf{D}_m \mathbf{D}_m^+ (\widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}}) \mathbf{D}_m^{+'} \mathbf{D}_m' (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \\ &\quad \cdot \text{vec}(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_1) \\ &+ \frac{n}{2} \text{vec}'(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_2) (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \mathbf{D}_m \mathbf{D}_m^+ (\widehat{\boldsymbol{\Delta}} \otimes \widehat{\boldsymbol{\Delta}}) \mathbf{D}_m^{+'} \mathbf{D}_m' (\widehat{\boldsymbol{\Delta}}^{-1} \otimes \widehat{\boldsymbol{\Delta}}^{-1}) \\ &\quad \cdot \text{vec}(\widehat{\boldsymbol{\Delta}} - \widehat{\boldsymbol{\Delta}}_2). \end{aligned}$$

Korzystając z faktu, że $\mathbf{D}_m \mathbf{D}_m^+ = \mathbf{N}_m$ (por. Magnus, Neudecker, 1986) oraz z zależ-

ności (1.17) i (1.18) mamy

$$\begin{aligned}
\text{RST} &= \frac{n(q-1)}{2} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_1)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1})\mathbf{N}_m(\widehat{\Delta} \otimes \widehat{\Delta})\mathbf{N}_m(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \\
&\quad \cdot \text{vec}(\widehat{\Delta} - \widehat{\Delta}_1) \\
&+ \frac{n}{2} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_2)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1})\mathbf{N}_m(\widehat{\Delta} \otimes \widehat{\Delta})\mathbf{N}_m(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \\
&\quad \cdot \text{vec}(\widehat{\Delta} - \widehat{\Delta}_2) \\
&= \frac{n(q-1)}{8} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_1)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1})(\mathbf{I}_{m^2} + \mathbf{K}_{m,m})(\widehat{\Delta} \otimes \widehat{\Delta})(\mathbf{I}_{m^2} + \mathbf{K}_{m,m}) \\
&\quad \cdot (\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \text{vec}(\widehat{\Delta} - \widehat{\Delta}_1) \\
&+ \frac{n}{8} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_2)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1})(\mathbf{I}_{m^2} + \mathbf{K}_{m,m})(\widehat{\Delta} \otimes \widehat{\Delta})(\mathbf{I}_{m^2} + \mathbf{K}_{m,m}) \\
&\quad \cdot (\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \text{vec}(\widehat{\Delta} - \widehat{\Delta}_2) \\
&= \frac{n(q-1)}{2} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_1)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1})(\widehat{\Delta} \otimes \widehat{\Delta})(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \\
&\quad \cdot \text{vec}(\widehat{\Delta} - \widehat{\Delta}_1) \\
&+ \frac{n}{2} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_2)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1})(\widehat{\Delta} \otimes \widehat{\Delta})(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \\
&\quad \cdot \text{vec}(\widehat{\Delta} - \widehat{\Delta}_2) \\
&= \frac{n(q-1)}{2} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_1)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \text{vec}(\widehat{\Delta} - \widehat{\Delta}_1) \\
&+ \frac{n}{2} \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_2)(\widehat{\Delta}^{-1} \otimes \widehat{\Delta}^{-1}) \text{vec}(\widehat{\Delta} - \widehat{\Delta}_2).
\end{aligned}$$

Z Własności 1.1(ii, iii) operatora wektoryzacji otrzymujemy dalej, że

$$\begin{aligned}
\text{RST} &= \frac{n}{2} \left\{ (q-1) \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_1) \text{vec}[\widehat{\Delta}^{-1}(\widehat{\Delta} - \widehat{\Delta}_1)\widehat{\Delta}^{-1}] \right. \\
&\quad \left. + \text{vec}'(\widehat{\Delta} - \widehat{\Delta}_2) \text{vec}[\widehat{\Delta}^{-1}(\widehat{\Delta} - \widehat{\Delta}_2)\widehat{\Delta}^{-1}] \right\} \\
&= \frac{n}{2} \left\{ (q-1) \text{Tr}[(\widehat{\Delta} - \widehat{\Delta}_1)\widehat{\Delta}^{-1}(\widehat{\Delta} - \widehat{\Delta}_1)\widehat{\Delta}^{-1}] \right. \\
&\quad \left. + \text{Tr}[(\widehat{\Delta} - \widehat{\Delta}_2)\widehat{\Delta}^{-1}(\widehat{\Delta} - \widehat{\Delta}_2)\widehat{\Delta}^{-1}] \right\} \\
&= \frac{n}{2} \left[(q-1) \text{Tr}(\mathbf{I}_m - 2\widehat{\Delta}_1\widehat{\Delta}^{-1} + \widehat{\Delta}_1\widehat{\Delta}^{-1}\widehat{\Delta}_1\widehat{\Delta}^{-1}) \right. \\
&\quad \left. + \text{Tr}(\mathbf{I}_m - 2\widehat{\Delta}_2\widehat{\Delta}^{-1} + \widehat{\Delta}_2\widehat{\Delta}^{-1}\widehat{\Delta}_2\widehat{\Delta}^{-1}) \right].
\end{aligned}$$

Z idempotentności projektorów \mathbf{Q}_q i \mathbf{P}_q , z faktu, że, $\mathbf{Q}_q + \mathbf{P}_q = \mathbf{I}_q$ oraz ponownie

wykorzystując Własność 1.1(iv), powyższe wyrażenie można przekształcić do postaci

$$\begin{aligned}
\text{RST} &= \frac{n}{2} \left[\text{Tr} \mathbf{Q}_q \text{Tr}(\mathbf{I}_m - 2\widehat{\Delta}_1 \widehat{\Delta}^{-1} + \widehat{\Delta}_1 \widehat{\Delta}^{-1} \widehat{\Delta}_1 \widehat{\Delta}^{-1}) \right. \\
&\quad \left. + \text{Tr} \mathbf{P}_q \text{Tr}(\mathbf{I}_m - 2\widehat{\Delta}_2 \widehat{\Delta}^{-1} + \widehat{\Delta}_2 \widehat{\Delta}^{-1} \widehat{\Delta}_2 \widehat{\Delta}^{-1}) \right] \\
&= \frac{n}{2} \left[\text{Tr}(\mathbf{Q}_q \otimes \mathbf{I}_m) - 2 \text{Tr}(\mathbf{Q}_q \otimes \widehat{\Delta}_1 \widehat{\Delta}^{-1}) + \text{Tr}(\mathbf{Q}_q \otimes \widehat{\Delta}_1 \widehat{\Delta}^{-1} \widehat{\Delta}_1 \widehat{\Delta}^{-1}) \right. \\
&\quad \left. + \text{Tr}(\mathbf{P}_q \otimes \mathbf{I}_m) - 2 \text{Tr}(\mathbf{P}_q \otimes \widehat{\Delta}_2 \widehat{\Delta}^{-1}) + \text{Tr}(\mathbf{P}_q \otimes \widehat{\Delta}_2 \widehat{\Delta}^{-1} \widehat{\Delta}_2 \widehat{\Delta}^{-1}) \right] \\
&= \frac{n}{2} \text{Tr} \left[\mathbf{I}_{mq} \right. \\
&\quad - 2(\mathbf{Q}_q \otimes \widehat{\Delta}_1)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1}) \\
&\quad + (\mathbf{Q}_q \otimes \widehat{\Delta}_1)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1})(\mathbf{Q}_q \otimes \widehat{\Delta}_1)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1}) \\
&\quad - 2(\mathbf{P}_q \otimes \widehat{\Delta}_2)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1}) \\
&\quad \left. + (\mathbf{P}_q \otimes \widehat{\Delta}_2)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1})(\mathbf{P}_q \otimes \widehat{\Delta}_2)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1}) \right].
\end{aligned}$$

Ostatecznie z ortogonalności \mathbf{Q}_q i \mathbf{P}_q otrzymujemy

$$\begin{aligned}
\text{RST} &= \frac{n}{2} \text{Tr} \left[\mathbf{I}_{mq} - 2(\mathbf{Q}_q \otimes \widehat{\Delta}_1 + \mathbf{P}_q \otimes \widehat{\Delta}_2)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1}) \right. \\
&\quad \left. + (\mathbf{Q}_q \otimes \widehat{\Delta}_1 + \mathbf{P}_q \otimes \widehat{\Delta}_2)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1})(\mathbf{Q}_q \otimes \widehat{\Delta}_1 + \mathbf{P}_q \otimes \widehat{\Delta}_2)(\mathbf{I}_q \otimes \widehat{\Delta}^{-1}) \right] \\
&= \frac{n}{2} \text{Tr} \left[\mathbf{I}_{mq} - 2\widehat{\Sigma}_{\text{BCS}}(\mathbf{I}_q \otimes \widehat{\Delta})^{-1} + \widehat{\Sigma}_{\text{BCS}}(\mathbf{I}_q \otimes \widehat{\Delta})^{-1} \widehat{\Sigma}_{\text{BCS}}(\mathbf{I}_q \otimes \widehat{\Delta})^{-1} \right] \\
&= \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - \widehat{\Sigma}_{\text{BCS}}(\mathbf{I}_q \otimes \widehat{\Delta})^{-1} \right]^2 \right\}.
\end{aligned}$$

□

Oznaczając estymatory największej wiarygodności macierzy kowariancji przy prawdziwości hipotezy zerowej i alternatywnej odpowiednio poprzez $\widehat{\Sigma}_0$ i $\widehat{\Sigma}_1$ łatwo zauważyć, że

$$\text{RST} = \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - \widehat{\Sigma}_1 \widehat{\Sigma}_0^{-1} \right]^2 \right\}. \quad (4.12)$$

Ze względu na zależność między parametryzacją (1.9a) i (1.9b), postać statystyki RST można sformułować w terminologii macierzy $\mathbf{\Gamma}_0$ oraz $\mathbf{\Gamma}_1$, co przedstawia poniższy wniosek.

Wniosek 4.1. *Statystyka wynikowa Rao dla testu hipotezy (4.2) ma postać*

$$\text{RST} = \frac{nq(q-1)}{2} \text{Tr} \left[(\hat{\Gamma}_0^{-1} \hat{\Gamma}_1)^2 \right],$$

gdzie $\hat{\Gamma}_0$ oraz $\hat{\Gamma}_1$ są estymatorami największej wiarygodności Γ_0 i Γ_1 wyrażonymi wzorami (2.9).

