

Monika Mokrzycka

**APROKSYMACJA MACIERZY KOWARIANCJI  
WYBRANYMI STRUKTURAMI W MODELACH  
PODWÓJNIE WIELOWYMIAROWYCH**

Praca przygotowana  
w Instytucie Genetyki Roślin  
Polskiej Akademii Nauk w Poznaniu  
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 matematyczne*

Dyscyplina: *matematyka*

POZNAŃ, 2021



# Podziękowania

Pragnę serdecznie podziękować Pani Promotor dr hab. inż. Katarzynie Filipiak, której pomoc i wsparcie było kluczowym elementem podczas powstawania mojej pracy. Pod czujnym okiem Pani Promotor, w przyjaznej atmosferze, mogłam rozwijać swoją wiedzę i umiejętności. Dziękuję za poświęcony czas oraz ogrom cierpliwości i życzliwości, która spotykała mnie na każdym kroku.

Serdeczne podziękowania kieruję również w stronę Pana Profesora Pawła Krajewskiego, dyrektora Instytutu Genetyki Roślin PAN, który mimo naglących terminów zadań projektowych wykonywanych w IGR pozwolił mi na rozwój matematycznych pasji.

Praca ta nie powstałaby również bez wsparcia Pana Profesora Augustyna Markiewicza z Uniwersytetu Przyrodniczego w Poznaniu, któremu dziękuję za cenne rady i wskazówki oraz przyjęcie mnie do grupy uczestników Środowiskowego Seminarium z Algebry Liniowej i jej Zastosowań, organizowanego przez Pana Profesora Augustyna Markiewicza, Panią dr hab. inż. Katarzynę Filipiak z Politechniki Poznańskiej oraz Pana Profesora Jana Hauke z Uniwersytetu im. Adama Mickiewicza w Poznaniu. Dziękuję również Koleżankom i Kolegom z tej grupy za wspólne rozwiązywanie zagadnień algebraicznych.

Chciałabym podziękować także mojej naukowej rodzinie, uczestnikom międzynarodowej grupy badawczej "Multivariate and Mixed Linear Models", organizowanej przez Pana Profesora Augustyna Markiewicza oraz Panią dr hab. inż. Katarzynę Filipiak, za inspiracje i wskazywanie drogi w statystycznym świecie. Szczególne podziękowania kieruję do dr Daniela Kleina z Uniwersytetu Pavla Jozefa Šafárika w Košicach na Słowacji za wiele godzin wspólnej pracy oraz dla dr Jolanty Pielaszkiewicz z Uniwersytetu w Linköping w Szwecji za komentarze i uwagi, które okazały się bardzo pomocne w pisaniu tej pracy.

Na koniec chciałabym podziękować moim bliskim, którzy nie do końca rozumiejąc czym się zajmuję, wspierali mnie każdego dnia.



# Spis treści

Podziękowania .....	3
Abstract .....	7
Wstęp .....	9
1. Wprowadzenie .....	11
1.1. Obserwacje .....	11
1.2. Macierze kowariancji.....	12
1.3. Funkcje rozbieżności .....	15
1.4. Operatory i operacje na macierzach .....	19
2. Aproksymacja entropijną funkcją straty .....	23
2.1. Rozdzielna struktura kowariancyjna.....	23
2.2. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze kompletnie symetrycznej .....	26
2.3. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze procesu autoregresji pierwszego rzędu .....	35
3. Aproksymacja kwadratową funkcją straty.....	39
3.1. Rozdzielna struktura kowariancyjna.....	39
3.2. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze kompletnie symetrycznej .....	43
3.3. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze procesu autoregresji pierwszego rzędu .....	45
4. Badania symulacyjne oraz zastosowanie minimum funkcji rozbieżności do danych rzeczywistych .....	47
4.1. Wprowadzenie do badań symulacyjnych .....	47
4.2. Identyfikacja struktury .....	50
4.3. Estymacja.....	55
4.4. Testowanie hipotez .....	63
4.5. Dane rzeczywiste.....	77
5. Podsumowanie.....	81
Bibliografia .....	83



# Abstract

The Ph.D. dissertation is dedicated to approximation of the discrepancy between two probability distributions with different covariance matrices, expressed by the entropy or quadratic loss function. Doubly multivariate models that allow for the analysis of data with two sources of variability, for example, when several characteristics are observed repeatedly in time, are considered. The natural dependency structure is then Kronecker product of two symmetric positive definite matrices. The study also considers cases, where one component has additional structure of a compound symmetry matrix or a first-order autoregression process.

The algorithms for determination of the minimum of the respective discrepancy function are applied in simulation studies to identify the covariance structure, to propose estimators of covariance structures and to study their statistical properties, as well as in the research on the power of the tests, to measure the discrepancy between sets of covariance structures under consideration. Algebraic results are interpreted on real data.



# Wstęp

W dzisiejszym, szybko rozwijającym się świecie, nowoczesne techniki badawcze pozwalają na zbieranie bardzo dużych, wielokrotnie wielowymiarowych zestawów danych, mających zwykle formę tensorową. Klasyczne metody analizy statystycznej w takich przypadkach często zawodzą, a ich wielowymiarowe odpowiedniki nie obejmują danych tensorowych. Pojawia się zatem naturalna potrzeba rozwoju metod eksploracji tego typu zestawów danych.

Analizując eksperymenty, w których wiele cech mierzonych jest w wielu lokalizacjach, na różnych głębokościach etc. lub pomiary powtarzane są tak często, jak to możliwe, należy uwzględnić istniejące zależności między zmiennymi. Interesująca jest więc identyfikacja macierzy kowariancji pozwalająca poszerzyć wiedzę o zachowaniu zmiennych oraz umożliwiającą analizę danych przy użyciu precyzyjnego modelu statystycznego o mniejszej liczbie parametrów kowariancyjnych, pozostawiając więcej swobody na estymację interesujących parametrów modelu. Co więcej, zmniejszenie liczby parametrów umożliwia wnioskowanie statystyczne w sytuacji, gdy pojawia się problem wysokiej wymiarowości (ang. *high-dimensionality*), to znaczy, gdy wielkość próby jest zbyt mała w porównaniu z liczbą zmiennych. W takich sytuacjach klasyczne metody takie jak na przykład testy ilorazu wiarygodności nie są dostępne, więc obniżenie liczby parametrów realnie wpływa na ilość możliwych do wykorzystania metod. Znajomość zależności kowariancyjnych jest istotna w wielu metodach analizy statystycznej, między innymi w analizie składowych głównych, liniowej i kwadratowej analizie dyskryminacyjnej, w analizie regresji czy analizie niezależności zmiennych.

Wielokierunkowo wielowymiarowe eksperymenty przeprowadzane są w prawie wszystkich dziedzinach nauki, na przykład w biologii, genetyce, rolnictwie, jak również w obszarach biomedycznych, medycznych, związanych z ochroną środowiska czy naukach inżynierskich. Dlatego celem niniejszej pracy jest zaproponowanie algebraicznych metod identyfikacji struktury kowariancyjnej poprzez aproksymację rozbieżności między dwoma rozkładami prawdopodobieństwa o różnych macierzach kowariancji, wyrażonej za pomocą entropijnej lub kwadratowej funkcji straty. Rozważane są modele podwójnie wielowymiarowe pozwalające na analizę danych, w których występują dwa źródła zmienności, na przykład, gdy kilka cech obserwowanych jest wielokrotnie w czasie (lokalizacjach, głębokościach etc.). Naturalną strukturą zależności jest wówczas iloczyn Kroneckera dwóch macierzy symetrycznych, określonych

dotatnio, będących macierzami kowariancji odpowiednio dla cech i dla punktów czasowych. W pracy rozważa się również przypadki, gdy jeden z czynników iloczynu Kroneckera ma dodatkowo jedną z powszechnie występujących struktur korelacyjnych: strukturę kompletnej symetrii lub procesu autoregresji pierwszego rzędu.

Rozprawa składa się z czterech rozdziałów. W rozdziale pierwszym zdefiniowane zostały obserwacje tensorowe, rozważane struktury kowariancyjne, funkcje rozbieżności oraz przypomniane zostały pojęcia algebraiczne związane z operatorem wektoryzowania, iloczynem Kroneckera oraz regułami różniczkowania macierzy.

Kolejne dwa rozdziały zawierają najważniejsze wyniki algebraiczne niniejszej pracy. W Twierdzeniach 2.1–2.3 oraz w Twierdzeniach 3.1–3.3 wyprowadzone zostały wzory pozwalające na wyznaczenie minimum odpowiednio entropijnej i kwadratowej funkcji straty na zbiorach rozdzielnych struktur kowariancyjnych, rozdzielnych struktur kowariancyjnych z czynnikiem o strukturze całkowicie symetrycznej oraz strukturze procesu autoregresji pierwszego rzędu. Zaproponowano również algorytmy numeryczne rozwiązywania układów równań przedstawionych w twierdzeniach.

Rozdział czwarty przedstawia zastosowanie zaproponowanych metod aproksymacji do zagadnień statystycznych. Za pomocą badań symulacyjnych pokazane zostało, że zarówno entropijna jak i kwadratowa funkcja straty prawidłowo rozpoznają prawdziwą strukturę kowariancyjną tensora obserwacji. Ponadto zbadane zostały statystyczne własności estymatorów macierzy kowariancji otrzymanych poprzez minimalizację entropijnej oraz kwadratowej funkcji straty. Algorytmy minimalizujące rozważane funkcje straty zostały również użyte jako miary rozbieżności w kontekście badania mocy testów, do pomiaru rozbieżności między zbiorami struktur kowariancyjnych rozważanych w procesie testowania hipotez za pomocą testu ilorazu wiarygodności i testu wynikowego Rao. Na koniec prezentowane metody zostały zilustrowane na przykładzie rzeczywistych danych.

Rozprawę kończy Podsumowanie, w którym przedstawione zostały najważniejsze wnioski z przeprowadzonych badań.

# 1. Wprowadzenie

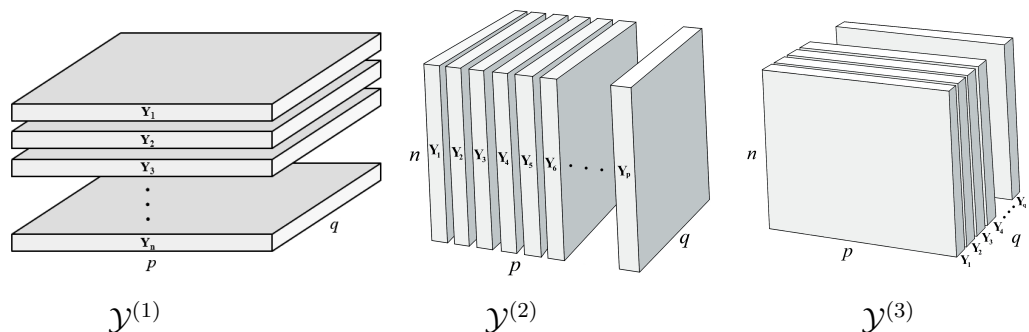
## 1.1. Obserwacje

Rozważmy eksperyment, w którym obserwuje się  $q$  charakterystyk w  $p$  punktach czasowych (lub lokalizacjach) dla każdego z  $n$  badanych obiektów. Zebrane dane można zaaranżować na wiele sposobów, m.in. tworząc trójindeksowe macierze (tensory),  $\mathcal{Y}$  (por. Rysunek 1.1).

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

Rysunek 1.1. Wizualizacja tensora  $\mathcal{Y} \in \mathbb{R}_{n \times q \times p}$ .

Rysunek 1.2 przedstawia tensor stopnia trzeciego  $\mathcal{Y} \in \mathbb{R}_{n \times q \times p}$  z uwzględnieniem trzech równoważnych zapisów macierzowych:  $\mathcal{Y}^{(1)}$  – macierze  $\mathbf{Y}_i$ ,  $i = 1, \dots, n$ , wymiaru  $q \times p$  ustawione jedna pod drugą,  $\mathcal{Y}^{(2)}$  – macierze  $\mathbf{Y}_k$ ,  $k = 1, \dots, p$ , wymiaru  $n \times q$  przedstawione jedna obok drugiej, lub  $\mathcal{Y}^{(3)}$  – macierze  $\mathbf{Y}_j$ ,  $j = 1, \dots, q$ , wymiaru  $n \times p$  zapisane jedna za drugą.



Rysunek 1.2. Graficzna reprezentacja sposobów macierzowania danych trójwymiarowych.

Tensor  $\mathcal{Y} \in \mathbb{R}_{n \times q \times p}$  można sprowadzić do macierzy  $n \times qp$ ,  $p \times nq$  lub  $q \times np$  dokonując wektoryzacji odpowiednich podmacierzy  $\mathbf{Y}_\bullet$  tensorów  $\mathcal{Y}^{(\tau)}$ ,  $\tau = 1, 2, 3$ , i odpowiednio je zapisując. Wektoryzując podmacierze  $\mathbf{Y}_i$ ,  $i = 1, \dots, n$ , macierzy  $\mathcal{Y}^{(1)}$  i zapisując ich transpozycje jedna pod drugą, otrzymamy macierz  $\mathbf{Y}^{(1)}$  wymiaru  $n \times qp$ . Wektoryzując podmacierze  $\mathbf{Y}_k$ ,  $k = 1, \dots, p$ , macierzy  $\mathcal{Y}^{(2)}$  i zapisując ich transpozycje jedna pod drugą, otrzymamy macierz  $\mathbf{Y}^{(2)}$  wymiaru  $p \times nq$ . Wektoryzując podmacierze  $\mathbf{Y}_j$ ,  $j = 1, \dots, q$ , macierzy  $\mathcal{Y}^{(3)}$  i zapisując ich transpozycje jedna pod drugą, otrzymamy macierz  $\mathbf{Y}^{(3)}$  wymiaru  $q \times np$ . Matematycznie operacje te można zapisać następująco (por. Kolda i Bader [19]):

$$\begin{aligned}\mathbf{Y}^{(1)} &= \sum_{i=1}^n \mathbf{e}_{i,n} \text{vec}' \mathbf{Y}_i = \sum_{i=1}^n \sum_{j=1}^q \sum_{k=1}^p y_{ijk} \mathbf{e}_{i,n} (\mathbf{e}'_{k,p} \otimes \mathbf{e}'_{j,q}), \\ \mathbf{Y}^{(2)} &= \sum_{k=1}^p \mathbf{e}_{k,p} \text{vec}' \mathbf{Y}_k = \sum_{i=1}^n \sum_{j=1}^q \sum_{k=1}^p y_{ijk} (\mathbf{e}_{j,q} \otimes \mathbf{e}_{i,n}) \mathbf{e}'_{k,p}, \\ \mathbf{Y}^{(3)} &= \sum_{j=1}^q \mathbf{e}_{j,q} \text{vec}' \mathbf{Y}_j = \sum_{i=1}^n \sum_{j=1}^q \sum_{k=1}^p y_{ijk} (\mathbf{e}_{k,p} \otimes \mathbf{e}_{i,n}) \mathbf{e}'_{j,q},\end{aligned}$$

gdzie  $\mathbf{e}_{\ell,t}$  jest  $\ell$ -tą kolumną macierzy jednostkowej stopnia  $t$ ,  $\mathbf{I}_t$ ,  $\text{vec}(\bullet)$  jest operatorem transformującym macierz  $\mathbf{A}$  wymiaru  $t_1 \times t_2$  w  $t_1 t_2$ -wymiarowy wektor poprzez zapisanie kolumn jedna pod drugą, natomiast symbol  $\otimes$  oznacza iloczyn Kroneckera. W każdym z powyższych wzorów druga równość wynika ze wzoru  $\text{vec}(\mathbf{ab}') = \mathbf{b} \otimes \mathbf{a}$  (por. Magnus i Neudecker [24]) zastosowanego do zapisu dowolnej macierzy  $\mathbf{A} : t_1 \times t_2$  w postaci  $\sum_{\ell} \sum_m a_{\ell m} \mathbf{e}_{\ell,t_1} \mathbf{e}'_{m,t_2}$ .

W dalszym ciągu niniejszej pracy będziemy wykorzystywać transformację  $\mathbf{Y}^{(1)}$  i dla uproszczenia będziemy ją oznaczać przez  $\mathbf{Y}$ .

Załóżmy, że każda z  $n$  niezależnych macierzy obserwacji  $\mathbf{Y}_i$ ,  $i = 1, \dots, n$ , ma macierzowy rozkład normalny z wartością oczekiwaną  $E(\mathbf{Y}_i) = \mathbf{M} \in \mathbb{R}_{q \times p}$  i symetryczną określoną dodatnio macierzą kowariancji  $D(\mathbf{Y}_i) = D(\text{vec } \mathbf{Y}_i) = \mathbf{\Omega} \in \mathbb{R}_{qp \times qp}^>$ . Zauważmy, że w powyższym modelu macierze  $\mathbf{M}$  i  $\mathbf{\Omega}$  są takie same dla każdego obiektu. Jeżeli  $\mathbf{y}_i = \text{vec } \mathbf{Y}_i$  to  $\mathbf{y}_i \sim N_{qp}(\text{vec } \mathbf{M}, \mathbf{\Omega})$  i  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_n)'$  ma macierzowy rozkład normalny

$$\mathbf{Y} \sim N_{n,qp}(\mathbf{1}_n \text{vec}' \mathbf{M}, \mathbf{I}_n, \mathbf{\Omega}), \quad (1.1)$$

gdzie  $\mathbf{1}_n$  jest  $n$ -wymiarowym wektorem jedynek.