Dowód. Uwzględniając we wzorze (4.12) parametryzację (1.9a) oraz korzystając z Własności 1.1(iv), a także z faktu, iż $\text{Tr} [(\mathbf{J}_q - \mathbf{I}_q)^2] = q(q-1)$, otrzymujemy

$$\begin{aligned} \text{RST} &= \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - (\mathbf{I}_q \otimes \hat{\Gamma}_0^{-1})(\mathbf{I}_q \otimes \hat{\Gamma}_0 + (\mathbf{J}_q - \mathbf{I}_q) \otimes \hat{\Gamma}_1) \right]^2 \right\} \\ &= \frac{n}{2} \text{Tr} \left\{ \left[(\mathbf{J}_q - \mathbf{I}_q) \otimes (\hat{\Gamma}_0^{-1} \hat{\Gamma}_1) \right]^2 \right\} = \frac{n}{2} \text{Tr} [(\mathbf{J}_q - \mathbf{I}_q)^2] \cdot \text{Tr} \left[(\hat{\Gamma}_0^{-1} \hat{\Gamma}_1)^2 \right] \\ &= \frac{nq(q-1)}{2} \text{Tr} \left[(\hat{\Gamma}_0^{-1} \hat{\Gamma}_1)^2 \right]. \end{aligned}$$

□

Na koniec przypomnijmy, że przy założeniu prawdziwości hipotezy zerowej i rosnącej wielkości próby, rozkład statystyki RST dąży do rozkładu χ^2 z $m(m+1)/2$ stopniami swobody.

4.2.5 Test Walda

Kolejnym rozpatrywanym testem jest test Walda. Mając na uwadze postać (1.15) statystyki testowej WT hipotezę (4.3) można zapisać w postaci

$$H_0 : \text{vech}(\Delta_1 - \Delta_2) = \mathbf{0}_{m(m+1)/2} \quad \text{vs} \quad H_1 : \text{vech}(\Delta_1 - \Delta_2) \neq \mathbf{0}_{m(m+1)/2}. \quad (4.13)$$

Hipotezę zerową (4.13) można równoważnie przedstawić w postaci

$$\left(\mathbf{I}_{m(m+1)/2}, -\mathbf{I}_{m(m+1)/2} \right) \boldsymbol{\theta} = \mathbf{0}_{m(m+1)/2}$$

gdzie $\boldsymbol{\theta} = (\text{vech}' \Delta_1, \text{vech}' \Delta_2)'$, a stąd w (1.15) otrzymujemy

$$\mathbf{w}(\boldsymbol{\theta}) = \left(\mathbf{I}_{m(m+1)/2}, -\mathbf{I}_{m(m+1)/2} \right) \boldsymbol{\theta}, \quad \mathbf{W}(\boldsymbol{\theta}) = \left(\mathbf{I}_{m(m+1)/2}, -\mathbf{I}_{m(m+1)/2} \right),$$

natomiast $\mathbf{F}^{-1}(\boldsymbol{\theta})$ jest macierzą blokowo-diagonalną o blokach postaci (4.10) i (4.11). Podobnie jak rozkłady LRT i RST, przy rosnącej wielkości próby również rozkład WT dąży do rozkładu χ^2 z $m(m+1)/2$ stopniami swobody, a statystyka Walda przedstawiona jest w poniższym twierdzeniu.

Twierdzenie 4.3. *Statystyka Walda dla testu hipotezy (4.13) ma postać*

$$\text{WT} = \frac{n}{2} \text{vec}'(\widehat{\boldsymbol{\Delta}}_1 - \widehat{\boldsymbol{\Delta}}_2) \left[\frac{1}{q-1} (\widehat{\boldsymbol{\Delta}}_1 \otimes \widehat{\boldsymbol{\Delta}}_1) + (\widehat{\boldsymbol{\Delta}}_2 \otimes \widehat{\boldsymbol{\Delta}}_2) \right]^{-1} \text{vec}(\widehat{\boldsymbol{\Delta}}_1 - \widehat{\boldsymbol{\Delta}}_2),$$

gdzie $\widehat{\boldsymbol{\Delta}}_1$ oraz $\widehat{\boldsymbol{\Delta}}_2$ są estymatorami największej wiarygodności $\boldsymbol{\Delta}_1$ i $\boldsymbol{\Delta}_2$ wyrażonymi wzorami (2.8).

Dowód. Na podstawie pracy Rao (2005) statystykę testową Walda można przedstawić korzystając z (1.15). Stąd,

$$\begin{aligned} \mathbf{A}(\boldsymbol{\theta}) &= \frac{2}{n} \left(\mathbf{I}_{m(m+1)/2}, -\mathbf{I}_{m(m+1)/2} \right) \begin{pmatrix} \frac{1}{q-1} \mathbf{D}_m^+(\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1) \mathbf{D}_m^{+'} & \mathbf{0}_{m(m+1)/2} \\ \mathbf{0}_{m(m+1)/2} & \mathbf{D}_m^+(\boldsymbol{\Delta}_2 \otimes \boldsymbol{\Delta}_2) \mathbf{D}_m^{+'} \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} \mathbf{I}_{m(m+1)/2} \\ -\mathbf{I}_{m(m+1)/2} \end{pmatrix} \\ &= \frac{2}{n} \mathbf{D}_m^+ \left[\frac{1}{q-1} (\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1) + (\boldsymbol{\Delta}_2 \otimes \boldsymbol{\Delta}_2) \right] \mathbf{D}_m^{+'}, \end{aligned}$$

i, wstawiając odwrotność $\mathbf{A}(\widehat{\boldsymbol{\theta}})$ do wzoru na statystykę testową Walda,

$$\text{WT} = \frac{n}{2} \text{vech}'(\widehat{\boldsymbol{\Delta}}_1 - \widehat{\boldsymbol{\Delta}}_2) \mathbf{D}_m' \left[\frac{1}{q-1} (\widehat{\boldsymbol{\Delta}}_1 \otimes \widehat{\boldsymbol{\Delta}}_1) + (\widehat{\boldsymbol{\Delta}}_2 \otimes \widehat{\boldsymbol{\Delta}}_2) \right]^{-1} \mathbf{D}_m \text{vech}(\widehat{\boldsymbol{\Delta}}_1 - \widehat{\boldsymbol{\Delta}}_2). \quad (4.14)$$

Korzystając z Definicji 1.3 macierzy duplikacji, otrzymujemy tezę twierdzenia. \square

Można zauważyć, że aby wyznaczyć statystykę testową Walda w terminologii (1.9a), w Twierdzeniu 4.3 wystarczy wyrazić $\widehat{\boldsymbol{\Delta}}_1$, $\widehat{\boldsymbol{\Delta}}_2$ za pomocą estymatorów $\widehat{\boldsymbol{\Gamma}}_0$ i $\widehat{\boldsymbol{\Gamma}}_1$; por. (1.10). W przeciwieństwie jednak do statystyki RST, gdzie wykorzystanie parametryzacji (1.9a) uprościło postać RST, postać statystyki WT będzie bardziej złożona.

Na koniec zwróćmy uwagę, że przyjmując we wzorze (4.14) estymator największej wiarygodności macierzy kowariancji przy prawdziwości hipotezy zerowej zamiast alternatywnej, otrzymamy statystykę testową przedstawioną przez Tsukada (2018). Należy jednak zauważyć, że takie podejście nie jest zgodne z definicją testu Walda podaną przez Rao (2005).

4.3 Badania symulacyjne

W niniejszym rozdziale za pomocą badań symulacyjnych, przy założeniu prawdziwości hipotezy zerowej w (4.4), wyznaczymy rozkłady empiryczne statystyk RST i WT i porównamy je z dokładnym rozkładem statystyki LRT pod względem zbieżności do granicznego rozkładu chi-kwadrat. Ponadto zweryfikujemy odporność wszystkich rozważanych testów na zaburzenie założenia o normalności rozkładu obserwacji oraz porównamy je pod kątem mocy. Ponieważ badania symulacyjne wymagają generowania danych dla zadanych parametrów, udowodnimy najpierw następujące twierdzenie.

Twierdzenie 4.4. *Przy prawdziwości hipotezy zerowej w (4.4), rozkłady statystyk LRT, RST i WT nie zależą od prawdziwych wartości $\boldsymbol{\mu}$ oraz $\boldsymbol{\Delta}$.*

Dowód. Ponieważ statystyki LRT, RST oraz WT przedstawione odpowiednio w (4.5), Twierdzeniu 4.2 i Twierdzeniu 4.3, nie zawierają $\boldsymbol{\mu}$, to rozkłady tych statystyk nie zależą od prawdziwej wartości $\boldsymbol{\mu}$. Pokażemy teraz, że rozkłady statystyk nie zależą również od $\boldsymbol{\Delta}$.

Macierz obserwacji \mathbf{Y} można zapisać w postaci $\mathbf{Y} = \mathbf{1}_n \boldsymbol{\mu}' + \mathbf{E} \boldsymbol{\Omega}^{1/2}$, gdzie $\boldsymbol{\Omega}^{1/2} = \mathbf{I}_q \otimes \boldsymbol{\Delta}^{1/2}$, $\boldsymbol{\Delta}^{1/2} \boldsymbol{\Delta}^{1/2} = \boldsymbol{\Delta}$ oraz $\mathbf{E} \sim N_{n,mq}(\mathbf{0}, \mathbf{I}_n, \mathbf{I}_{mq})$. Wówczas

$$\mathbf{Y}' \mathbf{Q}_n \mathbf{Y} = \boldsymbol{\Omega}^{1/2} \mathbf{E}' \mathbf{Q}_n \mathbf{E} \boldsymbol{\Omega}^{1/2}.$$

Stąd, dzięki (2.7) i Własności 1.2(iv), otrzymujemy, że

$$\begin{aligned}
\widehat{\Delta} &= \frac{1}{q} \text{BTr}_m(\mathbf{S}) = \frac{1}{nq} \text{BTr}_m(\mathbf{Y}'\mathbf{Q}_n\mathbf{Y}) = \\
&= \frac{1}{nq} \text{BTr}_m(\Omega^{1/2}\mathbf{E}'\mathbf{Q}_n\mathbf{E}\Omega^{1/2}) = \frac{1}{nq} \text{BTr}_m[(\mathbf{I}_q \otimes \Delta^{1/2})\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{I}_q \otimes \Delta^{1/2})] = \\
&= \Delta^{1/2} \frac{1}{nq} \text{BTr}_m[\mathbf{E}'\mathbf{Q}_n\mathbf{E}]\Delta^{1/2} = \\
&= \Delta^{1/2}\widehat{\Upsilon}_0\Delta^{1/2},
\end{aligned}$$

gdzie $\widehat{\Upsilon}_0$ nie zależy od prawdziwych wartości nieznanymi parametrów. Podobnie można wyrazić estymatory $\widehat{\Delta}_1, \widehat{\Delta}_2$ z (2.8). Z idempotentności projektorów \mathbf{Q}_q i \mathbf{P}_q otrzymujemy

$$\begin{aligned}
\widehat{\Delta}_1 &= \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{S}] = \frac{1}{q-1} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{Y}'\mathbf{Q}_n\mathbf{Y}] \\
&= \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\Omega^{1/2}\mathbf{E}'\mathbf{Q}_n\mathbf{E}\Omega^{1/2}] \\
&= \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)(\mathbf{I}_u \otimes \Delta^{1/2})\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{I}_u \otimes \Delta^{1/2})] \\
&= \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{I}_q \otimes \Delta^{1/2})(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{I}_q \otimes \Delta^{1/2})] \\
&= \Delta^{1/2} \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}]\Delta^{1/2} \\
&= \Delta^{1/2} \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{Q}_q \otimes \mathbf{I}_m)]\Delta^{1/2} = \Delta^{1/2}\widehat{\Upsilon}_1\Delta^{1/2},
\end{aligned}$$

oraz

$$\begin{aligned}
\widehat{\Delta}_2 &= \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{S}] = \frac{1}{n} \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{Y}'\mathbf{Q}_n\mathbf{Y}] \\
&= \frac{1}{n} \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\Omega^{1/2}\mathbf{E}'\mathbf{Q}_n\mathbf{E}\Omega^{1/2}] \\
&= \frac{1}{n} \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)(\mathbf{I}_q \otimes \Delta^{1/2})\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{I}_q \otimes \Delta^{1/2})] \\
&= \frac{1}{n} \text{BTr}_m[(\mathbf{I}_u \otimes \Delta^{1/2})(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{I}_u \otimes \Delta^{1/2})] \\
&= \Delta^{1/2} \frac{1}{n} \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}]\Delta^{1/2} \\
&= \Delta^{1/2} \frac{1}{n} \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{P}_q \otimes \mathbf{I}_m)]\Delta^{1/2} = \Delta^{1/2}\widehat{\Upsilon}_2\Delta^{1/2},
\end{aligned}$$

gdzie

$$\begin{aligned}
\widehat{\Upsilon}_1 &= \frac{1}{n(q-1)} \text{BTr}_m[(\mathbf{Q}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{Q}_q \otimes \mathbf{I}_m)], \\
\widehat{\Upsilon}_2 &= \frac{1}{n} \text{BTr}_m[(\mathbf{P}_q \otimes \mathbf{I}_m)\mathbf{E}'\mathbf{Q}_n\mathbf{E}(\mathbf{P}_q \otimes \mathbf{I}_m)]
\end{aligned}$$

nie zależą od prawdziwych wartości nieznanymi parametrów. Stąd estymator $\widehat{\Sigma}_{\text{BCS}}$ zdefiniowany w (4.1) możemy przedstawić jako

$$\widehat{\Sigma}_{\text{BCS}} = (\mathbf{I}_q \otimes \Delta^{1/2}) \left(\mathbf{Q}_q \otimes \widehat{\Upsilon}_1 + \mathbf{P}_q \otimes \widehat{\Upsilon}_2 \right) (\mathbf{I}_q \otimes \Delta^{1/2})$$

i dalej, iloraz wiarygodności przyjmuje postać

$$\Lambda = \left(\frac{|\widehat{\Upsilon}_1|^{q-1} |\widehat{\Upsilon}_2|}{|\widehat{\Upsilon}_0|^q} \right)^{n/2},$$

natomiast statystyki testowe RST i WT wyrażają się jako

$$\text{RST} = \frac{n}{2} \text{Tr} \left\{ \left[\mathbf{I}_{mq} - (\mathbf{Q}_q \otimes \widehat{\Upsilon}_1 + \mathbf{P}_q \otimes \widehat{\Upsilon}_2) (\mathbf{I}_q \otimes \widehat{\Upsilon}_0)^{-1} \right]^2 \right\}$$

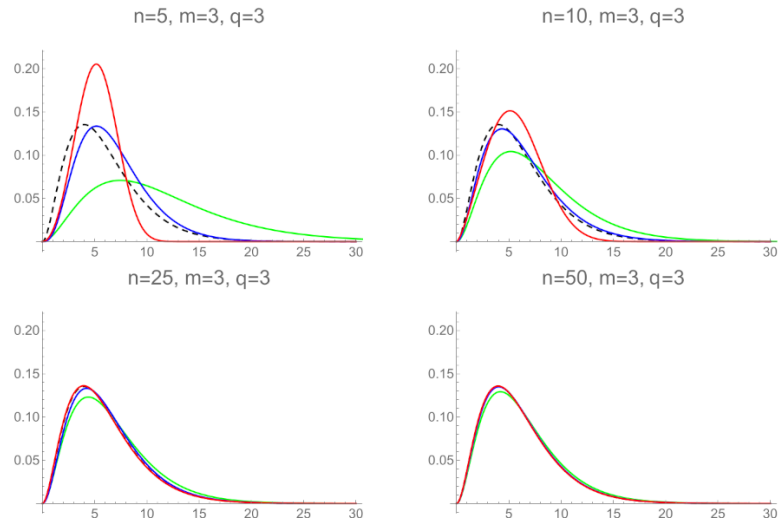
oraz

$$\text{WT} = \frac{n}{2} \text{vec}' \left(\widehat{\Upsilon}_1 - \widehat{\Upsilon}_2 \right) \left[\frac{1}{q-1} (\widehat{\Upsilon}_1 \otimes \widehat{\Upsilon}_1) + (\widehat{\Upsilon}_2 \otimes \widehat{\Upsilon}_2) \right]^{-1} \text{vec} \left(\widehat{\Upsilon}_1 - \widehat{\Upsilon}_2 \right),$$

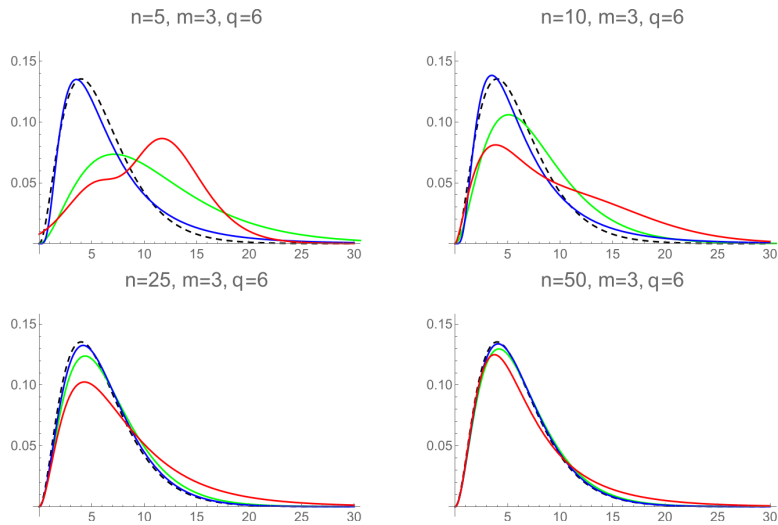
co daje tezę twierdzenia. □

4.3.1 Zbieżność do rozkładu granicznego

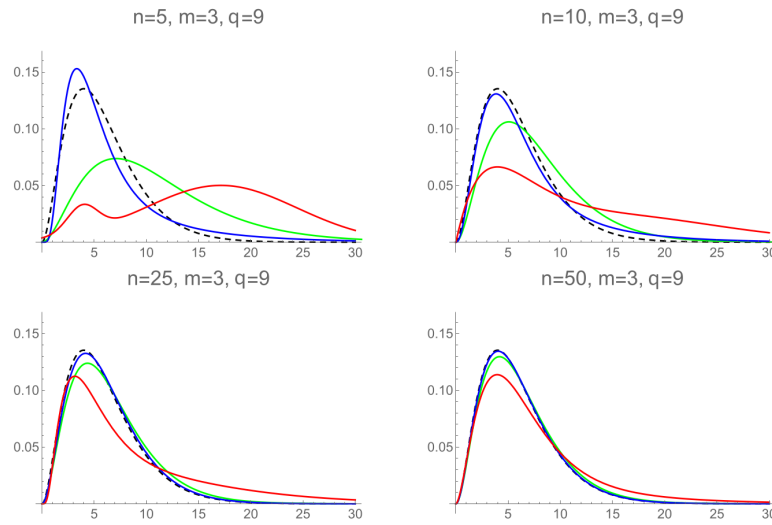
Przypomnijmy, że warunkiem koniecznym przeprowadzenia testów LRT, RST i WT dla hipotezy (4.4) jest aby wielkość próby przekraczała liczbę cech, $n > m$. Ponadto, liczba stopni swobody granicznego rozkładu chi-kwadrat jest równa $m(m+1)/2$. W badaniach symulacyjnych przyjmować będziemy, że $(m, q) \in \{(3, 3), (3, 6), (3, 9), (6, 3), (9, 3)\}$, natomiast $n \in \{10, 25, 50\}$ dla każdej wartości m , oraz dodatkowo $n = 5$ dla $m = 3$ i $n = 7$ dla $m = 6$. Dla każdego zestawu parametrów wygenerowaliśmy losowo 50 000 macierzy obserwacji o rozkładzie $N_{n,mq}(\mathbf{0}, \mathbf{I}_n, \mathbf{I}_{mq})$, dla każdej z nich obliczyliśmy wartości statystyk testowych, i na ich podstawie wyznaczyliśmy empiryczne rozkłady statystyk RST i WT. Rozkład statystyki LRT wyznaczony został z wykorzystaniem pakietu `CharFunToolR` oprogramowania R (por. Gajdoš, 2018). Otrzymane rozkłady, wraz z rozkładem granicznym, zaprezentowane zostały na Rysunkach 8–12.