## 1.2. Macierze kowariancji

Liczba nieznanych parametrów w macierzy  $\mathbf{\Omega}$  wynosi  $pq(pq + 1)/2$ . Do najpopularniejszych estymatorów  $\mathbf{\Omega}$  należy macierz kowariancji z próby,  $\mathbf{S}^* = \frac{1}{n-1} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$ ,

oraz estymator największej wiarygodności,  $\mathbf{S} = \frac{1}{n} \mathbf{Y}' \mathbf{Q}_n \mathbf{Y}$ , gdzie  $\mathbf{Q}_n = \mathbf{I}_n - \mathbf{P}_n$  oraz  $\mathbf{P}_n = \frac{1}{n} \mathbf{1}_n \mathbf{1}_n'$ . Zauważmy, że w przypadku gdy  $n \leq pq$  obydwa estymatory są macierzami osobliwymi (pojawia się problem tak zwanej wysokiej wymiarowości). Ponadto, efektywność estymacji zmniejsza się gwałtownie wraz ze wzrostem  $p$  lub  $q$ . Jedną z metod rozwiązania problemu osobliwości estymatora macierzy kowariancji jest przyjęcie pewnej struktury pozwalającej na zmniejszenie liczby nieznanych parametrów. Najbardziej naturalnym dla modeli podwójnie wielowymiarowych wydaje się założenie, że cechy są skorelowane niezależnie od punktów czasowych (lokalizacji) oraz punkty czasowe (lokalizacje) są skorelowane niezależnie od cech. Wówczas macierz kowariancji macierzy  $\mathbf{Y}_i$ ,  $i = 1, \dots, n$ , wyraża się jako następujący iloczyn Kroneckera

$$\mathbf{\Omega} = \mathbf{\Psi} \otimes \mathbf{\Sigma},$$

gdzie  $\mathbf{\Psi} \in \mathbb{R}_{p \times p}^>$  jest macierzą kowariancji dla punktów czasowych (lokalizacji), taką samą dla każdej charakterystyki, natomiast  $\mathbf{\Sigma} \in \mathbb{R}_{q \times q}^>$  jest macierzą kowariancji dla charakterystyk, taką samą dla każdego punktu czasowego (lokalizacji). Taka postać macierzy kowariancji nosi nazwę struktury rozdzielnej (ang. *separable structure*). Podzbiór wszystkich macierzy kowariancji o strukturze rozdzielnej oznaczamy przez  $\mathcal{S}$ , to znaczy

$$\mathcal{S} = \{ \mathbf{\Psi} \otimes \mathbf{\Sigma} \in \mathbb{R}_{pq \times pq}^> : \mathbf{\Psi} \in \mathbb{R}_{p \times p}^>, \mathbf{\Sigma} \in \mathbb{R}_{q \times q}^> \}.$$

Ponieważ iloczyn Kroneckera nie jest jednoznaczny, bo

$$\mathbf{\Psi} \otimes \mathbf{\Sigma} = c \mathbf{\Psi} \otimes \frac{1}{c} \mathbf{\Sigma},$$

zazwyczaj zakłada się, że jeden z elementów macierzy  $\mathbf{\Psi}$  lub  $\mathbf{\Sigma}$  jest ustalony. W niniejszej pracy będziemy przyjmować, że  $\psi_{11} = 1$ .

Zauważmy, że liczba parametrów zmniejszyła się teraz znacząco w porównaniu z macierzą  $\mathbf{\Omega}$  nieposiadającą struktury, do  $\left( \frac{p(p+1)}{2} - 1 \right) + \frac{q(q+1)}{2}$ .

Estymacją największej wiarygodności struktury rozdzielnej zajmowali się Lu i Zimmerman [23] wyrażając estymatory poszczególnych elementów macierzy kowariancji,  $\hat{\psi}_{\ell m}$ ,  $\hat{\sigma}_{\ell^* m^*}$ ,  $\ell, m = 1, \dots, p$ ,  $\ell^*, m^* = 1, \dots, q$ , z wykorzystaniem wielokrotnych sum. Filipiak i in. [13] stosowali zapis wektorowy,  $\text{vec } \hat{\mathbf{\Psi}}$ ,  $\text{vec } \hat{\mathbf{\Sigma}}$ , natomiast Filipiak i Klein [8] zapisali estymator struktury rozdzielnej w naturalnej macierzowej formie,  $\hat{\mathbf{\Psi}}$ ,  $\hat{\mathbf{\Sigma}}$ . Warto zwrócić uwagę, że żaden znany w literaturze zapis nie podaje osobno estymatorów  $\mathbf{\Psi}$  i  $\mathbf{\Sigma}$  w postaci jawnej; estymator  $\mathbf{\Psi}$  zależy od estymatora  $\mathbf{\Sigma}$  i odwrotnie, czego konsekwencją jest potrzeba numerycznego rozwiązania odpowiedniego układu równań.

Naturalna struktura rozdzielna ma związek ze sposobem przeprowadzenia eksperymentu. Jednakże bardzo często, np. w doświadczeniach genetycznych, badacz nie posiada żadnej wiedzy na temat potencjalnej struktury zależności. Wówczas, w celu dostrzeżenia ewentualnej struktury, stosowane są różnego rodzaju metody graficzne na przykład sieci neuronowe (ang. *neural networks*), mapowanie (Gilson i in. [16]), graficzny algorytm lasso (Devijver i Gallopin [4]), lub opracowane w ostatnich latach metody algebraiczne bazujące na normie Frobeniusa (Cui i in. [3], Filipiak i Klein [7]) czy na entropijnej funkcji straty (Lin i in. [21], Filipiak i in. [9]). Ze statystycznego punktu widzenia właściwą metodą identyfikacji struktury zależności jest testowanie hipotez o (rozdzielnej) strukturze kowariancyjnej, na przykład poprzez testy oparte na ilorazie wiarygodności (Lu i Zimmerman [23], Roy i Khattree [32, 33], Mitchell i in. [27], Roy [31], Roy i Leiva [34], Srivastava i in. [37], Simpson [35], Mancaeur i Dutilleul [26], Filipiak i in. [13]) lub test wynikowy Rao (Filipiak i in. [13]).

Strukturę rozdzielną można dodatkowo doprecyzować, zakładając określoną strukturę jednego z czynników iloczynu Kroneckera. Ponieważ rolę macierzy  $\Psi$  i  $\Sigma$  można łatwo zamienić dokonując transformacji postaci

$$\Psi \otimes \Sigma = \mathbf{K}_{q,p} (\Sigma \otimes \Psi) \mathbf{K}_{p,q},$$

gdzie  $\mathbf{K}_{s,t} \in \mathbb{R}_{st \times st}$  jest macierzą komutacji transformującą macierz  $\mathbf{G} \in \mathbb{R}_{s \times t}$  zgodnie z formułą

$$\mathbf{K}_{s,t} \text{vec } \mathbf{G} = \text{vec } \mathbf{G}'$$

(por. Magnus i Neudecker [24], Kollo i von Rosen [20]), bez straty ogólności zakładać będziemy dodatkową strukturę macierzy  $\Psi$ . W niniejszej pracy rozważać będziemy następujące struktury korelacyjne:

- wszystkie obserwacje z różnych punktów czasowych są jednakowo skorelowane, co oznacza, że korelacje nie zmieniają się wraz z odległościami między punktami czasowymi - wówczas  $\Psi$  nazywana jest strukturą kompletnej symetrii (ang. *compound symmetry*), CS, i przyjmuje postać  $\Psi_{\text{CS}} = (1 - \varrho)\mathbf{I}_p + \varrho\mathbf{1}_p\mathbf{1}_p'$ ,
- korelacje między obserwacjami z różnych punktów czasowych zmieniają się wykładniczo wraz z odległościami między punktami czasowymi - wówczas  $\Psi$  nazywana jest strukturą procesu autoregresji pierwszego rzędu (ang. *first order autoregression*), AR(1), i przyjmuje postać  $\Psi_{\text{AR}} = \mathbf{I}_p + \sum_{i=1}^{p-1} \varrho^i (\mathbf{C}^i + \mathbf{C}^{i'})$ , gdzie  $\mathbf{C} = (c_{ij})$ ,  $c_{ij} = 1$  dla  $j - i = 1$  i 0 w przeciwnym przypadku,  $i, j = 1, \dots, p$ .

Aby zapewnić dodatnią określoność macierzy  $\Psi_{\text{CS}}$  i  $\Psi_{\text{AR}}$ , współczynnik korelacji  $\varrho$  musi należeć odpowiednio do przedziału  $\left(-\frac{1}{p-1}, 1\right)$  i  $(-1, 1)$ .

W niniejszej pracy zbiory powyższych struktur rozdzielnych definiować będziemy jako podzbiory  $\mathcal{S}$  w następujący sposób

$$\mathcal{S}_{\text{CS}} = \{\Psi_{\text{CS}} \otimes \Sigma \in \mathbb{R}_{pq \times pq}^> : \Psi_{\text{CS}} \in \mathbb{R}_{p \times p}^>, \Sigma \in \mathbb{R}_{q \times q}^>\},$$

$$\mathcal{S}_{\text{AR}} = \{\Psi_{\text{AR}} \otimes \Sigma \in \mathbb{R}_{pq \times pq}^> : \Psi_{\text{AR}} \in \mathbb{R}_{p \times p}^>, \Sigma \in \mathbb{R}_{q \times q}^>\}.$$

Warto zauważyć, że zbiór  $\mathcal{S}$  oraz jego podzbiory  $\mathcal{S}_{\text{CS}}$  i  $\mathcal{S}_{\text{AR}}$  nie są zbiorami wypukłymi. Ponadto tylko zbiory  $\mathcal{S}$  i  $\mathcal{S}_{\text{CS}}$  są zamknięte ze względu na obustronne mnożenie macierzy, to znaczy

- (i)  $\mathbf{A}, \mathbf{B} \in \mathcal{S} \quad \Rightarrow \quad \mathbf{ABA} \in \mathcal{S},$
- (ii)  $\mathbf{A}, \mathbf{B} \in \mathcal{S}_{\text{CS}} \quad \Rightarrow \quad \mathbf{ABA} \in \mathcal{S}_{\text{CS}},$
- (iii)  $\mathbf{A}, \mathbf{B} \in \mathcal{S}_{\text{AR}} \quad \not\Rightarrow \quad \mathbf{ABA} \in \mathcal{S}_{\text{AR}}.$

Zauważmy ponadto, że dla  $\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{A}_2$ ,  $|\mathbf{A}_1| \neq 0$ ,  $|\mathbf{A}_2| \neq 0$  (przy czym  $\mathbf{A}_1$ ,  $\mathbf{A}_2$  niekoniecznie są określone dodatnio) zachodzą następujące własności:

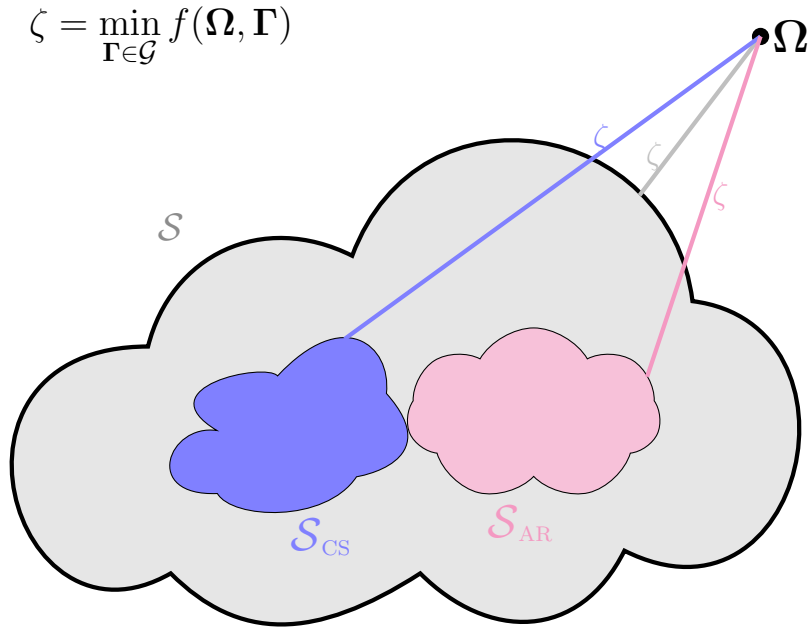
- (i')  $\mathbf{B} \in \mathcal{S} \quad \Rightarrow \quad \mathbf{ABA}' \in \mathcal{S},$
- (ii')  $\mathbf{B} \in \mathcal{S}_{\text{CS}} \wedge \mathbf{A}_1 = \mathbf{1}_p \mathbf{1}'_p + \beta \mathbf{B}_1, \mathbf{B}_1 \mathbf{B}'_1 = \mathbf{Q}_p \quad \Rightarrow \quad \mathbf{ABA}' \in \mathcal{S}_{\text{CS}},$
- (iii')  $\mathbf{B} \in \mathcal{S}_{\text{AR}} \wedge \mathbf{A}_1 = \mathbf{I}_p \quad \Rightarrow \quad \mathbf{ABA}' \in \mathcal{S}_{\text{AR}}.$

Można również zauważyć, że dla  $\mathbf{B} \in \mathcal{S}$  ( $\mathbf{B} \in \mathcal{S}_{\text{CS}}$ ) oraz macierzy  $\mathbf{A}$  spełniającej (i') (odpowiednio (ii')) macierz  $\mathbf{ABA}'$  generuje cały zbiór  $\mathcal{S}$  (odpowiednio  $\mathcal{S}_{\text{CS}}$ ). Własność ta nie zachodzi dla  $\mathbf{B} \in \mathcal{S}_{\text{AR}}$ .

### 1.3. Funkcje rozbieżności

Głównym celem niniejszej rozprawy jest zaproponowanie algebraicznych technik identyfikacji najbardziej odpowiedniej zależności spośród zbioru potencjalnych struktur. Identyfikacja może być traktowana jako pierwszy, wstępny krok pozwalający na sformułowanie właściwej hipotezy w procesie testowania. Zauważmy, że nie dysponując wystarczającą wiedzą a priori o zależnościach między obserwacjami przyjmuje się, że macierz kowariancji  $\Omega$  nie posiada żadnej struktury i estymuje się ją na przykład za pomocą macierzy  $\mathbf{S}$ . Problem identyfikacji struktury sprowadza się wówczas do wyznaczenia macierzy ze zbioru  $\mathcal{S}$ ,  $\mathcal{S}_{\text{CS}}$  i  $\mathcal{S}_{\text{AR}}$ , dla której rozbieżność

od macierzy  $\mathbf{S}$ , w sensie pewnej funkcji rozbieżności,  $f(\mathbf{S}, \mathbf{\Gamma})$ ,  $\mathbf{\Gamma} \in \mathcal{G}$ , gdzie  $\mathcal{G} = \mathcal{S}$ ,  $\mathcal{G} = \mathcal{S}_{CS}$  lub  $\mathcal{G} = \mathcal{S}_{AR}$ , jest najmniejsza. Przypomnijmy, że  $\mathcal{S}_{CS}, \mathcal{S}_{AR} \subset \mathcal{S}$ , dlatego najmniejszą rozbieżność otrzymamy zawsze dla struktury ze zbioru  $\mathcal{S}$ . Niemniej jednak, ze względu na niewielką liczbę parametrów, również struktura ze zbioru  $\mathcal{S}_{CS}$  lub  $\mathcal{S}_{AR}$  jest interesująca (więcej stopni swobody pozostawionych zostaje na estymację parametrów głównych modelu). Wizualizację badanego problemu dla dowolnej macierzy  $\mathbf{\Omega}$  przedstawia Rysunek 1.3. Warto przypomnieć, że żaden ze zbiorów  $\mathcal{S}$ ,  $\mathcal{S}_{CS}$  czy  $\mathcal{S}_{AR}$  nie jest wypukły. Proces identyfikacji struktury kowariancyjnej bywa nazywany w literaturze regularyzacją; por. Lin et al. [21], Cui et al. [3].



Rysunek 1.3. Wizualizacja procesu identyfikacji struktury macierzy kowariancji.

Do identyfikacji struktury zależności Filipiak i in. [7] jako funkcję rozbieżności wykorzystali normę Frobeniusa definiowaną

$$f_F(\mathbf{\Omega}, \mathbf{\Gamma}) = \|\mathbf{\Omega} - \mathbf{\Gamma}\|_F^2 = \text{tr} [(\mathbf{\Omega} - \mathbf{\Gamma})(\mathbf{\Omega} - \mathbf{\Gamma})']. \quad (1.2)$$

W niniejszej pracy jako miarę rozbieżności będziemy używać

- entropijną funkcję straty (por. Stein [38], Dey i Srinivasan [5], James i Stein [18], Lin i in. [21]) postaci

$$f_E(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr}(\mathbf{\Omega}^{-1}\mathbf{\Gamma}) - \ln |\mathbf{\Omega}^{-1}\mathbf{\Gamma}| - pq, \quad (1.3)$$

oraz

- kwadratową funkcję straty (por. Anderson [1])

$$f_Q(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr} \left[ (\mathbf{\Omega}^{-1} \mathbf{\Gamma} - \mathbf{I}_{pq})^2 \right]. \quad (1.4)$$

Entropijna funkcja straty znana jest również jako rozbieżność Kullbacka-Leiblera między dwoma wielowymiarowymi rozkładami, które różnią się macierzami kowariancji; Pan i Fang [29].

Dla zadanej macierzy  $\mathbf{\Omega}$  oznaczmy funkcję  $f_k(\mathbf{\Omega}, \mathbf{\Gamma})$  przez  $L_{\mathbf{\Omega}}^{(k)}(\mathbf{\Gamma})$ ,  $k \in \{E, Q\}$ . Interesuje nas wyznaczenie

$$\zeta^{(k)} = \min_{\mathbf{\Gamma} \in \mathcal{G}} L_{\mathbf{\Omega}}^{(k)}(\mathbf{\Gamma}), \quad k \in \{E, Q\}. \quad (1.5)$$

Zarówno entropijna funkcja straty, jak i kwadratowa funkcja straty są nieujemne (dla  $\mathbf{\Gamma}, \mathbf{\Omega}$  określonych dodatnio), przyjmują wartość równą 0 dla  $\mathbf{\Gamma} = \mathbf{\Omega}$ , jednakże nie są symetryczne,  $f_k(\mathbf{\Omega}, \mathbf{\Gamma}) \neq f_k(\mathbf{\Gamma}, \mathbf{\Omega})$ ,  $k \in \{E, Q\}$ , nie mogą więc być traktowane jak odległości. Niemniej jednak wartość  $\zeta^{(k)}$ ,  $k \in \{E, Q\}$ , może być rozumiana jako miara rozbieżności pomiędzy daną macierzą  $\mathbf{\Omega}$  a odpowiednim zbiorem struktur; Lin i in. [21].

Niech  $\mathbf{\Omega}$  będzie dowolną symetryczną macierzą określoną dodatnio, niech  $\mathcal{A}$  będzie podzbiorem wszystkich macierzy symetrycznych określonych dodatnio oraz niech  $\mathbf{\Gamma} \in \mathcal{A}$ . Zarówno entropijna jak i kwadratowa funkcja straty są wypukłe ze względu na  $\mathbf{\Gamma}$  (Stein [39]; Muirhead [28]) oraz antysymetryczne, to znaczy, dla  $k \in \{E, Q\}$

$$f_k(\mathbf{\Omega}, \mathbf{\Gamma}) = f_k(\mathbf{\Gamma}, \mathbf{\Omega}) \quad \Rightarrow \quad \mathbf{\Omega} = \mathbf{\Gamma},$$

oraz

$$\mathbf{\Omega} = \mathbf{\Gamma} \quad \Leftrightarrow \quad f_k(\mathbf{\Omega}, \mathbf{\Gamma}) = f_k(\mathbf{\Gamma}, \mathbf{\Omega}) = 0.$$

Ponadto, zarówno entropijna jak i kwadratowa funkcja straty są niezmiennicze ze względu na grupę przekształceń liniowych (James i Stein [18], Haff [17]), to znaczy dla dowolnej nieosobliwej macierzy  $\mathbf{A}$

$$f_k(\mathbf{A}\mathbf{\Omega}\mathbf{A}', \mathbf{A}\mathbf{\Gamma}\mathbf{A}') = f_k(\mathbf{\Omega}, \mathbf{\Gamma}). \quad (1.6)$$

Nieosobliwość macierzy  $\mathbf{A}$  jest wymagana ze względu na konieczność obliczenia odwrotności macierzy  $\mathbf{A}\mathbf{\Omega}\mathbf{A}'$  występującej w  $f_k(\mathbf{A}\mathbf{\Omega}\mathbf{A}', \mathbf{A}\mathbf{\Gamma}\mathbf{A}')$ . Warto zauważyć, że  $\mathbf{A}\mathbf{\Omega}\mathbf{A}'$  jest przekształceniem liniowym macierzy  $\mathbf{\Omega}$  z wykorzystaniem operatora  $\mathbf{A} \otimes \mathbf{A}$ , działającego z  $(pq)^2$ -wymiarowej przestrzeni wektorowej na przestrzeń wektorową tego samego wymiaru, to jest  $\mathbf{A} \otimes \mathbf{A} : \mathbb{R}_{pq} \rightarrow \mathbb{R}_{pq}$ . Stąd, każdy

element tej przestrzeni można zapisać w postaci macierzy stopnia  $pq$ , to znaczy  $\mathbf{A} \otimes \mathbf{A}(\boldsymbol{\Omega}) = \mathbf{A}\boldsymbol{\Omega}\mathbf{A}'$ . Przekształcenie to może być równoważnie wyrażone jako  $\text{vec}(\mathbf{A}\boldsymbol{\Omega}\mathbf{A}') = (\mathbf{A} \otimes \mathbf{A}) \text{vec } \boldsymbol{\Omega}$ .

Przypomnijmy, że dla danej macierzy  $\boldsymbol{\Omega}$  oznaczamy  $f_k(\boldsymbol{\Omega}, \boldsymbol{\Gamma})$  jako  $L_{\boldsymbol{\Omega}}^{(k)}(\boldsymbol{\Gamma})$ . Wyznaczenie minimum z (1.5) można sprowadzić do wskazania  $\hat{\boldsymbol{\Gamma}} \in \mathcal{G}$  takiego, że

$$\hat{\boldsymbol{\Gamma}} = \underset{\boldsymbol{\Gamma} \in \mathcal{G}}{\text{argmin}} L_{\boldsymbol{\Omega}}^{(k)}(\boldsymbol{\Gamma}). \quad (1.7)$$

Ponieważ zarówno entropijna jak i kwadratowa funkcja straty nie są symetryczne, czyli  $f_k(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) \neq f_k(\boldsymbol{\Gamma}, \boldsymbol{\Omega})$ , nie są one miarami odległości pomiędzy  $\boldsymbol{\Omega}$  i  $\boldsymbol{\Gamma}$ . Niemniej jednak wartość  $L_{\boldsymbol{\Omega}}^{(k)}(\hat{\boldsymbol{\Gamma}})$  z  $\hat{\boldsymbol{\Gamma}}$  będącym najlepszym przybliżeniem macierzy  $\boldsymbol{\Omega}$ , jest traktowana jako miara rozbieżności między daną macierzą  $\boldsymbol{\Omega}$  i zbiorem symetrycznych, określonych dodatnio macierzy,  $\mathcal{G}$  (Lin i in. [21]). Dlatego jeśli  $\boldsymbol{\Omega} \in \mathcal{G}$  to jest ona najlepszym przybliżeniem samej siebie,  $\hat{\boldsymbol{\Gamma}} = \boldsymbol{\Omega}$ , oraz rozbieżność między  $\boldsymbol{\Omega}$  i  $\mathcal{G}$  jest równa 0, to znaczy

$$\min_{\boldsymbol{\Gamma} \in \mathcal{G}} L_{\boldsymbol{\Omega}}^{(k)}(\boldsymbol{\Gamma}) = L_{\boldsymbol{\Omega}}^{(k)}(\hat{\boldsymbol{\Gamma}}) = 0, \quad k \in \{E, Q\}. \quad (1.8)$$

Rozważmy teraz własność (1.6) i jej związek z (1.7).

**Lemat 1.1.** *Niech  $\mathbf{A}$  będzie macierzą nieosobliwą taką, że operator liniowy  $\mathbf{A} \otimes \mathbf{A}$  przekształca zbiór macierzy symetrycznych, określonych dodatnio,  $\mathcal{G}$ , na  $\mathcal{G}$ . Ponadto niech  $\hat{\boldsymbol{\Gamma}}$  będzie zdefiniowane jak w (1.7) i niech  $\hat{\boldsymbol{\Delta}} \in \mathcal{G}$  będzie minimum na zbiorze  $\mathcal{G}$  z funkcji  $L_{\mathbf{A}\boldsymbol{\Omega}\mathbf{A}'}^{(k)}(\boldsymbol{\Delta})$ , gdzie  $k \in \{E, Q\}$ . Wtedy prawdziwe jest jedno z poniższych stwierdzeń:*

- (i)  $\boldsymbol{\Omega}, \mathbf{A}\boldsymbol{\Omega}\mathbf{A}' \in \mathcal{G}$  i  $L_{\boldsymbol{\Omega}}^{(k)}(\hat{\boldsymbol{\Gamma}}) = L_{\mathbf{A}\boldsymbol{\Omega}\mathbf{A}'}^{(k)}(\hat{\boldsymbol{\Delta}}) = 0$ ;
- (ii)  $\boldsymbol{\Omega}, \mathbf{A}\boldsymbol{\Omega}\mathbf{A}' \notin \mathcal{G}$  i jeżeli  $\mathbf{A}\hat{\boldsymbol{\Gamma}}\mathbf{A}' \in \mathcal{G}$  to  $L_{\boldsymbol{\Omega}}^{(k)}(\hat{\boldsymbol{\Gamma}}) = L_{\mathbf{A}\boldsymbol{\Omega}\mathbf{A}'}^{(k)}(\hat{\boldsymbol{\Delta}})$ .

*Dowód.* Z założenia, że operator liniowy  $\mathbf{A} \otimes \mathbf{A}$  jest bijekcją wynika, że  $\boldsymbol{\Omega} \in \mathcal{G}$  wtedy i tylko wtedy, gdy  $\mathbf{A}\boldsymbol{\Omega}\mathbf{A}' \in \mathcal{G}$  oraz, że  $\boldsymbol{\Omega} \notin \mathcal{G}$  wtedy i tylko wtedy, gdy  $\mathbf{A}\boldsymbol{\Omega}\mathbf{A}' \notin \mathcal{G}$ .

(i) Z równania (1.8) mamy, że  $\hat{\boldsymbol{\Gamma}} = \boldsymbol{\Omega}$  i  $\hat{\boldsymbol{\Delta}} = \mathbf{A}\boldsymbol{\Omega}\mathbf{A}'$ , co jest równoważne  $\mathbf{A}\hat{\boldsymbol{\Gamma}}\mathbf{A}' = \hat{\boldsymbol{\Delta}}$ . Stąd  $L_{\boldsymbol{\Omega}}^{(k)}(\hat{\boldsymbol{\Gamma}}) = L_{\mathbf{A}\boldsymbol{\Omega}\mathbf{A}'}^{(k)}(\hat{\boldsymbol{\Delta}}) = 0$ .

(ii) Jeżeli  $\mathbf{A}\hat{\boldsymbol{\Gamma}}\mathbf{A}' \in \mathcal{G}$  to  $\hat{\boldsymbol{\Delta}} = \mathbf{A}\hat{\boldsymbol{\Gamma}}\mathbf{A}'$  oraz, z (1.6),  $L_{\boldsymbol{\Omega}}^{(k)}(\hat{\boldsymbol{\Gamma}}) = L_{\mathbf{A}\boldsymbol{\Omega}\mathbf{A}'}^{(k)}(\hat{\boldsymbol{\Delta}})$ . ■

Przypomnijmy, że w niniejszej pracy rozważamy  $\mathcal{G} = \mathcal{S}$ ,  $\mathcal{G} = \mathcal{S}_{\text{CS}}$  lub  $\mathcal{G} = \mathcal{S}_{\text{AR}}$ .

## 1.4. Operatory i operacje na macierzach

W dalszej części rozprawy wykorzystywanych będzie wiele terminów algebraicznych oraz własności różnych operatorów, które teraz podamy.

W trakcie przekształceń wykorzystywane są podstawowe własności macierzy, które znaleźć można w pracy Magnus i Neudecker [24]. Dla przypomnienia są one przedstawione w poniższych lematkach.

**Lemat 1.2.** *Dla dowolnych macierzy  $\mathbf{A}$  i  $\mathbf{B}$  tego samego stopnia zachodzi następujący związek między operatorem  $\text{vec}(\bullet)$  i śladem macierzy:*

$$\text{vec}' \mathbf{A} \cdot \text{vec} \mathbf{B} = \text{tr}(\mathbf{A}'\mathbf{B}).$$

**Lemat 1.3.** *Dla dowolnych kwadratowych macierzy  $\mathbf{A}$  i  $\mathbf{B}$ , niekoniecznie tego samego stopnia, zachodzi związek*

$$\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr} \mathbf{A} \cdot \text{tr} \mathbf{B}.$$

**Lemat 1.4.** *Dla dowolnych kwadratowych nieosobliwych macierzy  $\mathbf{A}$  i  $\mathbf{B}$  zachodzi własność*

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}.$$

**Lemat 1.5.** *Dla dowolnych kwadratowych macierzy  $\mathbf{A}$  i  $\mathbf{B}$  stopnia  $m$  i  $n$  odpowiednio zachodzi własność*

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^n |\mathbf{B}|^m.$$

**Lemat 1.6.** *Niech  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  będą takimi macierzami, że iloczyn  $\mathbf{ABC}$  istnieje. Wtedy*

$$\text{vec} \mathbf{ABC} = (\mathbf{C}' \otimes \mathbf{A}) \text{vec} \mathbf{B}.$$

**Lemat 1.7.** *Niech  $\mathbf{A} : m \times n$  oraz  $\mathbf{B} : p \times q$ . Wtedy*

$$\text{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_n \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_p) (\text{vec} \mathbf{A} \otimes \text{vec} \mathbf{B}).$$

W celu wyznaczenia najlepszego przybliżenia macierzy  $\mathbf{\Omega}$  należy wyznaczyć minimum odpowiedniej funkcji straty. W rozprawie minimum wyznaczane jest poprzez różniczkowanie entropijnej lub kwadratowej funkcji straty względem nieznanymi zmiennymi macierzowymi. Wykorzystane zostaną następujące wzory na różniczkowanie macierzy; por. Fackler [6].

**Lemat 1.8.** Dla  $\mathbf{A} : m \times n$ ,  $\mathbf{B} : p \times q$  oraz  $\mathbf{X}$  odpowiedniego wymiaru zachodzą następujące zależności:

$$\begin{aligned}
\frac{d \operatorname{tr}(\mathbf{A}\mathbf{X})}{d\mathbf{X}} &= \operatorname{vec}'(\mathbf{A}') \\
\frac{d \ln |\mathbf{A}|}{d\mathbf{A}} &= \operatorname{vec}'\left[(\mathbf{A}^{-1})'\right] \\
\frac{d\mathbf{A}^{-1}}{d\mathbf{A}} &= -\left[(\mathbf{A}^{-1})' \otimes \mathbf{A}^{-1}\right] \\
\frac{d\mathbf{A}\mathbf{X}\mathbf{B}}{d\mathbf{X}} &= \mathbf{B}' \otimes \mathbf{A} \\
\frac{d\mathbf{A} \otimes \mathbf{B}}{d\mathbf{A}} &= (\mathbf{I}_n \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_p) (\mathbf{I}_{mn} \otimes \operatorname{vec} \mathbf{B}) \\
\frac{d\mathbf{A} \otimes \mathbf{B}}{d\mathbf{B}} &= (\mathbf{I}_n \otimes \mathbf{K}_{q,m} \otimes \mathbf{I}_p) (\operatorname{vec} \mathbf{A} \otimes \mathbf{I}_{pq}) \\
\frac{d\mathbf{A}^i}{d\mathbf{A}} &= \sum_{j=1}^i (\mathbf{A}')^{i-j} \otimes \mathbf{A}^{j-1}.
\end{aligned}$$

Jako że rozważana w rozprawie macierz kowariancji jest symetryczna, posiada ona  $pq(pq+1)/2$  niewiadomych. Powtarzające się elementy mogą być wyeleminowane przez wprowadzenie operatora  $\operatorname{vech}(\bullet)$  przekształcającego macierz kowariancji w  $pq(pq+1)/2$ -wymiarowy wektor poprzez eliminację naddiagonalnych elementów. Niech zatem  $\mathbf{A}$  będzie symetryczną macierzą stopnia  $m$ . Wówczas  $\operatorname{vech} \mathbf{A}$  zawiera tylko różne elementy macierzy  $\mathbf{A}$ . Ponieważ składowymi wektora  $\operatorname{vec} \mathbf{A}$  są składowe wektora  $\operatorname{vech} \mathbf{A}$  z odpowiednimi powtórzeniami, więc istnieje jednoznaczna macierz wymiaru  $m^2 \times \frac{1}{2}m(m+1)$ , która przekształca  $\operatorname{vech} \mathbf{A}$  w  $\operatorname{vec} \mathbf{A}$ . Macierz ta nazywana jest macierzą duplikacji i jest oznaczana jako  $\mathbf{D}_m$ . Zatem,

$$\mathbf{D}_m \cdot \operatorname{vech} \mathbf{A} = \operatorname{vec} \mathbf{A}.$$

Korzystając z reguły łańcuchowej różniczkowania funkcji złożonej macierzy symetrycznej (por. Magnus i Neudecker [24], Filipiak i in. [12]) w celu uzyskania odpowiednich pochodnych dla macierzy symetrycznych wystarczy kolejne formuły z Lematu 1.8 przemnożyć prawostronnie przez macierz duplikacji.

**Lemat 1.9.** Niech  $\mathbf{F}(\mathbf{A})$  będzie  $k \times l$  wymiarową funkcją macierzową argumentu  $\mathbf{A}$ .

(i) Jeśli  $\mathbf{A}$  jest macierzą  $m \times n$  wymiarową, to  $\partial \operatorname{vec} \mathbf{F}(\mathbf{A}) / \partial \operatorname{vec}' \mathbf{A}$  jest macierzą  $kl \times mn$  wymiarową taką, że jej  $(i, j)$ -ty element jest pochodną  $i$ -tej składowej wektora  $\operatorname{vec} \mathbf{F}(\mathbf{A})$  ze względu na  $j$ -tą składową wektora  $\operatorname{vec} \mathbf{A}$ .

(ii) Jeśli  $\mathbf{A}$  jest symetryczną macierzą stopnia  $m$ , to

$$\frac{\partial \operatorname{vec} \mathbf{F}(\mathbf{A})}{\partial \operatorname{vech}' \mathbf{A}} = \frac{\partial \operatorname{vec} \mathbf{F}(\mathbf{A})}{\partial \operatorname{vec}' \bar{\mathbf{A}}} \cdot \frac{\partial \operatorname{vec} \bar{\mathbf{A}}}{\partial \operatorname{vech}' \mathbf{A}} = \frac{\partial \operatorname{vec} \mathbf{F}(\mathbf{A})}{\partial \operatorname{vec}' \bar{\mathbf{A}}} \cdot \mathbf{D}_m,$$

gdzie  $\bar{\mathbf{A}}$  jest tym samym co  $\mathbf{A}$  lecz z zastosowaniem reguł różniczkowania dla macierzy niesymetrycznych.