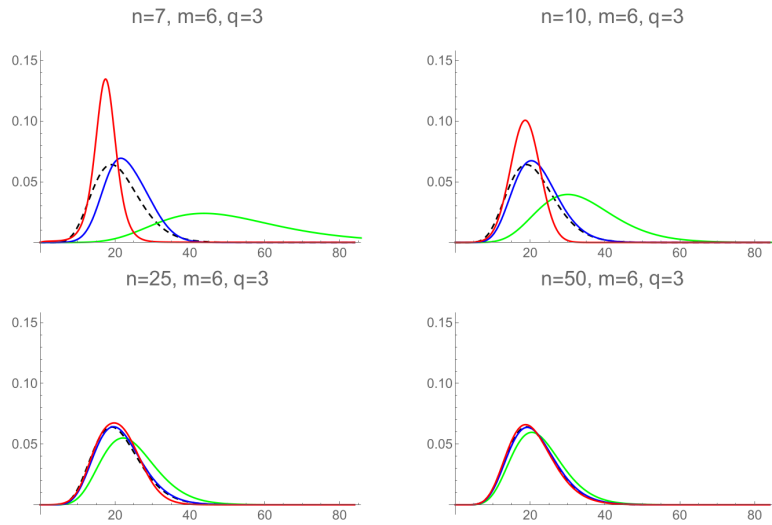
Rysunek 8: Rozkład dokładny LRT (zielony), rozkłady empiryczne RST (niebieski) i WT (czerwony) przy prawdziwości hipotezy zerowej wraz z rozkładem χ_6^2 (czarny przerywany) dla $m = q = 3$.



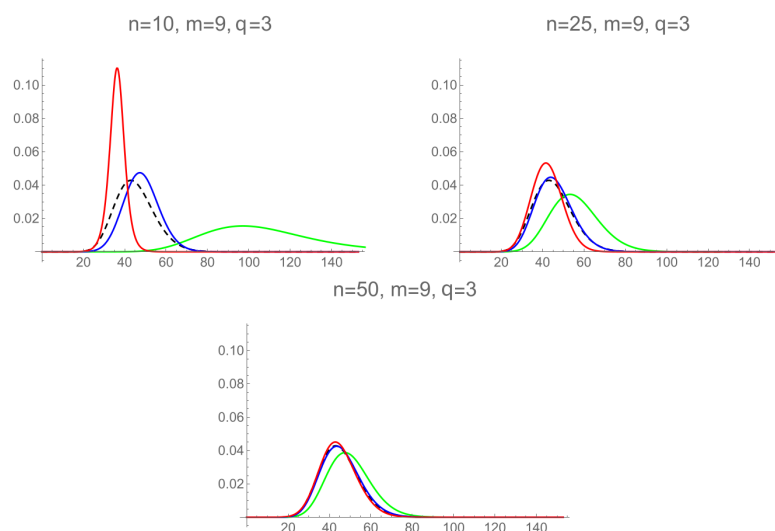
Rysunek 9: Rozkład dokładny LRT (zielony), rozkłady empiryczne RST (niebieski) i WT (czerwony) przy prawdziwości hipotezy zerowej wraz z rozkładem χ_6^2 (czarny przerywany) dla $m = 3, q = 6$.



Rysunek 10: Rozkład dokładny LRT (zielony), rozkłady empiryczne RST (niebieski) i WT (czerwony) przy prawdziwości hipotezy zerowej wraz z rozkładem χ_6^2 (czarny przerywany) dla $m = 3, q = 9$.



Rysunek 11: Rozkład dokładny LRT (zielony), rozkłady empiryczne RST (niebieski) i WT (czerwony) przy prawdziwości hipotezy zerowej wraz z rozkładem χ_{21}^2 (czarny przerywany) dla $m = 6, q = 3$.



Rysunek 12: Rozkład dokładny LRT (zielony), rozkłady empiryczne RST (niebieski) i WT (czerwony) przy prawdziwości hipotezy zerowej wraz z rozkładem χ_{45}^2 (czarny przerywany) dla $m = 9, q = 3$.

Wykresy te potwierdzają, że rozkłady wszystkich statystyk testowych dążą do rozkładu granicznego wraz ze wzrostem n , przy czym zbieżność rozkładu RST jest najszybsza i nawet dla stosunkowo małej liczebności próby, w przeciwieństwie do rozkładu statystyk LRT i WT, nie różni się istotnie od rozkładu granicznego. Można również zauważyć, że poza przypadkiem $m = 3, q \in \{6, 9\}$, dla małych wielkości próby test LRT zazwyczaj przeszacowuje wartości kwantyli rozkładu granicznego, natomiast test WT ich nie doszacowuje.

4.3.2 Odporność testów na zaburzenia rozkładu normalnego

Sprawdźmy teraz odporność rozważanych testów na pewne zaburzenia rozkładu normalnego obserwacji. W tym celu wygenerujemy dane z następujących rozkładów: wielowymiarowy rozkład t studenta z trzema lub pięcioma stopniami swobody (t_3, t_5), rozkład gamma z parametrami $(2, 1.5)$, a także jednostajny na przedziale $(0, 1)$. Przyjmiemy ponadto, że $(m, q) = (9, 3)$ oraz $n = \{10, 25\}$. Dla każdego zestawu

parametrów (m, q, n) wygenerowanych zostanie losowo 50 000 macierzy obserwacji o wspomnianych rozkładach, dla każdej z nich obliczymy wartość statystyk testowych, wyznaczymy empiryczne rozkłady statystyk LRT, RST i WT, FT oraz RLRT. Rozkłady statystyk wraz z odpowiadającymi rozkładami przy założeniu normalności przedstawimy na Rysunku 13.

Można zauważyć, że przy prawdziwości hipotezy zerowej empiryczne rozkłady wszystkich statystyk testowych wyznaczone na podstawie danych wygenerowanych ze wspomnianych rozkładów są dość zbliżone do swoich odpowiedników przy założeniu normalności rozkładu obserwacji. Podobnie, dla rosnącej wielkości próby i danych generowanych z rozkładu jednostajnego i gamma, nie ma istotnych różnic między empirycznymi rozkładami każdej z rozważanych statystyk testowych. Zauważmy jednak, że dla $n = 25$ i rozkładów generowanych z wielowymiarowych rozkładów t zaobserwować można istotne różnice między rozkładami statystyk testowych. Możemy zatem uznać, że wszystkie rozważane statystyki nie są odporne na zaburzenie normalności rozkładu obserwacji.

4.3.3 Analiza mocy testów

Do analizy funkcji mocy rozważanych testów w zależności od wielkości rozbieżności pomiędzy hipotezą zerową, a alternatywną w (4.4), wykorzystaliśmy miarę rozbieżności Kullbacka-Leiblera; por. (1.16). Dla rozważanej hipotezy i ustalonej macierzy Σ_{BCS} rozbieżność między hipotezami wyznaczymy jako

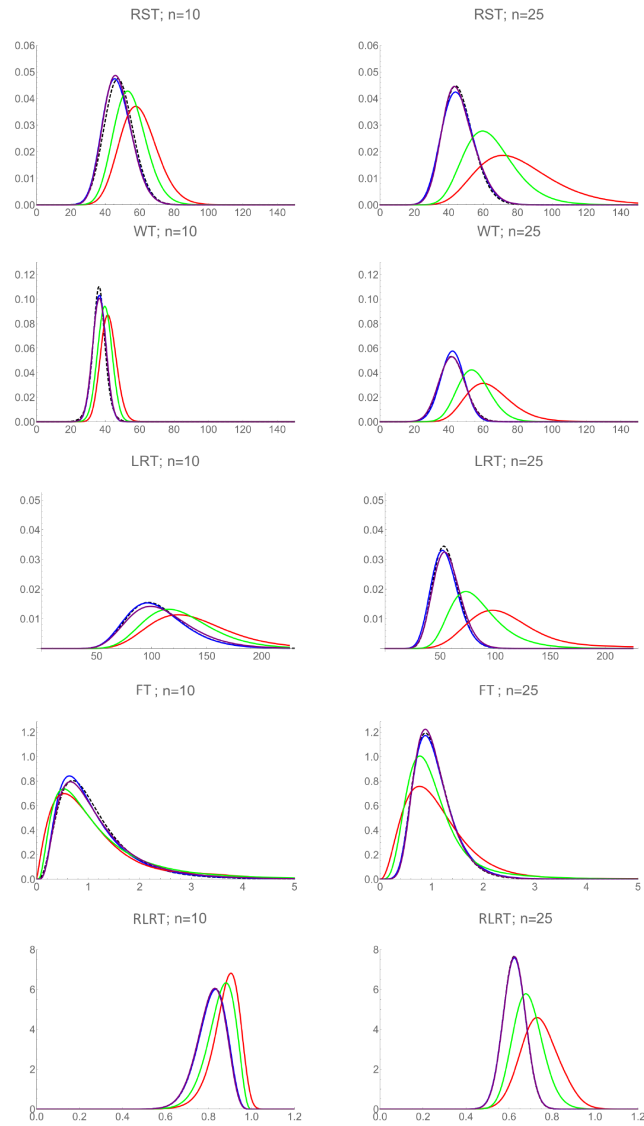
$$\xi(\tilde{\Delta}) = \min_{\Delta} f_E(\Delta, \Sigma_{\text{BCS}}),$$

gdzie

$$f_E(\Delta, \Sigma_{\text{BCS}}) = \text{Tr} \left[\Sigma_{\text{BCS}}^{-1} (\mathbf{I}_q \otimes \Delta) \right] - \ln |\Sigma_{\text{BCS}}^{-1}| - \ln |\mathbf{I}_q \otimes \Delta| - mq.$$

Różniczkując $f_E(\Delta, \Sigma_{\text{BCS}})$ ze względu na Δ , otrzymujemy

$$\frac{df_E(\Delta, \Sigma_{\text{BCS}})}{d\Delta} = \text{vec}' \left(\Sigma_{\text{BCS}}^{-1} - \mathbf{I}_q \otimes \Delta^{-1} \right) (\mathbf{I}_q \otimes \mathbf{K}_{m,q} \otimes \mathbf{I}_m) (\text{vec } \mathbf{I}_q \otimes \mathbf{I}_{m^2}) \mathbf{D}_m.$$



Rysunek 13: Rozkłady empiryczne statystyk testowych przy prawdziwości hipotezy zerowej (zaprezentowane w wierszach) dla danych o rozkładzie: normalnym (czarny przerywany), wielowymiarowym t_3 (czerwony), t_5 (zielony), gamma $G(2, 1.5)$ (niebieski) oraz jednostajnym $U(0, 1)$ (fioletowy), dla $m = 9$, $q = 3$, $n = 10$ (lewa kolumna), $n = 25$ (prawa kolumna).