W pracy będziemy również wykorzystywać własności operatorów śladów częściowych oraz śladu blokowego kwadratowej macierzy blokowej, przedstawiane przez Filipiak i in. [14] dla macierzy prostokątnych.

**Definicja 1.1.** Dla dowolnej macierzy blokowej  $\mathbf{A} = (\mathbf{A}_{ij})$  stopnia  $qp$ , gdzie  $\mathbf{A}_{ij}$  są blokami stopnia  $q$

- (i) operator **śladów częściowych**,  $\text{PTr}_q : \mathbb{R}_{qp \times qp} \rightarrow \mathbb{R}_{p \times p}$ , przekształca macierz  $\mathbf{A}$  w macierz stopnia  $p$ , w której bloki  $\mathbf{A}_{ij}$  zostały zastąpione przez swoje ślady:

$$\text{PTr}_q \mathbf{A} = (\text{tr } \mathbf{A}_{ij}), \quad i, j = 1, \dots, p,$$

- (ii) operator **śladu blokowego**,  $\text{BTr}_q : \mathbb{R}_{qp \times qp} \rightarrow \mathbb{R}_{q \times q}$ , przekształca macierz  $\mathbf{A}$  w macierz stopnia  $q$  powstałą z sumowania bloków diagonalnych:

$$\text{BTr}_q \mathbf{A} = \sum_{i=1}^p \mathbf{A}_{ii}.$$

Wybrane własności operatorów śladów częściowych oraz śladu blokowego pochodzą z pracy Filipiak i in. [14].

**Lemat 1.10.** Dla dowolnych macierzy  $\mathbf{A}$  stopnia  $pq$  oraz  $\mathbf{B}$  stopnia  $p$  zachodzą następujące własności:

$$\begin{aligned} \text{vec} \{ \text{PTr}_p [(\mathbf{I}_q \otimes \mathbf{B}') \mathbf{A}] \} &= (\mathbf{I}_{q^2} \otimes \text{vec}' \mathbf{B}) (\mathbf{I}_q \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_p) \text{vec } \mathbf{A}, \\ \text{vec} \{ \text{BTr}_q [(\mathbf{B}' \otimes \mathbf{I}_q) \mathbf{A}] \} &= (\text{vec}' \mathbf{B} \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec } \mathbf{A}. \end{aligned}$$

**Lemat 1.11.** Dla dowolnych macierzy  $\mathbf{A}$  stopnia  $pq$  oraz  $\mathbf{B}$  stopnia  $q$  zachodzą następujące własności:

$$\begin{aligned} \text{PTr}_p [(\mathbf{B} \otimes \mathbf{I}_p) \mathbf{A}] &= \mathbf{B} \cdot \text{PTr}_p (\mathbf{A}), \\ \text{BTr}_q [(\mathbf{I}_p \otimes \mathbf{B}) \mathbf{A}] &= \mathbf{B} \cdot \text{BTr}_q (\mathbf{A}). \end{aligned}$$

**Lemat 1.12.** Dla dowolnych macierzy  $\mathbf{A}$  stopnia  $pq$  oraz  $\mathbf{B}$  stopnia  $q$  i  $\mathbf{C}$  stopnia  $p$  zachodzą następujące własności:

$$\begin{aligned} \text{PTr}_q [\mathbf{A} (\mathbf{I}_p \otimes \mathbf{B})] &= \text{PTr}_q [(\mathbf{I}_p \otimes \mathbf{B}) \mathbf{A}], \\ \text{BTr}_q [\mathbf{A} (\mathbf{C} \otimes \mathbf{I}_q)] &= \text{BTr}_q [(\mathbf{C} \otimes \mathbf{I}_q) \mathbf{A}]. \end{aligned}$$

**Lemat 1.13.** Dla dowolnej macierzy  $\mathbf{A}$  stopnia  $pq$  zachodzą następujące własności:

$$\begin{aligned} \text{tr } \mathbf{A} &= \text{tr} (\text{PTr}_p \mathbf{A}) = \text{tr} (\text{PTr}_q \mathbf{A}) \\ \text{tr } \mathbf{A} &= \text{tr} (\text{BTr}_q \mathbf{A}) = \text{tr} (\text{BTr}_p \mathbf{A}). \end{aligned}$$

**Lemat 1.14.** *Dla dowolnych macierzy  $\mathbf{A}$  stopnia  $pq$  zachodzą następujące własności:*

- (i) jeśli  $\mathbf{A}$  jest macierzą symetryczną to operatory śladów częściowych i śladu blokowego są macierzami symetrycznymi,*
- (ii) jeśli  $\mathbf{A}$  jest macierzą określoną dodatnio to operatory śladów częściowych i śladu blokowego są macierzami określonymi dodatnio.*

## 2. Aproksymacja entropijną funkcją straty

W kolejnych podrozdziałach poszukiwać będziemy minimum entropijnej funkcji straty zdefiniowanej wzorem (1.3) w zbiorach  $\mathcal{S}$ ,  $\mathcal{S}_{CS}$  i  $\mathcal{S}_{AR}$  odpowiednio.

### 2.1. Rozdzielna struktura kowariancyjna

Dla dowolnej określonej dodatnio macierzy  $\Omega$  entropijna funkcja straty dla struktury  $\Psi \otimes \Sigma \in \mathcal{S}$  ma postać

$$f_E(\Omega, \Psi \otimes \Sigma) = \text{tr} [\Omega^{-1} (\Psi \otimes \Sigma)] - \ln |\Omega^{-1} (\Psi \otimes \Sigma)| - pq. \quad (2.1)$$

Najlepsze przybliżenie macierzy  $\Omega$  przez macierz o strukturze  $\Psi \otimes \Sigma$  podane jest w następującym twierdzeniu.

**Twierdzenie 2.1** (Filipiak i in. [9]). *Dla danej symetrycznej, określonej dodatnio macierzy  $\Omega$ , istnieje symetryczna, określona dodatnio macierz  $\Psi \otimes \Sigma$ , która minimalizuje entropijną funkcję straty (2.1) na zbiorze  $\mathcal{S}$  i minimum to jest osiągnięte dla  $\Psi, \Sigma$  spełniających układ równań*

$$\begin{cases} q \Psi^{-1} = \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] \\ p \Sigma^{-1} = \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}]. \end{cases} \quad (2.2)$$

*Dowód.* Zróżniczkujemy  $f_E(\Omega, \Psi \otimes \Sigma)$  względem  $\Psi$  oraz  $\Sigma$ . Stosując regułę łańcuchową różniczkowania funkcji złożonej macierzy symetrycznej

$$\begin{aligned} \frac{\partial f_E}{\partial \Psi} &= \frac{\partial f_E}{\partial (\Psi \otimes \Sigma)} \cdot \frac{\partial (\Psi \otimes \Sigma)}{\partial \Psi} \\ \frac{\partial f_E}{\partial \Sigma} &= \frac{\partial f_E}{\partial (\Psi \otimes \Sigma)} \cdot \frac{\partial (\Psi \otimes \Sigma)}{\partial \Sigma} \end{aligned}$$

(por. Magnus i Neudecker [24], Filipiak i in. [12]) oraz korzystając ze wzorów opisanych w Lematach 1.8 i 1.9 otrzymujemy

$$\begin{aligned} \frac{\partial f_E}{\partial \Psi} &= \text{vec}' (\Omega^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \cdot \mathbf{D}_p \\ &\quad - \text{vec}' (\Psi^{-1} \otimes \Sigma^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \cdot \mathbf{D}_p \\ &= \left[ (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} (\Omega^{-1} - \Psi^{-1} \otimes \Sigma^{-1})' \right]' \cdot \mathbf{D}_p, \\ \frac{\partial f_E}{\partial \Sigma} &= \text{vec}' (\Omega^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\text{vec } \Psi \otimes \mathbf{I}_{q^2}) \cdot \mathbf{D}_q \\ &\quad - \text{vec}' (\Psi^{-1} \otimes \Sigma^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\text{vec } \Psi \otimes \mathbf{I}_{q^2}) \cdot \mathbf{D}_q \\ &= \left[ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} (\Omega^{-1} - \Psi^{-1} \otimes \Sigma^{-1})' \right]' \cdot \mathbf{D}_q. \end{aligned}$$

Przyrównując powyższe pochodne do zera możemy zapisać

$$\begin{cases} (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} (\Omega^{-1} - \Psi^{-1} \otimes \Sigma^{-1}) = \mathbf{0} \\ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} (\Omega^{-1} - \Psi^{-1} \otimes \Sigma^{-1}) = \mathbf{0}. \end{cases}$$

Ponieważ macierze duplikacji powodują jedynie powielenie wybranych składowych wektorów pochodnych, nie muszą one być uwzględnione przy przyrównywaniu do zera.

Używając operatorów śladów częściowych oraz śladu blokowego zaprezentowanych w Definicji 1.1 oraz ich własności podanych w Lematach 1.10 – 1.14 można zapisać następujące przekształcenia powyższego układu równań:

$$\begin{cases} \text{vec} \left\{ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma') \cdot (\Omega^{-1} - \Psi^{-1} \otimes \Sigma^{-1})] \right\} = \mathbf{0} \\ \text{vec} \left\{ \text{BTr}_q [(\Psi' \otimes \mathbf{I}_q) \cdot (\Omega^{-1} - \Psi^{-1} \otimes \Sigma^{-1})] \right\} = \mathbf{0} \\ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] - \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) (\Psi^{-1} \otimes \Sigma^{-1})] = \mathbf{0} \\ \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}] - \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) (\Psi^{-1} \otimes \Sigma^{-1})] = \mathbf{0} \\ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] = \text{PTr}_q [\Psi^{-1} \otimes \mathbf{I}_q] \\ \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}] = \text{BTr}_q [\mathbf{I}_p \otimes \Sigma^{-1}] \\ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] = \text{PTr}_q [(\Psi^{-1} \otimes \mathbf{I}_q) \cdot \mathbf{I}_{pq}] \\ \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}] = \text{BTr}_q [(\mathbf{I}_p \otimes \Sigma^{-1}) \cdot \mathbf{I}_{pq}] \\ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] = \Psi^{-1} \cdot \text{PTr}_q (\mathbf{I}_{pq}) \\ \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}] = \Sigma^{-1} \cdot \text{BTr}_q (\mathbf{I}_{pq}) \\ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] = \Psi^{-1} \cdot q \cdot \mathbf{I}_p \\ \text{BTr}_q [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}] = \Sigma^{-1} \cdot p \cdot \mathbf{I}_q \end{cases}$$

co prowadzi do (2.2).

Zauważmy, że układ ten, na podstawie Lematu 1.12, można zapisać w postaci

$$\begin{cases} \Psi^{-1} = \frac{1}{q} \text{PTr}_q [(\mathbf{I}_p \otimes \mathbf{Z}'_1) \Omega^{-1} (\mathbf{I}_p \otimes \mathbf{Z}_1)] \\ \Sigma^{-1} = \frac{1}{p} \text{BTr}_q [(\mathbf{Z}'_2 \otimes \mathbf{I}_q) \Omega^{-1} (\mathbf{Z}_2 \otimes \mathbf{I}_q)], \end{cases}$$

gdzie  $\mathbf{Z}_1$  i  $\mathbf{Z}_2$  spełniają  $\mathbf{Z}_1 \mathbf{Z}'_1 = \Sigma$  i  $\mathbf{Z}_2 \mathbf{Z}'_2 = \Psi$ . Ponieważ argumenty operatorów śladów częściowych i śladu blokowego są symetryczne i określone dodatnio, zatem, na podstawie Lematu 1.14, macierze  $\Psi$  i  $\Sigma$  spełniające powyższy układ równań są symetryczne i określone dodatnio. ■

Rozwiązanie  $\check{\Psi}$  i  $\check{\Sigma}$  spełniające układ równań (2.2) może być traktowane jako estymator macierzy kowariancji o strukturze rozdzielnej. Estymator ten zapisywać będziemy w skrócie ELE (ang. *entropy loss estimator*).

Innym, powszechnie stosowanym estymatorem  $\Psi \otimes \Sigma$  jest estymator największej wiarygodności (ang. *maximum likelihood estimator*, MLE) postaci

$$\begin{cases} q \hat{\Psi} &= \text{PTr}_q \left[ \left( \mathbf{I}_p \otimes \hat{\Sigma}^{-1} \right) \Omega \right] \\ p \hat{\Sigma} &= \text{BTr}_q \left[ \left( \hat{\Psi}^{-1} \otimes \mathbf{I}_q \right) \Omega \right]; \end{cases} \quad (2.3)$$

por. Filipiak i in. [10]. Porównując układy równań (2.2) i (2.3) można zauważyć, że macierze  $\Omega$  występujące w (2.2) zastąpione zostały w (2.3) przez odwrotności oraz estymatory  $\Psi$  i  $\Sigma$  występujące w (2.2) zastąpione zostały w (2.3) przez odpowiednie odwrotności. Istnieje zatem relacja między estymatorami ELE i MLE. Przypomnijmy, że logarytm funkcji wiarygodności wykorzystywany do wyznaczenia MLE ma dla struktury rozdzielnej postać

$$\ln L(\Omega, \Psi \otimes \Sigma) = -\frac{npq}{2} \ln(2\pi) - \frac{n}{2} \ln |\Psi \otimes \Sigma| - \frac{n}{2} \text{tr} \left[ (\Psi^{-1} \otimes \Sigma^{-1}) \Omega \right]. \quad (2.4)$$

Zastępując macierze  $\Omega$  i  $\Psi \otimes \Sigma$  we wzorze (2.1) przez odpowiednie odwrotności, entropijną funkcję straty można wyrazić za pomocą funkcji wiarygodności odpowiednich argumentów w następujący sposób:

$$\begin{aligned} \ln f_E(\Omega^{-1}, \Psi^{-1} \otimes \Sigma^{-1}) &= \text{tr} \left[ \Omega (\Psi^{-1} \otimes \Sigma^{-1}) \right] - \ln |\Omega (\Psi^{-1} \otimes \Sigma^{-1})| - pq \\ &= \text{tr} \left[ (\Psi^{-1} \otimes \Sigma^{-1}) \Omega \right] + \ln |\Psi \otimes \Sigma| + pq \ln(2\pi) \\ &\quad - pq \ln(2\pi) - \ln |\Omega| - pq \\ &= -\frac{2}{n} \ln L(\Omega, \Psi \otimes \Sigma) - \ln |\Omega| - pq[1 + \ln(2\pi)]. \end{aligned} \quad (2.5)$$

Oczywiście ELE nie jest odwrotnością MLE, jednakże sposób rozwiązywania układu równań (2.2) jest taki sam jak układu (2.3).

Jak zostało wspomniane we Wprowadzeniu, Lu i Zimmerman [23] podali układ równań prowadzący do wyznaczenia MLE równoważny (2.3). Ponieważ algorytm wyznaczania MLE prezentowany w pracy Lu i Zimmerman [23] jest zawsze zbieżny, również algorytm rozwiązujący (2.3) jest zbieżny, jednakże bardzo trudno jest pokazać analitycznie, że zbiega on do MLE. Lu i Zimmerman [22] pokazali, że dla próby o rozmiarze 4 algorytm może zbiegać do kilku różnych estymatorów w zależności od punktu początkowego, ale wartość funkcji dla każdego z nich jest taka sama, podczas gdy dla większej próby wydaje się, że algorytm zbiega do jednego rozwiązania

niezależnie od punktu początkowego. Ze względu na relację (2.5) między entropijną funkcją straty a funkcją wiarygodności oczekiwać można takiego samego zachowania w przypadku rozwiązywania układu równań (2.2).

Warto zwrócić uwagę, że układ równań (2.2) nie jest podany w postaci jawnej, ale układ ten może być rozwiązany iteracyjnie za pomocą algorytmu nazywanego „flip-flop”: w pierwszym kroku ustalamy jeden z czynników iloczynu Kroneckera na przykład  $\Psi = \mathbf{I}_p$  i wstawiając do drugiego równania układu (2.2) otrzymujemy pierwsze przybliżenie  $\Sigma$ , które następnie wstawiamy do pierwszego równania w celu otrzymania pierwszego przybliżenia macierzy  $\Psi$ . Algorytm jest kontynuowany tak długo, aż różnice między kolejnymi przybliżeniami (w sensie pewnej funkcji) będą wystarczająco małe.

Z pracy Stein [39] wiadomo, że entropijna funkcja straty jest wypukła ze względu na macierz  $\Gamma$  należącą do wypukłego zbioru macierzy określonych dodatnio. Zatem w zbiorze tym istnieje minimum lokalne, które jest jednocześnie minimum globalnym. Załóżmy teraz, że  $\Gamma \in \mathcal{S}$ . Jeżeli ustalimy jeden z czynników iloczynu Kroneckera, to mówimy o *profilowej* entropijnej funkcji straty, która podobnie jak poprzednio jest funkcją wypukłą ze względu na drugi czynnik. Zatem dla ustalonej macierzy  $\Psi$  (odpowiednio  $\Sigma$ ) profilowa entropijna funkcja straty jest wypukła ze względu na  $\Sigma$  (odpowiednio  $\Psi$ ), a więc istnieją lokalne minima funkcji profilowych. Ze względu na konstrukcję algorytmu flip-flop, minima te prowadzą do minimum lokalnego z  $L_{\Omega}^{(E)}(\Psi \otimes \Sigma)$ .

## 2.2. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze kompletnie symetrycznej

Dla dowolnej określonej dodatnio macierzy  $\Omega$  entropijna funkcja straty dla struktury  $\Psi_{CS} \otimes \Sigma \in \mathcal{S}_{CS}$  ma postać

$$f_E(\Omega, \Psi_{CS} \otimes \Sigma) = \text{tr} [\Omega^{-1} (\Psi_{CS} \otimes \Sigma)] - \ln |\Omega^{-1} (\Psi_{CS} \otimes \Sigma)| - pq. \quad (2.6)$$

Najlepsze przybliżenie macierzy  $\Omega$  przez macierz o strukturze  $\Psi_{CS} \otimes \Sigma$  podane jest w następującym twierdzeniu.

**Twierdzenie 2.2** (Filipiak i in. [10]). *Dla danej symetrycznej, określonej dodatnio macierzy  $\Omega$ , istnieje wyznaczona jednoznacznie symetryczna, określona dodatnio*

macierz  $\Psi_{\text{CS}} \otimes \Sigma$ , która minimalizuje entropijną funkcję straty postaci (2.6) w zbiorze  $\mathcal{S}_{\text{CS}}$  i minimum to jest osiągnięte dla  $\varrho$ ,  $\Sigma$  spełniających układ równań

$$\begin{cases} \varrho = \frac{-(p-2)\alpha - pq(p-1) + \sqrt{[(p-2)\alpha + pq(p-1)]^2 + 4(p-1)\alpha^2}}{-2(p-1)\alpha} \\ p \Sigma^{-1} = \text{BTr}_q [(\Psi_{\text{CS}} \otimes \mathbf{I}_q) \Omega^{-1}], \end{cases} \quad (2.7)$$

gdzie  $\alpha = \text{tr} \{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \Sigma] \Omega^{-1} \}$ .

*Dowód.* Różniczkując (2.6) ze względu na  $\varrho$  (por. Magnus i Neudecker [24], Fackler [6] lub Wprowadzenie) z zastosowaniem reguły łańcuchowej różniczkowania funkcji złożonej postaci

$$\frac{\partial f_E}{\partial \varrho} = \frac{\partial f_E}{\partial (\Psi_{\text{CS}} \otimes \Sigma)} \cdot \frac{\partial (\Psi_{\text{CS}} \otimes \Sigma)}{\partial \Psi_{\text{CS}}} \cdot \frac{\partial \Psi_{\text{CS}}}{\partial \varrho}$$

otrzymujemy

$$\begin{aligned} \frac{\partial f_E}{\partial \varrho} &= \text{vec}'(\Omega^{-1}) \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \text{vec}(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \\ &\quad - \text{vec}'(\Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \text{vec}(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \\ &= \text{vec}'[\Omega^{-1} - (\Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})]' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \text{vec}(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \\ &= \text{vec}'(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec}(\Omega^{-1} - \Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1}). \end{aligned}$$

Wyznaczenie pochodnej ze względu na  $\Sigma$  przebiega analogicznie jak w rozdziale 2.1, gdzie  $\Psi$  zostało zastąpione przez  $\Psi_{\text{CS}}$ , więc zostanie pominięte.

Stosując własności operatora śladów częściowych (Lemat 1.10 i 1.11) oraz własności śladu macierzy podanych w Lematach 1.2 i 1.3 możemy zapisać następujące przekształcenia pochodnej funkcji  $f_E$  ze względu na  $\varrho$ :

$$\begin{aligned} \frac{\partial f_E}{\partial \varrho} &= \text{vec}'(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \cdot \text{vec} \{ \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma') \cdot (\Omega^{-1} - \Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})] \} \\ &= \text{tr} \left\{ (\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \text{PTr}_q [(\mathbf{I}_p \otimes \Sigma) \cdot (\Omega^{-1} - \Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})] \right\} \\ &= \text{tr} \left[ \text{PTr}_q \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p)' \otimes \mathbf{I}_q] \cdot (\mathbf{I}_p \otimes \Sigma') \cdot (\Omega^{-1} - \Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})' \right\} \right] \\ &= \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p)' \otimes \mathbf{I}_q] \cdot (\mathbf{I}_p \otimes \Sigma') \cdot (\Omega^{-1} - \Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})' \right\} \\ &= \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \Sigma'] \cdot (\Omega^{-1})' - [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \Sigma'] (\Psi_{\text{CS}}^{-1} \otimes \Sigma^{-1})' \right\} \\ &= \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \Sigma] \cdot \Omega^{-1} \right\} - \text{tr} [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \Psi_{\text{CS}}^{-1} \otimes \mathbf{I}_q] \\ &= \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \Sigma] \cdot \Omega^{-1} \right\} - \text{tr} [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \cdot \Psi_{\text{CS}}^{-1}] \cdot \text{tr}(\mathbf{I}_q) \\ &= \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \Sigma] \cdot \Omega^{-1} \right\} - q \cdot [\text{tr}(\mathbf{1}_p \mathbf{1}'_p \Psi_{\text{CS}}^{-1}) - \text{tr}(\Psi_{\text{CS}}^{-1})]. \end{aligned}$$

Przyrównując powyższą pochodną do zera możemy zapisać

$$\text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \boldsymbol{\Sigma}'] \cdot \boldsymbol{\Omega}^{-1} \right\} - q \cdot \text{tr} (\mathbf{1}_p \mathbf{1}'_p \boldsymbol{\Psi}_{\text{CS}}^{-1} - \boldsymbol{\Psi}_{\text{CS}}^{-1}) = 0.$$

Zapisując odwrotność macierzy o strukturze kompletnej symetrii jako

$$\boldsymbol{\Psi}_{\text{CS}}^{-1} = \frac{1}{1-\varrho} \left[ \mathbf{I}_p - \frac{\varrho}{1+(p-1)\varrho} \mathbf{1}_p \mathbf{1}'_p \right]$$

można obliczyć

$$\begin{aligned} \text{tr} (\mathbf{1}_p \mathbf{1}'_p \boldsymbol{\Psi}_{\text{CS}}^{-1}) &= \text{tr} \left\{ \mathbf{1}_p \mathbf{1}'_p \frac{1}{1-\varrho} \left[ \mathbf{I}_p - \frac{\varrho}{1+(p-1)\varrho} \mathbf{1}_p \mathbf{1}'_p \right] \right\} = \frac{1}{1-\varrho} \left[ p - \frac{p^2 \varrho}{1+(p-1)\varrho} \right] = \frac{p}{1+(p-1)\varrho}, \\ \text{tr} (\boldsymbol{\Psi}_{\text{CS}}^{-1}) &= \frac{1}{1-\varrho} \left[ \text{tr} (\mathbf{I}_p) - \frac{\varrho}{1+(p-1)\varrho} \cdot \text{tr} (\mathbf{1}_p \mathbf{1}'_p) \right] = \frac{1}{1-\varrho} \left[ p - \frac{p\varrho}{1+(p-1)\varrho} \right] \\ &= \frac{p[1+(p-2)\varrho]}{(1-\varrho)[1+(p-1)\varrho]}. \end{aligned}$$

Podsumowując, równania pozwalające wyznaczyć minimum entropijnej funkcji straty w zbiorze  $\mathcal{S}_{\text{CS}}$  można zapisać następująco:

$$\begin{cases} -(p-1)\alpha\varrho^2 + [(p-2)\alpha + pq(p-1)]\varrho + \alpha = 0 \\ p \boldsymbol{\Sigma}^{-1} = \text{BTr}_q [(\boldsymbol{\Psi}_{\text{CS}} \otimes \mathbf{I}_q) \boldsymbol{\Omega}^{-1}], \end{cases} \quad (2.8)$$

gdzie  $\alpha = \text{tr} \{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \boldsymbol{\Sigma}] \boldsymbol{\Omega}^{-1} \}$  oraz, podobnie jak w Rozdziale 2.1, rozwiązanie drugiego równania jest symetryczne i określone dodatnio.

Należy teraz pokazać, że pierwsze równanie ma pierwiastek należący do przedziału  $\left(-\frac{1}{p-1}, 1\right)$ , co zagwarantuje, że rozwiązanie układu (2.8) należy do  $\mathcal{S}_{\text{CS}}$ . Oznaczmy więc  $-(p-1)\alpha\varrho^2 + [(p-2)\alpha + pq(p-1)]\varrho + \alpha$  przez  $h(\varrho)$ . Ponieważ  $h\left(-\frac{1}{p-1}\right) = -pq < 0$  oraz  $h(1) = pq(p-1) > 0$ , funkcja  $h(\varrho)$  posiada pierwiastek w przedziale  $\left(-\frac{1}{p-1}, 1\right)$  minimalizujący funkcję (2.6) ze względu na  $\varrho$ . Zauważmy ponadto, że dla  $p \geq 2$  wyróżnik jest zawsze dodatni więc funkcja  $h(\varrho)$  ma dwa pierwiastki. Dla  $\alpha > 0$  ( $\alpha < 0$ ) trójmian jest wklęsły (wypukły) i rozwiązanie  $\check{\varrho}$  jest określone przez mniejszy (większy) pierwiastek równania kwadratowego. W obu przypadkach  $\check{\varrho}$  ma taką samą postać. Stąd najlepsze przybliżenie macierzy  $\boldsymbol{\Omega}$  może być wyznaczone poprzez rozwiązanie układu równań (2.7). ■

W poprzednim rozdziale pokazaliśmy związek (2.5) między entropijną funkcją straty a funkcją wiarygodności dla struktury  $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ . Można by zatem oczekiwać,

że relacja ta nadal jest spełniona, gdy  $\Psi = \Psi_{CS}$ . Estymatory MLE otrzymujemy rozwiązując układ równań

$$\begin{cases} (p-1)k_0\varrho^3 + [k_0 - (p-1)k_0 + (p-1)^2a - (p-1)b] \varrho^2 + [2(p-1)a - k_0] \varrho + (a-b) = 0 \\ p\Sigma = \text{BTr}_q [(\Psi_{CS}^{-1} \otimes \mathbf{I}_q)\Omega] \end{cases} \quad (2.9)$$

gdzie

$$a = \text{tr} [(\mathbf{I}_p \otimes \Sigma^{-1})\Omega], \quad b = \text{tr} [(\mathbf{1}_p \mathbf{1}'_p \otimes \Sigma^{-1})\Omega], \quad k_0 = nq(p-1)p;$$

por. Filipiak i in. [13].

Zwróćmy uwagę, że zarówno w układzie (2.8) jak i (2.9) pierwsze równanie zależy od  $\Sigma$ , a drugie równanie zależy od  $\varrho$ . W związku z tym układy te można rozwiązać iteracyjnie za pomocą algorytmu flip-flop.

Porównując (2.8) z (2.9) widoczna jest podobna relacja między estymatorami ELE i MLE macierzy  $\Sigma$  jak w przypadku struktury rozdzielnej. Żeby zauważyć podobieństwo między estymatorami  $\varrho$  warto zastosować inne podejście do minimalizacji entropijnej funkcji straty / maksymalizacji funkcji wiarygodności.

Dla dowolnej macierzy  $\mathbf{X}$  stopnia  $pq$  zachodzi związek  $\text{tr} \mathbf{X} = \text{tr}(\text{BTr}_q \mathbf{X})$ ; por. Lemat 1.13. Zatem, dla  $\Psi = \Psi_{CS}$  entropijna funkcja straty może być zapisana w postaci

$$f_E(\Omega, \Psi_{CS} \otimes \Sigma) = \text{tr} \{ \text{BTr}_q [\Omega^{-1}(\Psi_{CS} \otimes \Sigma)] \} - \ln |\Omega^{-1}(\Psi_{CS} \otimes \Sigma)| - pq,$$

a następnie, używając Lematu 1.11,

$$f_E(\Omega, \Psi_{CS} \otimes \Sigma) = \text{tr} \{ \Sigma \cdot \text{BTr}_q [\Omega^{-1}(\Psi_{CS} \otimes \mathbf{I}_q)] \} - \ln |\Omega^{-1}(\Psi_{CS} \otimes \Sigma)| - pq. \quad (2.10)$$

Stosując rozkład spektralny macierzy  $\Psi_{CS}$  postaci

$$\Psi_{CS} = c_1 \mathbf{P}_p + c_2 \mathbf{Q}_p, \quad (2.11)$$

gdzie  $c_1 = 1 + (p-1)\varrho$  oraz  $c_2 = 1 - \varrho$ , operator śladu blokowego użyty w (2.10) może być zapisany jako

$$\text{BTr}_q [\Omega^{-1}(\Psi_{CS} \otimes \mathbf{I}_q)] = c_1 \text{BTr}_q [(\mathbf{P}_p \otimes \mathbf{I}_q) \Omega^{-1}] + c_2 \text{BTr}_q [(\mathbf{Q}_p \otimes \mathbf{I}_q) \Omega^{-1}].$$

Dla skrócenia zapisów wprowadźmy oznaczenia

$$\mathbf{C} = \text{BTr}_q [(\mathbf{P}_p \otimes \mathbf{I}_q) \Omega^{-1}], \quad \mathbf{D} = \text{BTr}_q [(\mathbf{Q}_p \otimes \mathbf{I}_q) \Omega^{-1}]. \quad (2.12)$$

Ponieważ  $|\Psi_{\text{CS}}| = c_1 c_2^{(p-1)}$ , entropijna funkcja straty może być zapisana jako

$$\begin{aligned} f_E(\Omega, \Psi_{\text{CS}} \otimes \Sigma) \\ = c_1 \text{tr}(\Sigma \mathbf{C}) + c_2 \text{tr}(\Sigma \mathbf{D}) + \ln |\Omega| - q \ln(c_1) - q(p-1) \ln(c_2) - p \ln |\Sigma| - pq. \end{aligned} \quad (2.13)$$

Różniczkując powyższą funkcję ze względu na  $\Sigma$  i przyrównując otrzymaną pochodną do zera otrzymujemy

$$p \check{\Sigma}^{-1} = c_1 \mathbf{C} + c_2 \mathbf{D} \quad (2.14)$$

lub równoważnie

$$\check{\Sigma} = p \text{BTr}_q^{-1} [(\Psi_{\text{CS}} \otimes \mathbf{I}_q) \Omega^{-1}], \quad (2.15)$$

gdzie  $\text{BTr}_q^{-1}(\bullet)$  oznacza odwrotność  $\text{BTr}_q(\bullet)$ . Istnienie ostatniej odwrotności jest zagwarantowane przez dodatnią określoność operatora śladu blokowego; por. Filipiak i in. [14] lub Lemat 1.14.

Różniczkując teraz entropijną funkcję straty postaci (2.13) ze względu na  $\varrho$  i przyrównując otrzymane wyrażenie do zera możemy zapisać

$$(p-1) \text{tr}(\Sigma \mathbf{C}) - \text{tr}(\Sigma \mathbf{D}) - \frac{q(p-1)}{c_1} + \frac{q(p-1)}{c_2} = 0.$$

Po wstawieniu (2.15) do powyższego równania otrzymujemy równość

$$q(p-1)\varrho + c_1 c_2 (p-1) \text{tr} [(c_1 \mathbf{C} + c_2 \mathbf{D})^{-1} \mathbf{C}] - c_1 c_2 \text{tr} [(c_1 \mathbf{C} + c_2 \mathbf{D})^{-1} \mathbf{D}] = 0, \quad (2.16)$$

która nie zależy od  $\Sigma$ . Należy teraz pokazać, że rozwiązanie równania (2.16) należy do przedziału  $(-1/(p-1), 1)$ . W tym celu oznaczmy lewą stronę tego równania przez  $\ell(\varrho)$ . Wówczas  $\ell(-\frac{1}{p-1}) = -q < 0$  oraz  $\ell(1) = q(p-1) > 0$ . Stąd, co najmniej jedno rozwiązanie  $\ell(\varrho) = 0$  należy do wymaganego przedziału. Aby otrzymać estymator macierzy  $\Sigma$  wystarczy wstawić rozwiązanie  $\check{\varrho}$  do równania (2.15).

Warto zauważyć, że może istnieć kilka pierwiastków równania (2.16). W takim przypadku wybieramy ten, dla którego wartość  $f_E(\Omega, \Psi_{\text{CS}} \otimes \Sigma)$  jest najmniejsza. Ponieważ dla ustalonego  $\varrho$  (odpowiednio  $\Sigma$ ) profilowa entropijna funkcja straty jest wypukła, podobnie jak w Rozdziale 2.1, punkt stacjonarny uzyskany za pomocą algorytmu flip-flop jest minimum lokalnym z  $L_{\Omega}^{(E)}(\Psi_{\text{CS}} \otimes \Sigma)$ .