Przyrównując powyższe do zera i wykorzystując Własności 1.2(i), (ii) oraz kwadratość podprzestrzeni struktur BCS, mamy

$$\begin{aligned} \text{vec}' \left[\text{BTr}_m \left(\Sigma_{\text{BCS}}^{-1} \right) \right] &= \text{vec}' \left[\text{BTr}_m \left(\mathbf{I}_q \otimes \Delta^{-1} \right) \right] \\ \text{BTr}_m \left(\Sigma_{\text{BCS}}^{-1} \right) &= \text{BTr}_m \left(\mathbf{I}_q \otimes \Delta^{-1} \right) \\ \left[\text{Tr}(\mathbf{Q}_q) \cdot \Delta_1^{-1} + \text{Tr}(\mathbf{P}_q) \cdot \Delta_2^{-1} \right] &= \text{Tr}(\mathbf{I}_q) \cdot \Delta^{-1} \\ \frac{1}{q} \left[(q-1) \Delta_1^{-1} + \Delta_2^{-1} \right] &= \Delta^{-1}. \end{aligned}$$

Stąd minimum $f_E(\Delta, \Sigma_{\text{BCS}})$ osiągnane jest dla

$$\tilde{\Delta} = \left(\frac{q-1}{q} \Delta_1^{-1} + \frac{1}{q} \Delta_2^{-1} \right)^{-1}. \quad (4.15)$$

Należy zauważyć, że (4.15) określa strukturę diagonalną bloku, która jest najbliższa w sensie (1.16) i nie musi być taka sama jak bloki diagonalne w hipotezie alternatywnej. Przypomnijmy również, że ponieważ wartość $\xi(\tilde{\Delta})$ nie jest ograniczona z góry, do analizy mocy testów została wykorzystana skorygowana rozbieżność $\eta(\tilde{\Delta}) = 1 - \frac{1}{1+\xi(\tilde{\Delta})}$; por. Rozdział 1.3.

Do wyznaczenia rozbieżności między hipotezami wykorzystaliśmy parametryzację (1.9a) oraz zależności (1.10). Przykładowo, dla $m = q = 3$ wygenerowaliśmy macierze

$$\mathbf{\Gamma}_0 = \begin{pmatrix} 88.910 & -13.002 & 14.855 \\ -13.002 & 84.921 & 5.285 \\ 14.855 & 5.285 & 120.934 \end{pmatrix}, \quad \mathbf{\Gamma}_1 = \begin{pmatrix} 26.195 & -0.231 & -4.579 \\ -0.231 & 2.357 & -1.647 \\ -4.579 & -1.647 & 3.495 \end{pmatrix}, \quad (4.16)$$

dla których minimum funkcji rozbieżności jest osiągnięte dla

$$\tilde{\Delta} = \begin{pmatrix} 75.995 & -13.2497 & 17.5422 \\ -13.2497 & 84.7433 & 5.49956 \\ 17.5422 & 5.49956 & 120.196 \end{pmatrix}$$

dając wartość $\eta = 0.2012$.

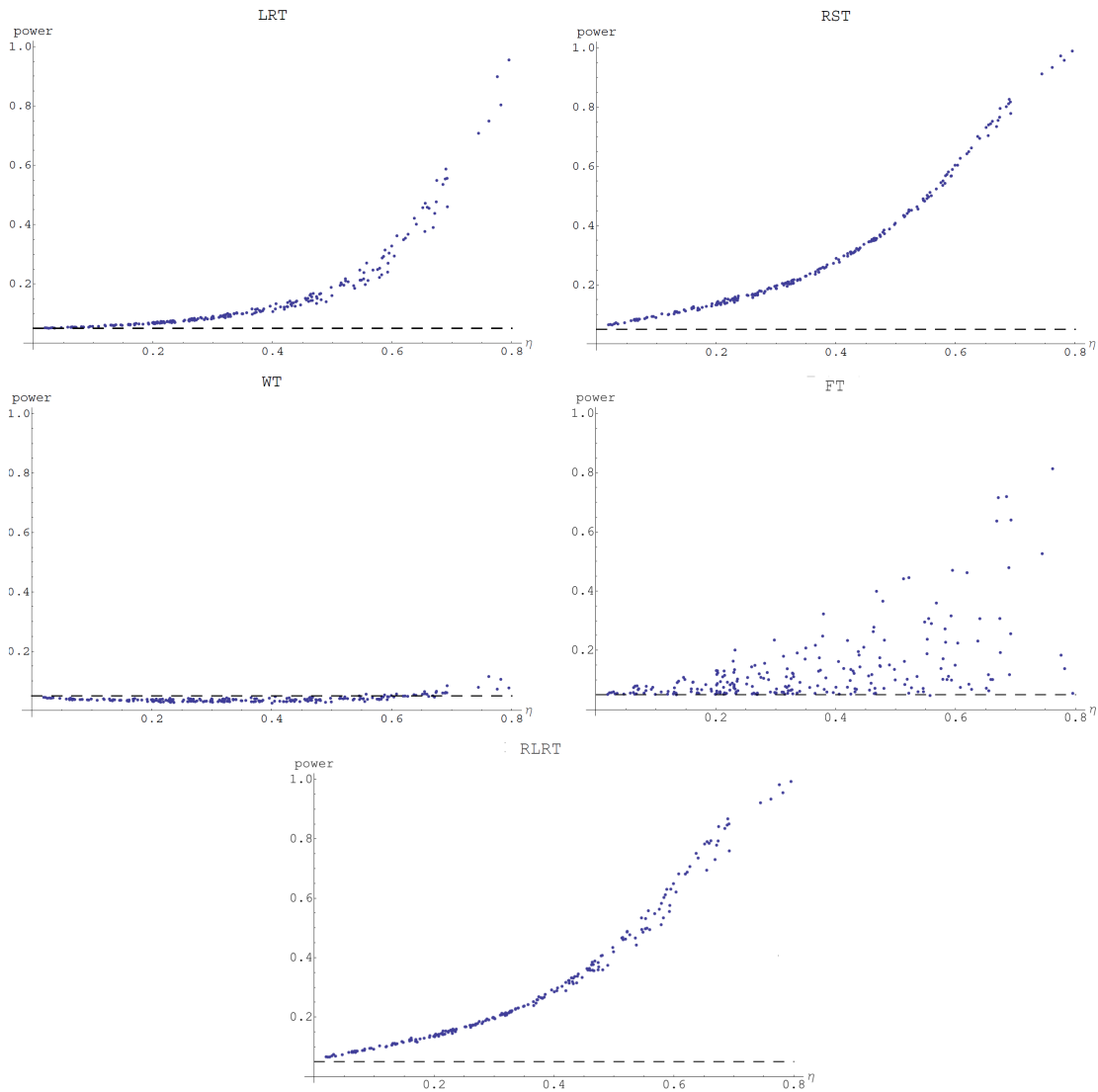
W pierwszym etapie badań symulacyjnych dla parametrów $m = q = 3$, $n = 5$ i macierzy $\mathbf{\Gamma}_0$ podanej w (4.16) wygenerowaliśmy losowo 300 macierzy $\mathbf{\Gamma}_1$ i sprawdziliśmy, że w 237 przypadkach otrzymana macierzy Σ_{BCS} jest określona dodatnio. Następnie dla każdej określonej dodatnio macierzy Σ_{BCS} obliczyliśmy rozbieżność $\eta(\tilde{\Delta})$

oraz wygenerowaliśmy 50 000 macierzy $\mathbf{Y} \sim N_{n,mq}(\mathbf{0}, \mathbf{I}_n, \boldsymbol{\Sigma}_{\text{BCS}})$, dla których wyznaczyliśmy wartości wszystkich statystyk testowych. Do odrzucenia hipotezy zerowej zostały użyte odpowiednio rozkład dokładny LRT (przedstawiony w Twierdzeniu 4.1 i obliczony za pomocą pakietu `CharFunToolR`), rozkłady empiryczne przy prawdziwości hipotezy zerowej RST i WT, rozkład F z $n - 1$ i $(n - 1)(q - 1)$ stopniami swobody dla FT i rozkład RLR z m , $(n - 1)(q - 1)$ i $n - 1$ stopniami swobody (wyznaczone za pomocą algorytmu zaproponowanego w pracy Chiani, 2016) dla RLRT. Podobnie jak w Fonseca i in. (2018) oraz Kozioł i in. (2021), do obliczenia FT wybrano $\mathbf{v} = \mathbf{1}_m$. We wszystkich obliczeniach przyjęliśmy poziom istotności równy 0.05. Moce testów, wyznaczone jako stosunek odrzuconych hipotez zerowych do liczby wszystkich macierzy obserwacji, 50 000, w zależności od skorygowanej rozbieżności $\eta(\widetilde{\Delta})$ przedstawiono na Rysunku 14.

Widocznym jest, że wartości mocy LRT, RST i RLRT rosną wraz ze wzrostem wartości rozbieżności, w przeciwieństwie do FT, dla którego dla dwóch równorzędnych struktur BCS moc różni się istotnie, oraz WT, dla którego moc jest często poniżej nominalnego poziomu istotności (test Walda jest testem obciążonym). Z tego powodu, do dalszego porównania mocy testów wzięliśmy pod uwagę tylko LRT, RST i RLRT.

Dla każdej pary $(m, q) \in \{(3, 3), (3, 6), (3, 9), (6, 3), (6, 6), (6, 9), (9, 3), (9, 6)\}$ wybraliśmy macierze $\boldsymbol{\Gamma}_0$ i $\boldsymbol{\Gamma}_1$ w taki sposób, aby wartość skorygowanej rozbieżności η wynosiła 0.2 oraz 0.4. Następnie, podobnie jak poprzednio, dla wybranych liczebności próby zapewniających istnienie wszystkich trzech testów ($n > m$) wygenerowaliśmy 50 000 macierzy obserwacji, dla których obliczyliśmy moce empiryczne testów LRT, RST oraz RLRT. Wykresy przedstawiające zależność mocy empirycznych kolejnych testów od wielkości próby przedstawiono na Rysunkach 15–17.

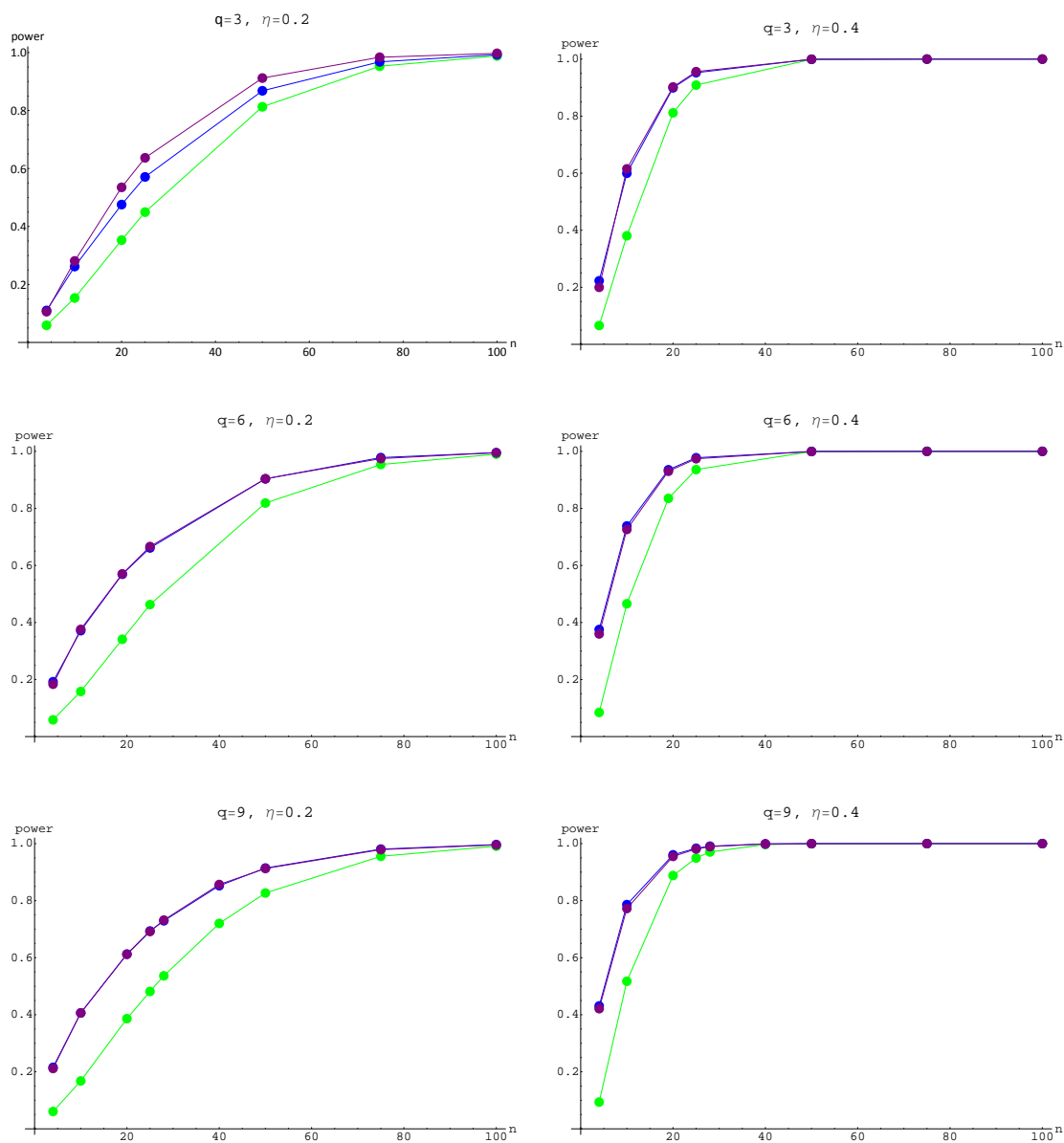
Można zauważyć, że zarówno dla $\eta = 0.2$ jak i $\eta = 0.4$ moce RST (linia niebieska) i RLRT (linia fioletowa) przewyższają moc LRT (linia zielona) dla każdej rozpatrywanej wielkości próby, jednak różnice te nie są znaczące. Co więcej, dla $\eta = 0.4$ nie widać istotnych różnic między mocą RST i RLRT.



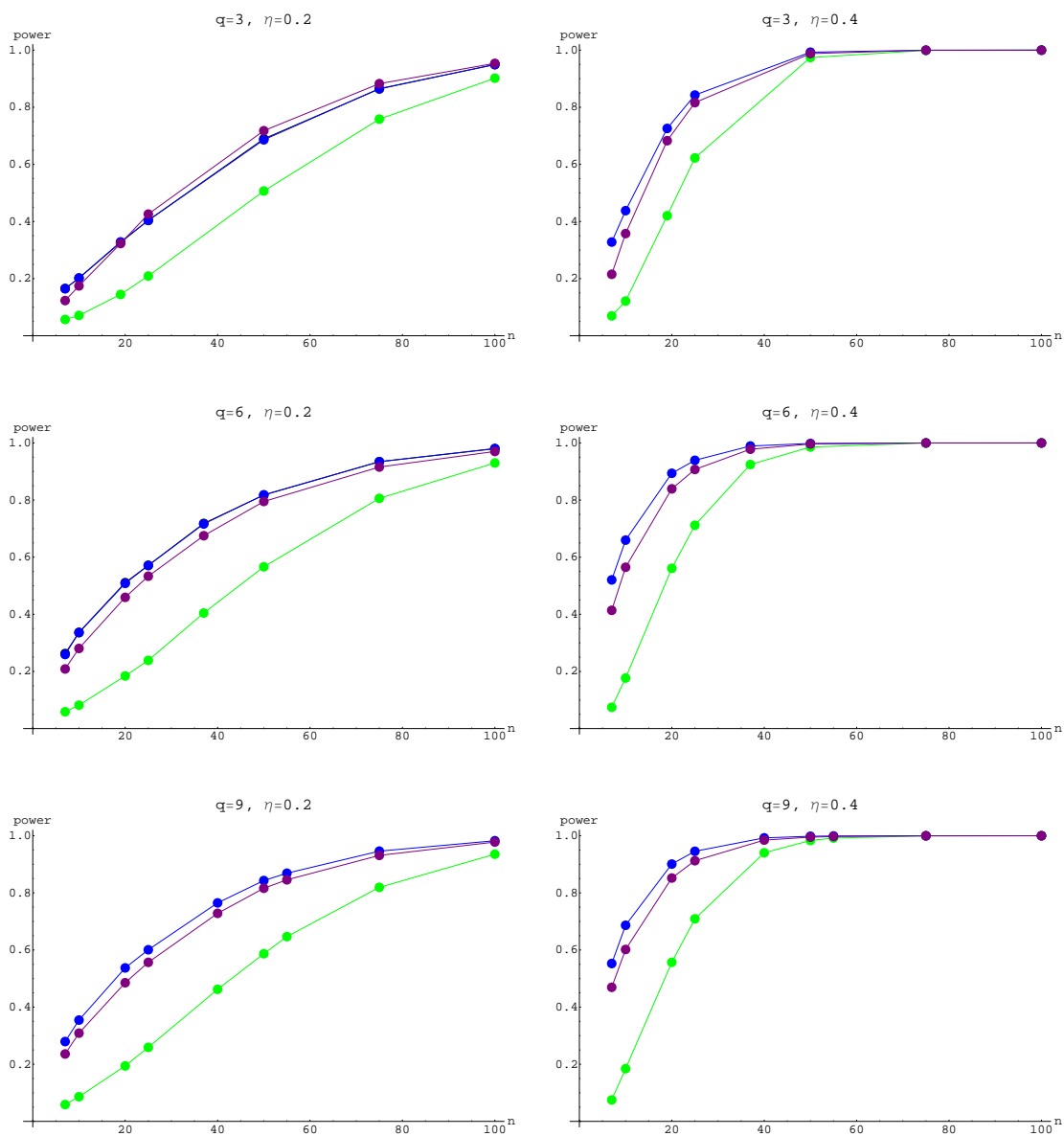
Rysunek 14: Moce empiryczne testów LRT, RST, WT, FT oraz RLRT dla $m = q = 3$, $n = 5$, w zależności od wartości rozbieżności η .

4.3.4 Przykład rzeczywisty

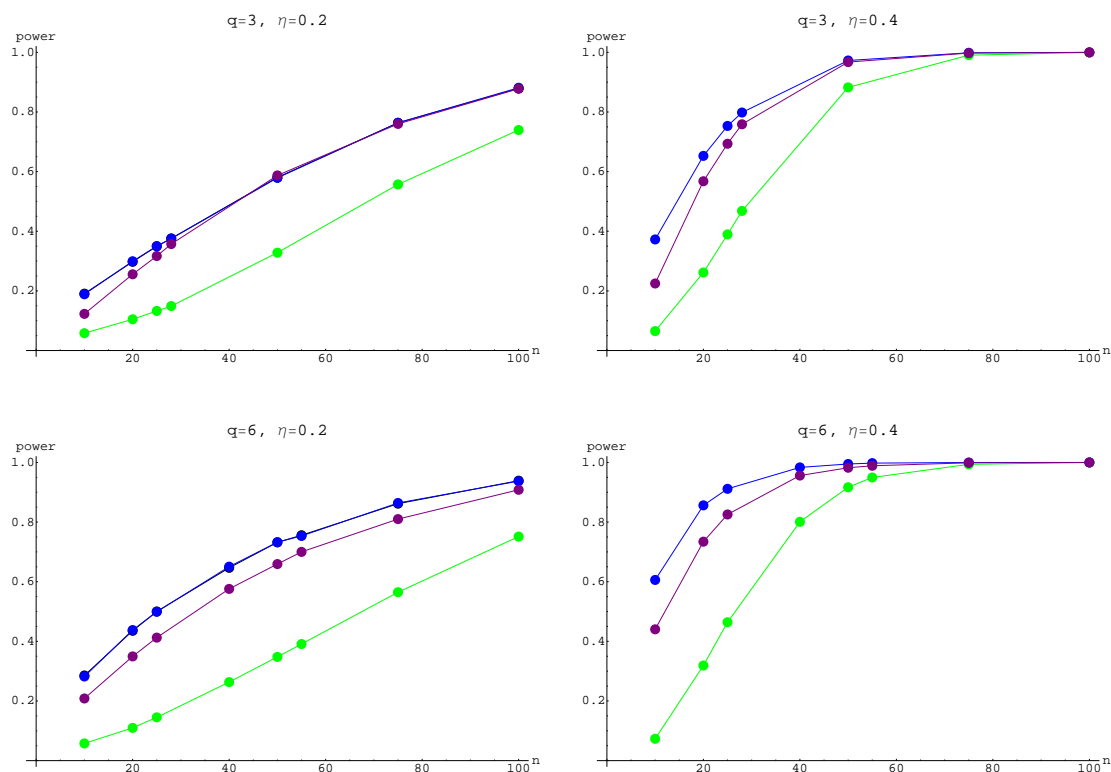
W niniejszym rozdziale zastosujemy omówione testy do weryfikacji hipotezy (4.4) dla rzeczywistych danych opisanych w Rozdziale 1.5. Zauważmy, że hipoteza zerowa dotyczy niezależności długości płatków pomiędzy dowolnymi dwoma kwia-



Rysunek 15: Moce empiryczne testów: LRT (zielony), RST (niebieski) oraz RLRT (fioletowy) w zależności od n dla $m = 3$, $q \in \{3, 6, 9\}$ oraz dla $\eta = 0.2$ (lewa kolumna) i $\eta = 0.4$ (prawa kolumna).