Podsumowując, najlepsze przybliżenie macierzy  $\Omega$  strukturą  $\Psi_{\text{CS}} \otimes \Sigma$  można wyznaczyć rozwiązując następujący układ równań

$$\begin{cases} q(p-1)\varrho + c_1 c_2 (p-1) \text{tr} [(c_1 \mathbf{C} + c_2 \mathbf{D})^{-1} \mathbf{C}] - c_1 c_2 \text{tr} [(c_1 \mathbf{C} + c_2 \mathbf{D})^{-1} \mathbf{D}] = 0 \\ \Sigma = p \text{BTr}_q^{-1} [(\Psi_{\text{CS}} \otimes \mathbf{I}_q) \Omega^{-1}], \end{cases} \quad (2.17)$$

który jest równoważny układowi (2.7). Łatwo zauważyć, że drugie równania układów (2.7) i (2.17) reprezentują ten sam estymator macierzy  $\Sigma$ . Ponadto, podstawiając (2.14) możemy zapisać następujące przekształcenia pierwszego równania układu (2.17):

$$\begin{aligned} q(p-1)\varrho + c_1c_2(p-1) \operatorname{tr} \left( \frac{1}{p} \Sigma \mathbf{C} \right) - c_1c_2 \operatorname{tr} \left( \frac{1}{p} \Sigma \mathbf{D} \right) &= 0 \\ q(p-1)\varrho + c_1c_2 \operatorname{tr} (\Sigma \mathbf{C}) - \frac{1}{p}c_1c_2 \operatorname{tr} (\Sigma \mathbf{C}) - \frac{1}{p}c_1c_2 \operatorname{tr} (\Sigma \mathbf{D}) &= 0 \\ q(p-1)\varrho + c_1c_2 \operatorname{tr} (\Sigma \mathbf{C}) - \frac{1}{p}c_1c_2 \operatorname{tr} [\Sigma (\mathbf{C} - \mathbf{D})] &= 0. \end{aligned}$$

Podstawiając (2.12) oraz stosując Lematy 1.11 i 1.13, występujące w powyższym równaniu ślady macierzy możemy zapisać jako

$$\begin{aligned} \operatorname{tr} (\Sigma \mathbf{C}) &= \operatorname{tr} \{ \Sigma \cdot \operatorname{BTr}_q [(\mathbf{P}_p \otimes \mathbf{I}_q) \Omega^{-1}] \} = \operatorname{tr} \{ \operatorname{BTr}_q [(\mathbf{P}_p \otimes \Sigma) \Omega^{-1}] \} \\ &= \operatorname{tr} [(\mathbf{P}_p \otimes \Sigma) \Omega^{-1}] = \operatorname{tr} \left[ \left( \frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \otimes \Sigma \right) \Omega^{-1} \right], \\ \operatorname{tr} [\Sigma (\mathbf{C} - \mathbf{D})] &= \operatorname{tr} \{ \Sigma \cdot \operatorname{BTr}_q \{ [(\mathbf{P}_p - \mathbf{Q}_p) \otimes \mathbf{I}_q] \Omega^{-1} \} \} \\ &= \operatorname{tr} \{ \Sigma \cdot \operatorname{BTr}_q [(\mathbf{I}_p \otimes \mathbf{I}_q) \Omega^{-1}] \} = \operatorname{tr} \{ \operatorname{BTr}_q [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] \} \\ &= \operatorname{tr} [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}], \end{aligned}$$

a samo równanie w postaci

$$q(p-1)\varrho + \frac{1}{p}c_1c_2 \operatorname{tr} [(\mathbf{1}_p \mathbf{1}_p' \otimes \Sigma) \Omega^{-1}] - \frac{1}{p}c_1c_2 \operatorname{tr} [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] = 0$$

i ostatecznie

$$q(p-1)\varrho + \frac{1}{p}c_1c_2 \operatorname{tr} \{ [(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \otimes \Sigma] \Omega^{-1} \} = 0.$$

Podstawiając  $c_1$  i  $c_2$  zdefiniowane pod wzorem (2.11) otrzymujemy pierwsze równanie układu (2.8), którego rozwiązaniem jest  $\varrho$  występujące w układzie (2.7). Zatem układ (2.17) jest równoważny układowi (2.7).

Z pracy Filipiak i in. [10] wiadomo, że stosując podobną procedurę do maksymalizacji funkcji wiarygodności otrzymamy następujący układ równań, równoważny (2.9), pozwalający na wyznaczenie MLE:

$$\begin{cases} q(p-1)\varrho + (p-1)c_2^2 \operatorname{tr} [(c_2 \mathbf{A} + c_1 \mathbf{B})^{-1} \mathbf{A}] - c_1^2 \operatorname{tr} [(c_2 \mathbf{A} + c_1 \mathbf{B})^{-1} \mathbf{B}] = 0 \\ \Sigma = \frac{1}{p} \operatorname{BTr}_q [(\Psi_{\text{CS}}^{-1} \otimes \mathbf{I}_q) \Omega], \end{cases} \quad (2.18)$$

gdzie  $\mathbf{A} = \operatorname{BTr}_q [(\mathbf{P}_p \otimes \mathbf{I}_q) \Omega]$  i  $\mathbf{B} = \operatorname{BTr}_q [(\mathbf{Q}_p \otimes \mathbf{I}_q) \Omega]$ .

Tabela 2.1 podsumowuje otrzymane układy równań dla minimalizacji entropijnej funkcji straty i maksymalizacji funkcji wiarygodności.

Tabela 2.1. Układy równań pozwalające wyznaczyć estymatory ELE oraz MLE.

ELE	$\mathbf{C} = \text{BTr}_q [(\mathbf{P}_p \otimes \mathbf{I}_q) \boldsymbol{\Omega}^{-1}]$ $\mathbf{D} = \text{BTr}_q [(\mathbf{Q}_p \otimes \mathbf{I}_q) \boldsymbol{\Omega}^{-1}]$ $\boldsymbol{\Sigma} = p \text{BTr}_q^{-1} [(\boldsymbol{\Psi}_{\text{CS}} \otimes \mathbf{I}_q) \boldsymbol{\Omega}^{-1}]$ $q(p-1)\varrho + (p-1)c_2 \text{tr} [(c_1 \mathbf{C} + c_2 \mathbf{D})^{-1} c_1 \mathbf{C}] - c_1 \text{tr} [(c_1 \mathbf{C} + c_2 \mathbf{D})^{-1} c_2 \mathbf{D}] = 0$
MLE	$\mathbf{A} = \text{BTr}_q [(\mathbf{P}_p \otimes \mathbf{I}_q) \boldsymbol{\Omega}]$ $\mathbf{B} = \text{BTr}_q [(\mathbf{Q}_p \otimes \mathbf{I}_q) \boldsymbol{\Omega}]$ $\boldsymbol{\Sigma} = \frac{1}{p} \text{BTr}_q [(\boldsymbol{\Psi}_{\text{CS}}^{-1} \otimes \mathbf{I}_q) \boldsymbol{\Omega}]$ $q(p-1)\varrho + (p-1)c_2 \text{tr} [(c_2 \mathbf{A} + c_1 \mathbf{B})^{-1} c_2 \mathbf{A}] - c_1 \text{tr} [(c_2 \mathbf{A} + c_1 \mathbf{B})^{-1} c_1 \mathbf{B}] = 0$

Można zauważyć, że podobnie jak w Rozdziale 2.1 istnieje relacja między rozwiązaniami powyższych układów równań: macierze  $\boldsymbol{\Omega}$  w (2.18) zostają zastąpione w (2.17) przez odwrotności, podobnie jak  $\boldsymbol{\Psi}_{\text{CS}}$  i  $\boldsymbol{\Sigma}$  w (2.17) zostają zastąpione w (2.18) przez odpowiednie odwrotności.

Warto zauważyć, że ostatnie z omówionych podejść nie tylko pozwala na dostrzeżenie podobieństw między ELE i MLE, ale również na wyznaczenie rozwiązania bez zastosowania algorytmu iteracyjnego.

W pracy Filipiak i in. [10] zaproponowano jeszcze trzecią metodę wyznaczania minimum entropijnej funkcji straty. Zastosowany został rozkład spektralny macierzy  $\boldsymbol{\Psi}_{\text{CS}}$  postaci  $\mathbf{U}\mathbf{G}\mathbf{U}'$ , gdzie  $\mathbf{U}$  jest macierzą ortogonalną stopnia  $p$  z kolumnami odpowiadającymi zortogonalizowanym wektorom własnym macierzy  $\boldsymbol{\Psi}_{\text{CS}}$  i nie zależy od  $\varrho$ , natomiast  $\mathbf{G} = \text{diag}[1 + (p-1)\varrho, 1 - \varrho, \dots, 1 - \varrho]$ . Oznaczając  $\boldsymbol{\Lambda} = (\mathbf{U}' \otimes \mathbf{I}_q) \boldsymbol{\Omega}^{-1} (\mathbf{U} \otimes \mathbf{I}_q)$  otrzymamy macierz blokową  $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda}_{ij})_{1 \leq i, j \leq p}$ . Wówczas entropijna funkcja straty (1.3) może być zapisana w postaci

$$\begin{aligned}
 f_E(\boldsymbol{\Lambda}, \boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}) &= \text{tr} [\boldsymbol{\Lambda} (\mathbf{G} \otimes \boldsymbol{\Sigma})] - \ln |\mathbf{G} \otimes \boldsymbol{\Sigma}| - \ln |\boldsymbol{\Lambda}| - pq \\
 &= \text{tr} \left\{ [1 + (p-1)\varrho] \boldsymbol{\Lambda}_{11} \boldsymbol{\Sigma} + \sum_{i=2}^p (1 - \varrho) \boldsymbol{\Lambda}_{ii} \boldsymbol{\Sigma} \right\} \\
 &\quad - q \ln [1 + (p-1)\varrho] - (p-1)q \ln(1 - \varrho) - p \ln |\boldsymbol{\Sigma}| - \ln |\boldsymbol{\Lambda}| - pq \\
 &= \text{tr} \{ [p\varrho \boldsymbol{\Lambda}_{11} + (1 - \varrho) \text{BTr}_q \boldsymbol{\Lambda}] \boldsymbol{\Sigma} \} \\
 &\quad - q \ln [1 + (p-1)\varrho] - (p-1)q \ln(1 - \varrho) - p \ln |\boldsymbol{\Sigma}| - \ln |\boldsymbol{\Lambda}| - pq.
 \end{aligned}$$

Różniczkując powyższą funkcję ze względu na  $\varrho$  i  $\Sigma$  (por. Magnus i Neudecker [24], Fackler [6], Wprowadzenie) otrzymujemy

$$\begin{cases} \frac{\partial f_E}{\partial \varrho} &= \text{tr} [(p\mathbf{\Lambda}_{11} - \text{BTr}_q \mathbf{\Lambda}) \Sigma] - \frac{q(p-1)}{1+(p-1)\varrho} + \frac{q(p-1)}{1-\varrho} \\ \frac{\partial f_E}{\partial \Sigma} &= \text{vec}' [p\varrho\mathbf{\Lambda}_{11} + (1-\varrho) \text{BTr}_q \mathbf{\Lambda}] - p \cdot \text{vec}' \Sigma^{-1}. \end{cases}$$

Oznaczając  $\text{tr}[(p\mathbf{\Lambda}_{11} - \text{BTr}_q \mathbf{\Lambda})\Sigma]$  przez  $\beta$  i przyrównując powyższe pochodne do zera możemy zapisać

$$\begin{cases} -(p-1)\beta \cdot \varrho^2 + [(p-2)\beta - pq(p-1)] \cdot \varrho + \beta = 0 \\ \Sigma^{-1} = \frac{1}{p} [p\varrho\mathbf{\Lambda}_{11} + (1-\varrho) \text{BTr}_q \mathbf{\Lambda}]. \end{cases}$$

Podobnie jak w poprzednich metodach należy pokazać, że  $\varrho$  minimalizujące funkcję (1.3) należy do przedziału  $(-\frac{1}{p-1}; 1)$ . Niech  $u(\varrho) = -(p-1)\beta \cdot \varrho^2 + [(p-2)\beta - pq(p-1)] \cdot \varrho + \beta$ . Ponieważ  $u(-\frac{1}{p-1}) = pq > 0$  i  $u(1) = -pq(p-1) < 0$  więc  $\check{\varrho} \in (-1/(p-1), 1)$  wyznacza minimum. Ponadto, funkcja ta ma dwa pierwiastki, ponieważ dla  $p \geq 2$  jej wyróżnik jest zawsze dodatni. Jeśli  $\beta < 0$  wtedy wielomian  $u(\varrho)$  jest wypukły i rozwiązanie  $\check{\varrho}$  jest wyznaczone przez większy pierwiastek  $u(\varrho)$ , podczas gdy dla  $\beta > 0$ , przez mniejszy pierwiastek  $u(\varrho)$ . W obu przypadkach  $\check{\varrho}$  ma taką samą postać i najlepsze przybliżenie macierzy  $\mathbf{\Omega}$  może być wyznaczone poprzez rozwiązanie następującego układu równań

$$\begin{cases} \varrho = \frac{-(p-2)\beta + pq(p-1) + \sqrt{[(p-2)\beta + pq(p-1)]^2 + 4(p-1)\beta^2}}{-2(p-1)\beta} \\ \Sigma = p [p\varrho\mathbf{\Lambda}_{11} + (1-\varrho) \text{BTr}_q \mathbf{\Lambda}]^{-1}. \end{cases} \quad (2.19)$$

Podobnie jak w przypadku pierwszej metody,  $\Sigma$  występuje w równaniu na  $\varrho$ , a  $\varrho$  występuje w równaniu na  $\Sigma$ . Układ (2.19) może być zatem rozwiązany iteracyjnie stosując algorytm flip-flop.

Pojawia się teraz pytanie, która z metod minimalizacji entropijnej funkcji straty jest najmniej czasochłonna. W celu porównania czasu potrzebnego do wyznaczenia minimum entropijnej funkcji straty dla  $\varrho$  i  $\Sigma$  przy użyciu standardowej iteracyjnej procedury (2.7), procedury bezpośredniej rozwiązywania układu (2.17) oraz procedury spektralnej (2.19), przeprowadzono badania symulacyjne. Wygenerowano 1000 prób o rozmiarze  $n = 100$  z wielowymiarowego rozkładu normalnego  $N_{pq}(\mathbf{0}, \Psi_{CS} \otimes \Sigma)$

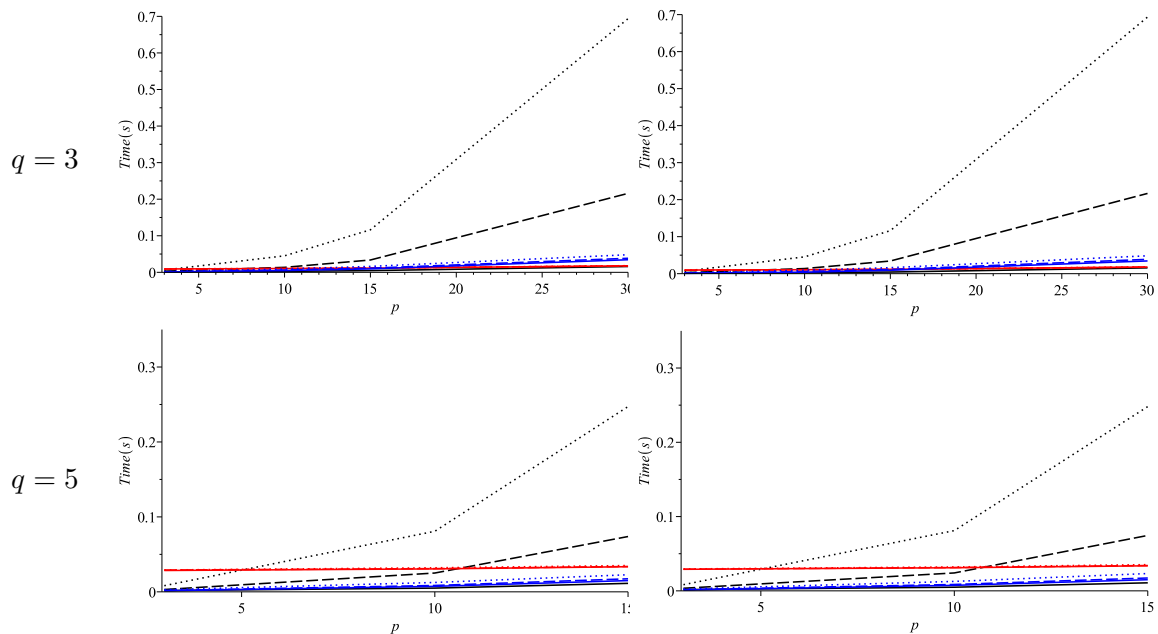
dla  $p = 3, 10, 15, 30$  i  $q = 3$  oraz  $p = 3, 10, 15$  i  $q = 5$ , przyjmując  $\varrho = 0, 0.5, 0.9$  i  $\Sigma = \mathbf{I}_q$  oraz wygenerowane losowo

$$\Sigma = \begin{pmatrix} 6.020 & 4.124 & 1.584 \\ 4.124 & 4.503 & 1.039 \\ 1.584 & 1.039 & 6.326 \end{pmatrix} \quad \text{i} \quad \Sigma = \begin{pmatrix} 6.467 & 0.438 & 1.877 & -1.049 & 2.235 \\ 0.438 & 6.188 & 0.039 & 0.854 & 0.055 \\ 1.877 & 0.039 & 5.971 & 0.245 & 0.000 \\ -1.049 & 0.854 & 0.245 & 4.129 & -0.651 \\ 2.235 & 0.055 & 0.000 & -0.651 & 7.885 \end{pmatrix}.$$

Szczegółowy opis symulacji znajduje się w Rozdziale 4.

Średni czas (w sekundach) wyznaczania minimum entropijnej funkcji straty został obliczony na średniej klasy komputerze, a wyniki przedstawia Rysunek 2.1.

Rysunek 2.1. Średni czas obliczania minimum entropijnej funkcji straty przy użyciu trzech metod (kolor czarny dla procedury iteracyjnej, kolor czerwony dla podejścia bezpośredniego i niebieski dla spektralnego) ze względu na  $p$ ,  $\varrho$  i  $\Sigma$ :  $\varrho = 0$  – linia ciągła,  $\varrho = 0.5$  – linia przerywana,  $\varrho = 0.9$  – linia kropkowana;  $\Sigma = \mathbf{I}_q$  – lewa kolumna,  $\Sigma$  generowana losowo – prawa kolumna.