Rysunek 16: Moce empiryczne testów: LRT (zielony), RST (niebieski) oraz RLRT (fioletowy) w zależności od n dla $m = 6, q \in \{3, 6, 9\}$ oraz dla $\eta = 0.2$ (lewa kolumna) i $\eta = 0.4$ (prawa kolumna).



Rysunek 17: Moce empiryczne testów: LRT (zielony), RST (niebieski) oraz RLRT (fioletowy) w zależności od n dla $m = 9$, $q \in \{3, 6\}$ oraz dla $\eta = 0.2$ (lewa kolumna) i $\eta = 0.4$ (prawa kolumna).

tami. W przypadku testu LRT do podjęcia decyzji wykorzystamy kwantyle rozkładu dokładnego prezentowanego w Twierdzeniu 4.1 (obliczone przy użyciu pakietu CharFunToolR) oraz kwantyle granicznego rozkładu χ_{10}^2 , natomiast dla statystyk RST i WT wykorzystamy kwantyle empiryczne rozkładów statystyk RST i WT oraz rozkładu granicznego χ_{10}^2 . Do zilustrowania działania testu F użyjemy trzech postaci wektora \mathbf{v} : $\mathbf{1}_m$, \mathbf{v}_{\max} będącego wektorem własnym odpowiadającym największej wartości własnej $\widehat{\Delta}_2 \widehat{\Delta}_1^{-1}$, a także wektora wygenerowanego losowo $\mathbf{v}_g = (0.859853, 0.175291, 0.011513, 0.405039)'$, a do podjęcia decyzji wykorzystamy kwantyle rozkładu F ze stopniami swobody 10, 20. Przeprowadzimy również test RLRT z wykorzystaniem kwantyli rozkładu RLR ze stopniami swobody 4, 20, 10

(z użyciem algorytmu przedstawionego w pracy Chiani, 2016). Wartości statystyk testowych wraz z odpowiadającymi im p -wartościami podane są w Tabeli 4.

	LRT	RST	WT	FT($\mathbf{1}_m$)	FT(\mathbf{v}_{\max})	FT(\mathbf{v}_g)	RLRT
statystyki testowe	11.532	5.974	8.600	0.700	0.978	0.144	0.601
p -wartość - dokładny rozkład	0.565	0.879	0.569	0.714	0.491	0.998	0.402
p -wartość - graniczny rozkład χ^2	0.318	0.817	0.570	—	—	—	—

Tabela 4: Wartości statystyki LRT, RST, WT, FT($\mathbf{1}_m$), FT(\mathbf{v}_{\max}), FT(\mathbf{v}_g) oraz RLRT wraz z odpowiednimi p -wartościami.

Na poziomie istotności 0.05 p -wartości obliczone z odpowiednich rozkładów wszystkich statystyk testowych nie odrzucają hipotezy zerowej o niezależności długości płatków dla różnych kwiatów. Ponadto decyzja podjęta na podstawie rozkładu granicznego χ^2 dla LRT, RST i WT pozostaje taka sama, przy czym w przypadku RST i WT empiryczne p -wartości i p -wartości obliczone na podstawie rozkładu granicznego są zbliżone, co nie ma miejsca w przypadku LRT. Fakt ten nie jest zaskakujący, gdyż na podstawie badań symulacyjnych (Rysunki 8–12) mogliśmy zaobserwować wolniejszą zbieżność rozkładu statystyki LRT do rozkładu granicznego dla małej wielkości próby.

W przypadku testu F można zauważyć, że wartość statystyki FT zależy od wyboru wektora \mathbf{v} i stąd różne p -wartości. W rozważanym przykładzie decyzja pozostaje taka sama, jednak nietrudno sobie wyobrazić, że przy (nie)odpowiednim doborze wektora \mathbf{v} p -wartość nie będzie przekraczać zadanego poziomu istotności. Na koniec przypomnijmy, że ponieważ statystyka Roya jest transformacją statystyki F wyzna-

czonej dla $\mathbf{v} = \mathbf{v}_{\max}$, stąd podobne p -wartości dla obu testów.

Na koniec zauważmy, że przeprowadzenie testu LRT bazującego na rozkładzie dokładnym przedstawionym w Twierdzeniu 4.1 oraz przeprowadzenie testu RLRT wymagają wykorzystania specjalistycznego oprogramowania (pakiet `CharFunTool` dla LRT czy algorytm wyznaczający p -wartości na podstawie rozkładu RLR, jako że dostępne w literaturze tablice zawierają kwantyle tylko dla dość szczególnych wartości parametrów eksperymentu), co może być dość uciążliwe dla praktyków. Z drugiej strony, ponieważ empiryczne rozkłady RST i WT nie różnią się istotnie od granicznych wartości rozkładu χ^2 , kwantyle rozkładu granicznego mogą być z powodzeniem wykorzystywane przez badaczy. Zwróćmy jednak uwagę, że ze względu na obciążoność testu WT, błąd drugiego rodzaju może być znacznie wyższy niż możliwy do zaakceptowania, dlatego najbardziej polecanym testem wydaje się RST.

5 Podsumowanie

W rozprawie doktorskiej analizowaliśmy wielowymiarowe i podwójnie wielowymiarowe modele normalne z macierzą kowariancji o zadanej strukturze liniowej.

W Rozdziale 2 pokazaliśmy, że w modelu z nieznaną wartością oczekiwaną, w którym przestrzeń wartości oczekiwanej komutuje z przestrzenią macierzy kowariancji, warunkiem równoważnym do istnienia jawnej postaci estymatora największej wiarygodności macierzy kowariancji jest fakt, iż przestrzeń macierzy kowariancji jest podprzestrzenią kwadratową. Ponadto pokazaliśmy, że estymator ten jest rzutem ortogonalnym estymatora największej wiarygodności macierzy kowariancji bez struktury na przestrzeń macierzy o zadanej strukturze. Przedstawione wyniki zilustrowaliśmy przykładami wybranych struktur kowariancyjnych w standardowym modelu wielowymiarowym, w modelu podwójnie wielowymiarowym oraz w modelu krzywych wzrostu.

W przypadku, gdy liczba rozważanych cech zwiększa się do nieskończoności lub też gdy jest większa od liczebności próby, ocena macierzy kowariancji nieposiadającej żadnej struktury, niezbędna do estymacji struktury kowariancyjnej, może być źle uwarunkowana lub osobliwa. W takiej sytuacji można poprawić własności estymatorów stosując przykładowo metody kurczenia (ang. *shrinkage methods*). W przypadku macierzy kowariancji o zadanej strukturze metody te, po odpowiedniej modyfikacji, mogą być również zastosowane, jednak będzie to przedmiotem dalszych badań.

W sytuacji, gdy macierz kowariancji o zadanej strukturze nie należy do podprzestrzeni kwadratowej (na przykład wstępowe macierze Toeplitza), estymator największej wiarygodności nie jest wyznaczany za pomocą rzutowania (rzutowanie nie zapewnia określoności macierzy), wyznacza się go numerycznie; por. Markiewicz i Mieldzioc (2023). Można również proponować inne metody estymacji, na przykład rzutowanie na stożek macierzy określonych nieujemnie o zadanej strukturze; por. Filipiak i in. (2018b) lub metodę kurczenia; por. John i Mieldzioc (2020). Wyznaczanie estymatorów macierzy kowariancji nienależących do podprzestrzeni kwadratowych, o dobrych własnościach statystycznych, będzie przedmiotem dalszych badań.

W Rozdziale 3 rozprawy zaproponowaliśmy testy hipotez związanych ze strukturami kowariancyjnymi należącymi do podprzestrzeni kwadratowych w modelu podwójnie wielowymiarowym. Wyprowadziliśmy test ilorazu wiarygodności i test wynikowy Rao dotyczący zarówno struktur międzyblokowych jak i wewnątrzblokowych. Z przeprowadzonych badań symulacyjnych wynika, że zbieżność rozkładu statystyki wynikowej Rao do granicznego rozkładu chi-kwadrat następuje już dla niewielkiej liczebności próby, czego nie można powiedzieć o rozkładzie statystyki ilorazu wiarygodności. Porównując natomiast moce obu testów ze względu na miarę rozbieżności możemy powiedzieć, że oba są zbliżone. Zaletą testu ilorazu wiarygodności jest znajomość rozkładu dokładnego statystyki testowej, jednak jest to przewaga tylko pozorna, gdyż jego wyznaczenie wymaga użycia specjalistycznego oprogramowania, podczas gdy w teście wynikowym Rao wykorzystać można kwantyle rozkładu granicznego już dla niewielkiej wielkości próby. W niniejszej pracy nie zbadaliśmy odporności proponowanych testów na zaburzenie normalności rozkładu obserwacji. Będzie to przedmiotem dalszych badań.

Warto zwrócić uwagę, że hipoteza rozważana w Rozdziale 3 nie obejmowała testowania międzyblokowej niezależności (struktura blokowo diagonalna macierzy kowariancji) przy założeniu blokowej struktury kompletnej symetrii (hipoteza ta nie spełnia warunku o jednakowej przestrzeni zależności międzyblokowej w hipotezie zerowej i alternatywnej). Zagadnienie to zostało rozważone w Rozdziale 4 rozprawy, gdzie oprócz testu ilorazu wiarygodności i testu wynikowego Rao zaproponowaliśmy także test Walda oraz porównaliśmy ich własności z testami istniejącymi w literaturze. Podobnie jak w poprzednim rozdziale, symulacje pokazały, że rozkład statystyki wynikowej Rao dąży do rozkładu granicznego już dla niewielkiego rozmiaru próby, a jego moc jest jedną z wyższych wśród wszystkich rozważanych testów, co wydaje się wystarczająco rekompensować nieznaną jej dokładnego rozkładu. Ponadto okazało się, że żaden z rozważanych testów nie jest odporny na zaburzenie normalności macierzy obserwacji.