Zgodnie z oczekiwaniami, na Rysunku 2.1 można zaobserwować, że czas wyznaczania minimum zależy od parametrów  $p$  i  $q$  dla wszystkich metod oraz, zgodnie z intuicją, dla  $q = 5$  jest dłuższy niż dla  $q = 3$ . Ponadto, w przypadku podejścia bezpośredniego czas obliczeń nie zależy od  $\varrho$  (czerwone linie nakładają się na siebie), natomiast dla procedury spektralnej wydłuża się nieznacznie wraz ze wzrostem  $\varrho$ . We wszystkich przypadkach czas nie zależy od  $\Sigma$ . Interesujące jest to, że dla  $\varrho = 0$

procedura iteracyjna może być szybsza od pozostałych. Ponadto, procedura spektralna może trwać krócej niż procedura bezpośrednia gdy  $q = 3$  i  $p$  jest stosunkowo małe. Podobny fenomen zaobserwować można dla  $q = 5$  i wszystkich badanych  $p$ .

Warto zauważyć, że algorytmy flip-flop rozwiązujące (2.7) i (2.19) są zbieżne. Wynika to z faktu, że przy ustalonym  $\varrho$  (odpowiednio  $\Sigma$ ) entropijna funkcja straty jest wypukła, a zatem uzyskiwany w każdym kroku algorytmu punkt stacjonarny  $\Sigma$  (odpowiednio  $\varrho$ ) jest minimum lokalnym. Ponieważ metoda bezpośrednia bazująca na (2.17) daje równoważne rozwiązanie, ono również jest minimum globalnym.

### 2.3. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze procesu autoregresji pierwszego rzędu

Dla dowolnej określonej dodatnio macierzy  $\Omega$  entropijna funkcja straty dla struktury  $\Psi_{AR} \otimes \Sigma \in \mathcal{S}_{AR}$  ma postać

$$f_E(\Omega, \Psi_{AR} \otimes \Sigma) = \text{tr} [\Omega^{-1} (\Psi_{AR} \otimes \Sigma)] - \ln |\Omega^{-1} (\Psi_{AR} \otimes \Sigma)| - pq. \quad (2.20)$$

Najlepsze przybliżenie macierzy  $\Omega$  przez macierz o strukturze  $\Psi_{AR} \otimes \Sigma$  podane jest w następującym twierdzeniu.

**Twierdzenie 2.3.** *Dla zadanej określonej dodatnio macierzy  $\Omega$ , istnieje określona dodatnio macierz  $\Psi_{AR} \otimes \Sigma$ , która minimalizuje entropijną funkcję straty (2.20) na zbiorze  $\mathcal{S}_{AR}$  i minimum to jest osiągnięte dla  $\varrho, \Sigma$  spełniających układ równań*

$$\begin{cases} (1 - \varrho^2) \cdot \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] + 2\varrho q(p - 1) = 0 \\ \Sigma^{-1} = \frac{1}{p} \cdot \text{BTr}_q [(\Psi_{AR} \otimes \mathbf{I}_q) \Omega^{-1}], \end{cases} \quad (2.21)$$

gdzie  $\mathbf{F} = \sum_{i=1}^{p-1} i\varrho^{i-1}(\mathbf{C}^i + \mathbf{C}^{i'})$ .

*Dowód.* Różniczkując (2.20) ze względu na  $\varrho$  (por. Magnus i Neudecker [24], Fackler [6] lub Wprowadzenie) z zastosowaniem reguły łańcuchowej różniczkowania funkcji złożonej postaci

$$\frac{\partial f_E}{\partial \varrho} = \frac{\partial f_E}{\partial (\Psi_{AR} \otimes \Sigma)} \cdot \frac{\partial (\Psi_{AR} \otimes \Sigma)}{\partial \Psi_{AR}} \cdot \frac{\partial \Psi_{AR}}{\partial \varrho}$$

otrzymujemy

$$\begin{aligned} \frac{\partial f_E}{\partial \varrho} &= \text{vec}' (\Omega^{-1})' \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \text{vec } \mathbf{F} \\ &\quad - \text{vec}' (\Psi_{AR}^{-1} \otimes \Sigma^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \cdot \text{vec } \mathbf{F} \\ &= \text{vec}' (\Omega^{-1} - \Psi_{AR}^{-1} \otimes \Sigma^{-1})' (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \cdot \text{vec } \mathbf{F} \\ &= \left[ \text{vec}' \mathbf{F} (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} (\Omega^{-1} - \Psi_{AR}^{-1} \otimes \Sigma^{-1})' \right]', \end{aligned}$$

gdzie

$$\begin{aligned} \text{vec } \mathbf{F} &= \frac{\partial \Psi_{\text{AR}}}{\partial \varrho} = \text{vec}(\mathbf{C} + \mathbf{C}') + 2\varrho \text{vec}(\mathbf{C}^2 + \mathbf{C}'^2) + \dots \\ &\quad + (p-1)\varrho^{p-2} \cdot \text{vec}[\mathbf{C}^{p-1} + \mathbf{C}'^{(p-1)'}] = \text{vec} \sum_{i=1}^{p-1} i\varrho^{i-1} (\mathbf{C}^i + \mathbf{C}'^i). \end{aligned}$$

Wyznaczenie pochodnej ze względu na  $\Sigma$  przebiega analogicznie jak w Rozdziale 2.1, gdzie  $\Psi$  zostało zastąpione przez  $\Psi_{\text{AR}}$ , więc zostanie pominięte.

Stosując własności operatora śladów częściowych (Lemat 1.10 i 1.11) oraz własności śladu macierzy podanych w Lematach 1.2 i 1.3, możemy zapisać następujące przekształcenia pochodnej funkcji  $f_E$  ze względu na  $\varrho$ :

$$\begin{aligned} \frac{\partial f_E}{\partial \varrho} &= \left\{ \text{vec}' \mathbf{F} \cdot \text{vec} \left( \text{PTr}_q \left[ (\mathbf{I}_p \otimes \Sigma') \cdot (\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1})' \right] \right) \right\}' \\ &= \text{tr}' \left\{ \mathbf{F}' \cdot \text{PTr}_q \left[ (\mathbf{I}_p \otimes \Sigma') (\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1})' \right] \right\} \\ &= \text{tr} \left\{ \text{PTr}_q \left[ (\mathbf{F}' \otimes \mathbf{I}_q) (\mathbf{I}_p \otimes \Sigma') (\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1})' \right] \right\} \\ &= \text{tr} \left\{ \text{PTr}_q \left[ (\mathbf{F}' \otimes \Sigma') (\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1})' \right] \right\} \\ &= \text{tr} \left\{ \text{PTr}_q \left\{ [(\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1}) (\mathbf{F} \otimes \Sigma)]' \right\} \right\} \\ &= \text{tr} \left\{ \text{PTr}_q' \left[ (\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1}) (\mathbf{F} \otimes \Sigma) \right] \right\} \\ &= \text{tr} \left\{ \text{PTr}_q \left[ (\mathbf{F} \otimes \Sigma) (\Omega^{-1} - \Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1}) \right] \right\} \\ &= \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] - \text{tr} [(\mathbf{F} \otimes \Sigma) (\Psi_{\text{AR}}^{-1} \otimes \Sigma^{-1})] \\ &= \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] - \text{tr} (\mathbf{F} \cdot \Psi_{\text{AR}}^{-1} \otimes \mathbf{I}_q) \\ &= \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] - \text{tr} (\mathbf{F} \cdot \Psi_{\text{AR}}^{-1}) \cdot \text{tr} (\mathbf{I}_q) \\ &= \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] - \text{tr} (\mathbf{F} \cdot \Psi_{\text{AR}}^{-1}) \cdot q. \end{aligned}$$

Przyrównując powyższą pochodną do zera możemy zapisać

$$\text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] - \text{tr} (\mathbf{F} \cdot \Psi_{\text{AR}}^{-1}) \cdot q = 0.$$

Zapisując odwrotność macierzy o strukturze autoregresji pierwszego rzędu jako

$$\Psi_{\text{AR}}^{-1} = \frac{1}{1-\varrho^2} (\mathbf{I}_p + \varrho^2 \mathbf{C}_1 - \varrho \mathbf{C}_2)$$

można zapisać, że

$$\begin{aligned} \text{tr} (\mathbf{F} \cdot \Psi_{\text{AR}}^{-1}) &= \text{tr} \left\{ \mathbf{F} \cdot \left[ \frac{1}{1-\varrho^2} (\mathbf{I}_p + \varrho^2 \mathbf{C}_1 - \varrho \mathbf{C}_2) \right] \right\} \\ &= \frac{1}{1-\varrho^2} \text{tr} [\mathbf{F} \cdot (\mathbf{I}_p + \varrho^2 \mathbf{C}_1 - \varrho \mathbf{C}_2)] \\ &= \frac{1}{1-\varrho^2} [\text{tr} (\mathbf{F}) + \varrho^2 \text{tr} (\mathbf{F} \mathbf{C}_1) - \varrho \text{tr} (\mathbf{F} \mathbf{C}_2)] \\ &= \frac{1}{1-\varrho^2} [0 + 0 - \varrho \cdot 2(p-1)] \\ &= -\frac{2\varrho(p-1)}{1-\varrho^2}. \end{aligned}$$

Otrzymujemy zatem układ równań (2.21).

Rozwiązanie układu równań (2.21) należy do  $\mathcal{S}_{\text{AR}}$  jeżeli odpowiednie równanie ma pierwiastek należący do przedziału  $(-1, 1)$ . Wprowadźmy zatem oznaczenie

$$W(\varrho) = (1 - \varrho^2) \cdot \text{tr} [(\mathbf{F} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1}] + 2\varrho q(p - 1).$$

Ponieważ  $W(-1) = -2q(p-1) < 0$  oraz  $W(1) = 2q(p-1) > 0$  funkcja  $W(\varrho)$  posiada co najmniej jeden pierwiastek w przedziale  $(-1, 1)$  minimalizujący funkcję (2.20) ze względu na  $\varrho$ . W przypadku więcej niż jednego rozwiązania  $W(\check{\varrho}) = 0$  w przedziale  $(-1, 1)$  wybierane jest  $\check{\varrho}$  minimalizujące  $f_E(\mathbf{\Omega}, \mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})$ . Z własności operatora śladu blokowego (Lemat 1.14) wynika, że również  $\check{\mathbf{\Sigma}}$  jest macierzą określoną dodatnio. ■

Układ równań (2.21) może być rozwiązany iteracyjnie z zastosowaniem algorytmu flip-flop, ponieważ rozwiązanie nie jest podane w postaci jawnej. Warto zauważyć, że pierwsze równanie układu (2.21) reprezentuje wielomian nawet stopnia  $p$ , a zatem obliczenia potrzebne do wyznaczenia numerycznego rozwiązania mogą być czasochłonne.

Na koniec zwróćmy uwagę, że algorytm flip-flop rozwiązujący (2.21) jest zbieżny. Wynika to z faktu, że przy ustalonym  $\varrho$  (odpowiednio  $\mathbf{\Sigma}$ ) entropijna funkcja straty jest wypukła, a zatem uzyskiwany w każdym kroku algorytmu punkt stacjonarny  $\mathbf{\Sigma}$  (odpowiednio  $\varrho$ ) jest minimum lokalnym.



### 3. Aproksymacja kwadratową funkcją straty

W kolejnych podrozdziałach poszukiwać będziemy minimum kwadratowej funkcji straty zdefiniowanej wzorem (1.4) w zbiorach  $\mathcal{S}$ ,  $\mathcal{S}_{CS}$  i  $\mathcal{S}_{AR}$  odpowiednio.

#### 3.1. Rozdzielna struktura kowariancyjna

Dla dowolnej określonej dodatnio macierzy  $\Omega$  kwadratowa funkcja straty dla struktury  $\Psi \otimes \Sigma$  ma postać

$$f_Q(\Omega, \Psi \otimes \Sigma) = \text{tr} \left\{ [\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}]^2 \right\}. \quad (3.1)$$

Najlepsze przybliżenie macierzy  $\Omega$  przez macierz o strukturze  $\Psi \otimes \Sigma$  podane jest w następującym twierdzeniu.

**Twierdzenie 3.1.** *Dla danej symetrycznej, określonej dodatnio macierzy  $\Omega$ , istnieje symetryczna macierz  $\Psi \otimes \Sigma$ , która minimalizuje kwadratową funkcję straty postaci (3.1) w zbiorze  $\mathcal{S}$  oraz minimum to jest osiągnięte dla  $\Psi$  i  $\Sigma$  spełniających układ równań*

$$\begin{cases} \text{vec } \Psi &= [\mathbf{A} (\Omega^{-1} \otimes \Omega^{-1}) \mathbf{A}']^{-1} \mathbf{A} \text{vec } \Omega^{-1} \\ \text{vec } \Sigma &= [\mathbf{B} (\Omega^{-1} \otimes \Omega^{-1}) \mathbf{B}']^{-1} \mathbf{B} \text{vec } \Omega^{-1}, \end{cases} \quad (3.2)$$

gdzie

$$\begin{aligned} \mathbf{A} &= (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \Sigma), \\ \mathbf{B} &= (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Psi \otimes \mathbf{I}_q). \end{aligned} \quad (3.3)$$

*Dowód.* Zróżniczkujemy  $f_Q(\Omega, \Psi \otimes \Sigma)$  względem  $\Psi$  oraz  $\Sigma$ . Zastosujmy oznaczenie  $\mathbf{Z} = \Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}$ . Stosując reguły różniczkowania dla macierzy symetrycznych

$$\begin{aligned} \frac{\partial f_Q}{\partial \Psi} &= \frac{\partial f_Q}{\partial \mathbf{Z}^2} \cdot \frac{\partial \mathbf{Z}^2}{\partial \mathbf{Z}} \cdot \frac{\partial \mathbf{Z}}{\partial (\Psi \otimes \Sigma)} \cdot \frac{\partial (\Psi \otimes \Sigma)}{\partial \Psi} \\ \frac{\partial f_Q}{\partial \Sigma} &= \frac{\partial f_Q}{\partial \mathbf{Z}^2} \cdot \frac{\partial \mathbf{Z}^2}{\partial \mathbf{Z}} \cdot \frac{\partial \mathbf{Z}}{\partial (\Psi \otimes \Sigma)} \cdot \frac{\partial (\Psi \otimes \Sigma)}{\partial \Sigma} \end{aligned}$$

otrzymujemy

$$\begin{aligned}
\frac{\partial f_Q}{\partial \Psi} &= \text{vec}'(\mathbf{I}_{pq}) \cdot \left\{ [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}] \right\} \\
&\quad \cdot (\mathbf{I}_{pq} \otimes \Omega^{-1}) \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \cdot \mathbf{D}_p \\
&= \left\{ (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot ([\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]) \cdot \text{vec}(\mathbf{I}_{pq}) \left. \right\}' \mathbf{D}_p \\
&= \left\{ (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot [\{[\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq}\} \text{vec}(\mathbf{I}_{pq}) \\
&\quad \left. + \{\mathbf{I}_{pq} \otimes [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \text{vec}(\mathbf{I}_{pq})\} \right\}' \mathbf{D}_p,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial f_Q}{\partial \Sigma} &= \text{vec}'(\mathbf{I}_{pq}) \cdot \left\{ [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}] \right\} \\
&\quad \cdot (\mathbf{I}_{pq} \otimes \Omega^{-1}) \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\text{vec } \Psi \otimes \mathbf{I}_{q^2}) \cdot \mathbf{D}_q \\
&= \left\{ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot ([\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]) \cdot \text{vec}(\mathbf{I}_{pq}) \left. \right\}' \mathbf{D}_q \\
&= \left\{ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot [\{[\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq}\} \text{vec}(\mathbf{I}_{pq}) \\
&\quad \left. + \{\mathbf{I}_{pq} \otimes [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \text{vec}(\mathbf{I}_{pq})\} \right\}' \cdot \mathbf{D}_q.
\end{aligned}$$

Stosując Lemat 1.6 powyższe pochodne można zapisać w postaci:

$$\begin{aligned}
\frac{\partial f_Q}{\partial \Psi} &= \left\{ (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot \left[ \text{vec}\{\mathbf{I}_{pq} [\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}]\}' + \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}] \mathbf{I}_{pq}\} \right] \left. \right\}' \cdot \mathbf{D}_p \\
&= \left\{ (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot \left[ \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} + \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \right] \left. \right\}' \cdot \mathbf{D}_p \\
&= \left\{ 2 (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \left. \right\}' \cdot \mathbf{D}_p \\
&= \left\{ 2 (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \cdot \text{vec}\{\Omega^{-1} [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \right\}' \cdot \mathbf{D}_p,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial f_Q}{\partial \Sigma} &= \left\{ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot \left[ \text{vec}\{\mathbf{I}_{pq} [\Omega^{-1}(\Psi \otimes \Sigma) - \mathbf{I}_{pq}]\}' + \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}] \mathbf{I}_{pq}\} \right] \left. \right\}' \cdot \mathbf{D}_q \\
&= \left\{ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\
&\quad \cdot \left[ \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} + \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \right] \left. \right\}' \cdot \mathbf{D}_q \\
&= \left\{ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \cdot (\mathbf{I}_{pq} \otimes \Omega^{-1}) \cdot 2 \right. \\
&\quad \cdot \text{vec}\{[(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \left. \right\}' \cdot \mathbf{D}_q \\
&= \left\{ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \cdot 2 \text{vec}\{\Omega^{-1} [(\Psi \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}]\} \right\}' \cdot \mathbf{D}_q.
\end{aligned}$$

Przyrównując powyższe pochodne do zera możemy zapisać

$$\begin{cases} (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \cdot \text{vec} \{ \Omega^{-1} [(\Psi \otimes \Sigma) \Omega^{-1} - \mathbf{I}_{pq}] \} = \mathbf{0} \\ (\text{vec}' \Psi \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \cdot \text{vec} \{ \Omega^{-1} [(\Psi \otimes \Sigma) \Omega^{-1} - \mathbf{I}_{pq}] \} = \mathbf{0}. \end{cases}$$

Ponieważ macierze duplikacji powodują jedynie powielenie wybranych składowych wektorów pochodnych, nie muszą one być uwzględnione przy przyrównywaniu do zera.