Przyszłe badania dotyczące testowania struktur kowariancyjnych należących do podprzestrzeni kwadratowych związane będą z rozszerzeniem wyników na modele, w których normalność rozkładu obserwacji nie jest wymagana, oraz na modele, w których liczba cech dąży do nieskończoności (zagadnienie wysokiej-wymiarowości). Interesujące będzie również testowanie hipotez o jednoczesnej strukturze macierzy kowariancji oraz wartości oczekiwanej. Wstępne wyniki związane z ostatnim zagadnieniem znaleźć można w Liang i in. (2021).

Bibliografia

- T.W. Anderson, *An Introduction to Multivariate Statistical Analysis* (3rd Ed.), John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.
- S. F. Arnold, Application of the theory of products of problems to certain patterned covariance matrices, *Ann. Statist.*, 1(4) (1973) 682–699.
- S. F. Arnold, Linear models with exchangeably distributed errors, *J. Amer. Statist. Assoc.*, 74 (1979) 194–199.
- A. Basilevsky, *Applied matrix algebra in the statistical sciences*, North-Holland, New York, 1983.
- P. Billingsley, *Probability and measure* (3rd ed.), John Wiley & Sons, Inc., New York, 1995.
- M. Chiani, Distribution of the largest root of a matrix for Roy’s test in multivariate analysis of variance, *J. Multivariate Anal.*, 143 (2016) 467–471.
- M. L. Eaton, *Multivariate statistics: a vector space approach*, John Wiley & Sons, Inc., New York, 1983.
- P. L. Fackler, Notes on matrix calculus,
<https://media.gradebuddy.com/documents/1897145/1ad5e235-824d-4bbf-81d6-ffd2040e37ec.pdf> (2005).

- K. Filipiak, M. John, D. Klein, Testing independence under a block compound symmetry covariance structure, *Statist. Papers*, 64(2) (2023) 677–704.
- K. Filipiak, M. John, Y. Liang, Testing covariance structures belonging to a quadratic subspace under a doubly multivariate model, *TEST*, 2024. Przyjęta do druku.
- K. Filipiak, M. John, A. Markiewicz, Comments on maximum likelihood estimation and projections under multivariate statistical models, In: T. Holgersson, M. Singull (Eds.), *Recent Developments in Multivariate and Random Matrix Analysis*, 51-66, Springer (2020).
- K. Filipiak, D. Klein, Estimation and testing of the covariance structure of doubly multivariate data, In: K. Filipiak, A. Markiewicz, D. von Rosen (Eds.), *Multivariate, Multilinear and Mixed Linear Models*, 131–155, Springer (2021).
- K. Filipiak, D. Klein, M. Mokrzycka, Separable covariance structure identification for doubly multivariate data. In: K. Filipiak, A. Markiewicz, D von Rosen (Eds.), *Multivariate, Multilinear and Mixed Linear Models*, 113–130. Springer (2021).
- K. Filipiak, D. Klein, M. Mokrzycka, Discrepancy between structured matrices in the context of power analysis, *Comp. Stat. Data Anal.* 192 (2024), 107907.
- K. Filipiak, D. Klein, A. Roy, Score test for a separable covariance structure with the first component as compound symmetric correlation matrix, *J. Multivariate Anal.*, 150 (2016) 105–124.
- K. Filipiak, D. Klein, E. Vojtková, The properties of partial trace and block trace operators of partitioned matrices, *Electron. J. Linear Algebra*, 33 (2018a) 3–15.
- K. Filipiak, A. Markiewicz, A. Mieldzioc, A. Sawikowska, On projection of a positive definite matrix on a cone of nonnegative definite Toeplitz matrices, *Electron. J. Linear Algebra*, 33 (2018b), 74–82.

- M. Fonseca, A. Koziół, R. Zmysłony, Testing hypotheses of covariance structure in multivariate data, *Electron. J. Linear Algebra*, 33 (2018) 53–62.
- A. Gajdoš, CharFunToolR: the characteristic functions toolbox (R), Available from: <https://github.com/gajdosandrej/CharFunToolR> (2018).
- S. Geisser, Multivariate analysis of variance for a special covariance case, *J. Amer. Statist. Assoc.*, 58 (1963) 660–669.
- M. A. Girshick, The distribution of the ellipticity statistic L_{η} when the hypothesis is false, *Terr. Magn. Atmos. Elect.*, 46 (1941) 455–457.
- L. Gleser, A note on the sphericity test, *Ann. Math. Statist.*, 37 (1966) 464–467.
- C. Hao, Y. Liang, A. Roy, Equivalency between vertices and centers-coupled-with-radii principal component analyses for interval data, *Statist. Probab. Lett.*, 106 (2015) 113–120.
- M. John, A. Mieldzioc, The comparison of the estimators of banded Toeplitz covariance structure under the high-dimensional multivariate model, *Comm. Statist. Simulation Comput.*, 49(3), 734–752.
- V. Kopčová, I. Žežula, On intraclass structure estimation in the growth curve model, *Statist. Papers*, 61 (2020) 1085–1106.
- T. Kollo, D. von Rosen, *Advanced Multivariate Statistics with Matrices*, Dordrecht, Springer, 2005.
- A. Koziół, A. Roy, R. Zmysłony, R. Leiva, M. Fonseca, Best unbiased estimates for parameters of three-level multivariate data with doubly exchangeable covariance structure, *Linear Algebra Appl.*, 535 (2017) 87–104.
- A. Koziół, A. Roy, R. Zmysłony, I. Žežula, M. Fonseca, Estimation and testing hypotheses in two-level and three-level multivariate data with block compound sym-

- metric covariance structure, in: K. Filipiak, A. Markiewicz, D. von Rosen (Eds.), *Multivariate, Multilinear and Mixed Linear Models*, 203–232, Springer (2021).
- P. R. Krishnaiah, J. C. Lee, On Covariance Structures. *Sankhya*, 38 (1976) 357–371.
- O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.*, 88 (2004), 365–411.
- Y. Liang, C.A. Coelho, T. von Rosen, Hypothesis testing in multivariate normal models with block circular covariance structures, *Biom. J.*, 64(3) (2022) 557–576.
- Y. Liang, D. von Rosen, T. von Rosen, On estimation in hierarchical models with block circular covariance structures, *Ann. Inst. Statist. Math.*, 67 (2015) 773–791.
- L. Lin, N. J. Higham, and J. Pan, Covariance structure regularization via entropy loss function, *Comput. Statist. Data Anal.*, 72 (2014), 315–327.
- J. Magnus, H. Neudecker, Symmetry, 0-1 matrices and Jacobians, a review, *Econom. Theory*, 2 (1986) 157–190.
- K. V. Mardia, T. Kent, M. Bibby, *Multivariate Analysis*, Academic Press, 1979.
- A. Markiewicz, A. Mieldzioc, Improved estimation of linearly structured covariance matrices, 2023. W recenzji.
- J. W. Mauchly, Significance test for sphericity of a normal n -variate distribution, *Ann. Math. Statist.*, 11 (1940) 204–209.
- R.J. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, Inc., New York, 1982.
- T. Nahtman, D. von Rosen, Shift permutation invariance in linear random factor models, *Math. Methods Statist.*, 17 (2008) 173–185.

- M. Ohlson, D. von Rosen, Explicit estimators of parameters in the growth curve model with linearly structured covariance matrices, *J. Multivariate Anal.*, 101, (2010) 1284–1295.
- I. Olkin, Inference for a normal population when the parameters exhibit some structure. Tech. Rep. No.77, Department of Statistics, Stanford University, 1973.
- I. Olkin, S. J. Press, Testing and estimation for a circular stationary model, *Ann. Math. Statist.*, 40 (1969) 1358–1373.
- C. R. Rao, Familial correlations or the multivariate generalizations of the intraclass correlation, *Current Science*, 14 (1945) 66–67.
- C. R. Rao, Discriminant functions for genetic differentiation and selection, *Sankhya*, 12 (1953) 229–246.
- C. R. Rao, Score test: Historical review and recent developments, In: *Advances in Ranking and Selection, Multiple Comparisons, and Reliability*, (2005) 3–20.
- A. Roy, R. Leiva, Estimating and testing a structured covariance matrix for three-level multivariate data, *Comm. Statist. Theory Methods*, 40 (2011) 1945–1963.
- A. Roy, K. Filipiak, D. Klein, Testing a block exchangeable covariance matrix, *Statistics*, 52 (2018) 393–408.
- A. Roy, R. Leiva, I. Žežula, D. Klein, Testing the equality of mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup, *J. Multivariate Anal.*, 137 (2015) 50–60.
- A. Roy, R. Zmysłony, M. Fonseca, R. Leiva, Optimal estimation for doubly multivariate data in blocked compound symmetric covariance structure, *J. Multivariate Anal.*, 144 (2016) 81–90.
- J. Seely, Quadratic subspaces and completeness. *Ann. Math. Statist.*, 42 (1971) 710–721.

- S. G. Self, K. Liang, Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *J. Amer. Statist. Assoc.*, 82 (1987) 605–610.
- C. Stein, *On Some Problems in Multivariate Analysis. Part I*, Department of Statistics, Stanford University, 1956
- G. Sun, J. Xie, Asymptotic normality and moderate deviation principle for high-dimensional likelihood ratio statistic on block compound symmetry covariance structure, *Statistics*, 54 (2020) 114–134.
- T. H. Szatrowski, Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. *Ann. Statist.*, 8 (1980) 802–810.
- S. Tsukada, Hypothesis testing for independence under blocked compound symmetric covariance structure, *Commun. Math. Stat.*, 6 (2018) 163–184.
- S. S. Wilks, Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distributions, *Ann. Math. Statist.*, 17 (1946) 257–281.
- V. Witkovský, CharFunTool: the characteristic functions tool-box (MATLAB), Available from: <https://github.com/witkovsky/CharFunTool> (2018).
- R. Zmysłony, I. Žežula, A. Koziół, Application of Jordan algebra for testing hypotheses about structure of mean vector in model with block compound symmetric covariance structure, *Electron. J. Linear Algebra*, 33 (2018) 41–52.
- P.W. Zehna, Invariance of maximum likelihood estimators, *Ann. Math. Statist.*, 37 (1966) 744.
- I. Žežula, Special variance structures in the growth curve model, *J. Multivariate Anal.*, 97 (2006) 606–618.

- I. Žežula, D. Klein, A. Roy, Testing of multivariate repeated measures data with block exchangeable covariance structure, *TEST*, 27 (2018) 360–378.