Używając operatorów śladów częściowych oraz śladu blokowego zaprezentowanych w Definicji 1.1 i ich własności podanej w Lemacie 1.10 oraz formuł podanych w Lematach 1.6 i 1.7 można zapisać następujące przekształcenia powyższego układu równań:

$$\begin{cases} \text{vec} (\text{PTr}_q \{ (\mathbf{I}_p \otimes \Sigma) \Omega^{-1} [(\Psi \otimes \Sigma) \Omega^{-1} - \mathbf{I}_{pq}] \}) = \mathbf{0} \\ \text{vec} (\text{BTr}_q \{ (\Psi \otimes \mathbf{I}_q) \Omega^{-1} [(\Psi \otimes \Sigma) \Omega^{-1} - \mathbf{I}_{pq}] \}) = \mathbf{0} \\ \text{vec} (\text{PTr}_q \{ (\mathbf{I}_p \otimes \mathbf{I}_q) (\mathbf{I}_p \otimes \Sigma) \Omega^{-1} [(\Psi \otimes \Sigma) \Omega^{-1} - \mathbf{I}_{pq}] \}) = \mathbf{0} \\ \text{vec} (\text{BTr}_q \{ (\mathbf{I}_p \otimes \mathbf{I}_q) (\Psi \otimes \mathbf{I}_q) \Omega^{-1} [(\Psi \otimes \Sigma) \Omega^{-1} - \mathbf{I}_{pq}] \}) = \mathbf{0} \\ (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1} (\Psi \otimes \Sigma) \Omega^{-1}] = \\ = (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] \\ (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} [(\Psi \otimes \mathbf{I}_q) \Omega^{-1} (\Psi \otimes \Sigma) \Omega^{-1}] = \\ = (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} [(\Psi \otimes \mathbf{I}_q) \Omega^{-1}] \\ (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} [(\mathbf{I}_p \otimes \Sigma) \Omega^{-1} (\mathbf{I}_p \otimes \Sigma) (\Psi \otimes \mathbf{I}_q) \Omega^{-1}] = \\ = (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \Sigma) \text{vec} \Omega^{-1} \\ (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) \text{vec} [(\Psi \otimes \mathbf{I}_q) \Omega^{-1} (\Psi \otimes \mathbf{I}_q) (\mathbf{I}_p \otimes \Sigma) \Omega^{-1}] = \\ = (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Psi \otimes \mathbf{I}_q) \text{vec} \Omega^{-1} \\ (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) [\Omega^{-1} \otimes (\mathbf{I}_p \otimes \Sigma) \Omega^{-1} (\mathbf{I}_p \otimes \Sigma)] \text{vec} (\Psi \otimes \mathbf{I}_q) = \\ = (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \Sigma) \text{vec} \Omega^{-1} \\ (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) [\Omega^{-1} \otimes (\Psi \otimes \mathbf{I}_q) \Omega^{-1} (\Psi \otimes \mathbf{I}_q)] \text{vec} (\mathbf{I}_p \otimes \Sigma) = \\ = (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Psi \otimes \mathbf{I}_q) \text{vec} \Omega^{-1} \\ (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) [\Omega^{-1} \otimes (\mathbf{I}_p \otimes \Sigma) \Omega^{-1} (\mathbf{I}_p \otimes \Sigma)] (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) \\ \cdot (\text{vec} \Psi \otimes \text{vec} \mathbf{I}_q) = (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \Sigma) \text{vec} \Omega^{-1} \\ (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) [\Omega^{-1} \otimes (\Psi \otimes \mathbf{I}_q) \Omega^{-1} (\Psi \otimes \mathbf{I}_q)] (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) \\ \cdot (\text{vec} \mathbf{I}_p \otimes \text{vec} \Sigma) = (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Psi \otimes \mathbf{I}_q) \text{vec} \Omega^{-1} \end{cases}$$

$$\left\{ \begin{array}{l} (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \boldsymbol{\Sigma}) (\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{\Omega}^{-1}) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \boldsymbol{\Sigma}) \\ \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec} \mathbf{I}_q) \text{vec} \boldsymbol{\Psi} \\ = (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{I}_p \otimes \boldsymbol{\Sigma}) \text{vec} \boldsymbol{\Omega}^{-1} \\ (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_q) (\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{\Omega}^{-1}) (\mathbf{I}_{pq} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_q) \\ \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\text{vec} \mathbf{I}_p \otimes \mathbf{I}_{q^2}) \text{vec} \boldsymbol{\Sigma} \\ = (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_q) \text{vec} \boldsymbol{\Omega}^{-1}. \end{array} \right.$$

Stosując oznaczenia (3.3) otrzymujemy układ równań (3.2). ■

Warto zwrócić uwagę, że układ równań (3.2) nie jest podany w postaci jawnej, ale może być rozwiązany iteracyjnie za pomocą algorytmu flip-flop: w pierwszym kroku ustalamy jeden z czynników iloczynu Kroneckera, na przykład  $\boldsymbol{\Psi} = \mathbf{I}_p$ , i wstawiając do drugiego równania układu (3.2) otrzymujemy pierwsze przybliżenie  $\boldsymbol{\Sigma}$ , które następnie wstawiamy do pierwszego równania w celu otrzymania pierwszego przybliżenia macierzy  $\boldsymbol{\Psi}$ . Algorytm jest kontynuowany tak długo, aż różnice między kolejnymi przybliżeniami (w sensie pewnej funkcji) będą wystarczająco małe. Ze względu na złożoność obliczeniową równań występujących w układzie (3.2) jego rozwiązanie jest bardzo czasochłonne.

Z pracy Haff [17] wiadomo, że kwadratowa funkcja straty jest wypukła ze względu na macierz  $\boldsymbol{\Gamma}$  należącą do wypukłego zbioru macierzy określonych dodatnio. Zatem w zbiorze tym istnieje minimum lokalne, które jest jednocześnie minimum globalnym. Załóżmy teraz, że  $\boldsymbol{\Gamma} \in \mathcal{S}$ . Jeżeli ustalimy jeden z czynników iloczynu Kroneckera, to mówimy o *profilowej* kwadratowej funkcji straty, która podobnie jak poprzednio jest funkcją wypukłą ze względu na drugi czynnik. Zatem dla ustalonej macierzy  $\boldsymbol{\Psi}$  (odpowiednio  $\boldsymbol{\Sigma}$ ) profilowa kwadratowa funkcja straty jest wypukła ze względu na  $\boldsymbol{\Sigma}$  (odpowiednio  $\boldsymbol{\Psi}$ ), a więc istnieją lokalne minima funkcji profilowych. Ze względu na konstrukcję algorytmu flip-flop, minima te prowadzą do minimum lokalnego z  $L_{\boldsymbol{\Omega}}^{(Q)}(\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$ .

Rozwiązanie  $\tilde{\boldsymbol{\Psi}}$  i  $\tilde{\boldsymbol{\Sigma}}$  spełniające układ równań (3.2) może być traktowane jako estymator macierzy kowariancji o strukturze rozdzielnej. Estymator ten zapisywać będziemy w skrócie QLE (ang. *quadratic loss estimator*).

### 3.2. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze kompletnie symetrycznej

Dla dowolnej określonej dodatnio macierzy  $\Omega$  kwadratowa funkcja straty dla struktury  $\Psi_{CS} \otimes \Sigma \in \mathcal{S}_{CS}$  ma postać

$$f_Q(\Omega, \Psi_{CS} \otimes \Sigma) = \text{tr} \left\{ \left[ \Omega^{-1}(\Psi_{CS} \otimes \Sigma) - \mathbf{I}_{pq} \right]^2 \right\}. \quad (3.4)$$

Najlepsze przybliżenie macierzy  $\Omega$  przez macierz o strukturze  $\Psi_{CS} \otimes \Sigma$  podane jest w następującym twierdzeniu.

**Twierdzenie 3.2.** *Dla danej określonej dodatnio macierzy  $\Omega$  istnieje symetryczna macierz  $\Psi_{CS} \otimes \Sigma$ , która minimalizuje kwadratową funkcję straty (3.4) w zbiorze  $\mathcal{S}_{CS}$  i minimum to jest osiągnięte dla  $\varrho, \Sigma$  spełniających układ równań*

$$\begin{cases} \varrho &= \frac{\text{tr}(\mathbf{M}) - \text{tr}[\mathbf{M}(\mathbf{I}_p \otimes \Sigma)\Omega^{-1}]}{\text{tr}(\mathbf{M}^2)} \\ \text{vec } \Sigma &= [\mathbf{B}(\Omega^{-1} \otimes \Omega^{-1})\mathbf{B}']^{-1} \mathbf{B} \text{vec } \Omega^{-1}, \end{cases} \quad (3.5)$$

gdzie

$$\begin{aligned} \mathbf{M} &= [(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \otimes \Sigma] \Omega^{-1}, \\ \mathbf{B} &= (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Psi_{CS} \otimes \mathbf{I}_q). \end{aligned} \quad (3.6)$$

*Dowód.* Różniczkując (3.4) ze względu na  $\varrho$  (por. Magnus i Neudecker [24], Fackler [6] lub Wprowadzenie) z zastosowaniem reguły łańcuchowej różniczkowania funkcji złożonej postaci

$$\frac{\partial f_Q}{\partial \varrho} = \frac{\partial f_Q}{\partial \mathbf{Z}^2} \cdot \frac{\partial \mathbf{Z}^2}{\partial \mathbf{Z}} \cdot \frac{\partial \mathbf{Z}}{\partial (\Psi_{CS} \otimes \Sigma)} \cdot \frac{\partial (\Psi_{CS} \otimes \Sigma)}{\partial \Psi_{CS}} \cdot \frac{\partial \Psi_{CS}}{\partial \varrho},$$

gdzie  $\mathbf{Z} = \Omega^{-1}(\Psi_{CS} \otimes \Sigma) - \mathbf{I}_{pq}$ , otrzymujemy

$$\begin{aligned} \frac{\partial f_Q}{\partial \varrho} &= \text{vec}'(\mathbf{I}_{pq}) \cdot \left\{ [(\Psi_{CS} \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [\Omega^{-1}(\Psi_{CS} \otimes \Sigma) - \mathbf{I}_{pq}] \right\} \\ &\quad \cdot (\mathbf{I}_{pq} \otimes \Omega^{-1}) \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \Sigma) \cdot \text{vec}(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \\ &= \left\{ \text{vec}'(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) (\mathbf{I}_{p^2} \otimes \text{vec}' \Sigma) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Omega^{-1}) \right. \\ &\quad \left. \cdot [(\Psi_{CS} \otimes \Sigma)\Omega^{-1} - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [\Omega^{-1}(\Psi_{CS} \otimes \Sigma) - \mathbf{I}_{pq}] \right\}' \text{vec}(\mathbf{I}_{pq}) \}. \end{aligned}$$

Wyznaczenie pochodnej ze względu na  $\Sigma$  przebiega analogicznie jak w Rozdziale 3.1, gdzie  $\Psi$  zostało zastąpione przez  $\Psi_{CS}$ , więc zostanie pominięte.

Stosując własności operatora śladów częściowych (Lemat 1.10, 1.11, 1.13) oraz własności śladu macierzy podanych w Lematach 1.2 i 1.3 możemy zapisać następujące przekształcenia pochodnej funkcji  $f_Q$  ze względu na  $\varrho$ :

$$\begin{aligned}
\frac{\partial f_Q}{\partial \varrho} &= \left\{ \text{vec}'(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \cdot \text{vec} \left( 2 \text{PTr}_q [(\mathbf{I}_p \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} \{(\boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} - \mathbf{I}_{pq}\}] \right) \right\}' \\
&= \text{tr} \left\{ (\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \cdot 2 \text{PTr}_q [(\mathbf{I}_p \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} \{(\boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} - \mathbf{I}_{pq}\}] \right\} \\
&= 2 \text{tr} \left\{ \text{PTr}_q \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \mathbf{I}_q] (\mathbf{I}_p \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} [(\boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} - \mathbf{I}_{pq}] \right\} \right\} \\
&= 2 \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \boldsymbol{\Sigma}] \boldsymbol{\Omega}^{-1} [(\boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} - \mathbf{I}_{pq}] \right\}.
\end{aligned}$$

Przyrównując powyższą pochodną do zera możemy zapisać

$$\text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \boldsymbol{\Sigma}] \boldsymbol{\Omega}^{-1} (\boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1} \right\} = \text{tr} \left\{ [(\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p) \otimes \boldsymbol{\Sigma}] \boldsymbol{\Omega}^{-1} \right\}.$$

Zapisując macierz o strukturze kompletnej symetrii jako

$$\boldsymbol{\Psi}_{\text{CS}} = \mathbf{I}_p + \varrho (\mathbf{1}_p \mathbf{1}'_p - \mathbf{I}_p)$$

oraz podstawiając  $\mathbf{M}$  zdefiniowane w (3.6) otrzymujemy

$$\text{tr} [\mathbf{M} (\mathbf{I}_p \otimes \boldsymbol{\Sigma}) \boldsymbol{\Omega}^{-1}] + \varrho \text{tr} (\mathbf{M}^2) = \text{tr} (\mathbf{M}).$$

Podsumowując, równania pozwalające wyznaczyć minimum kwadratowej funkcji straty w zbiorze  $\mathcal{S}_{\text{CS}}$  można zapisać w postaci (3.5) z uwzględnieniem (3.6). ■

Podobnie jak w Rozdziale 3.1 pierwsze równanie układu (3.5) zależy od  $\boldsymbol{\Sigma}$ , a drugie równanie zależy od  $\varrho$ . W związku z tym układ ten można rozwiązać iteracyjnie za pomocą algorytmu flip-flop. W pierwszym kroku ustalamy np.  $\boldsymbol{\Sigma} = \mathbf{I}_p$  i wstawiając do pierwszego równania układu (3.2) otrzymujemy pierwsze przybliżenie  $\varrho$ , które następnie wstawiamy do drugiego równania w celu otrzymania pierwszego przybliżenia macierzy  $\boldsymbol{\Sigma}$ . Algorytm jest kontynuowany tak długo, aż różnice między kolejnymi przybliżeniami (w sensie pewnej funkcji) będą wystarczająco małe.

Warto zauważyć, że algorytm flip-flop rozwiązujący (3.5) jest zbieżny. Wynika to z faktu, że przy ustalonym  $\varrho$  (odpowiednio  $\boldsymbol{\Sigma}$ ) kwadratowa funkcja straty jest wypukła, a zatem uzyskiwany w każdym kroku algorytmu punkt stacjonarny  $\boldsymbol{\Sigma}$  (odpowiednio  $\varrho$ ) jest minimum lokalnym.

Jak pokazały badania symulacyjne rozwiązanie układu równań (3.5) niekoniecznie jest określone dodatnio. Warto zwrócić uwagę, że podobny problem pojawił się

w przypadku standardowego modelu wielowymiarowego z macierzą kowariancji o liniowej strukturze Toeplitza, a rozwiązanie zaproponowane w pracy Filipiak i in. [15] jako rzut na stożek macierzy określonych nieujemnie jest w istocie rozwiązaniem numerycznym. Ponieważ struktura rozdzielna nie jest strukturą liniową, wyznaczenie warunkowego minimum jest jeszcze trudniejsze i wymaga dalszych badań.

### 3.3. Rozdzielna struktura kowariancyjna z czynnikiem o strukturze procesu autoregresji pierwszego rzędu

Dla dowolnej określonej dodatnio macierzy  $\Omega$  kwadratowa funkcja straty dla struktury  $\Psi_{AR} \otimes \Sigma \in \mathcal{S}_{AR}$  ma postać

$$f_Q(\Omega, \Psi_{AR} \otimes \Sigma) = \text{tr} \left\{ [\Omega^{-1}(\Psi_{AR} \otimes \Sigma) - \mathbf{I}_{pq}]^2 \right\}. \quad (3.7)$$

Najlepsze przybliżenie macierzy  $\Omega$  przez macierz o strukturze  $\Psi_{AR} \otimes \Sigma$  podane jest w następującym twierdzeniu.

**Twierdzenie 3.3.** *Dla danej określonej dodatnio macierzy  $\Omega$  istnieje symetryczna macierz  $\Psi_{AR} \otimes \Sigma$ , która minimalizuje kwadratową funkcję straty (3.7) w zbiorze  $\mathcal{S}_{AR}$  i minimum to jest osiągnane dla  $\varrho, \Sigma$  spełniających układ równań*

$$\begin{cases} \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1} (\Psi_{AR} \otimes \Sigma) \Omega^{-1}] = \text{tr} [(\mathbf{F} \otimes \Sigma) \Omega^{-1}] \\ \text{vec} \Sigma = [\mathbf{B} (\Omega^{-1} \otimes \Omega^{-1}) \mathbf{B}']^{-1} \mathbf{B} \text{vec} \Omega^{-1}, \end{cases} \quad (3.8)$$

gdzie macierz  $\mathbf{B}$  zdefiniowana jest jako

$$\mathbf{B} = (\text{vec}' \mathbf{I}_p \otimes \mathbf{I}_{q^2}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \Psi_{AR} \otimes \mathbf{I}_q) \quad (3.9)$$

oraz  $\mathbf{F} = \sum_{i=1}^{p-1} i \varrho^{i-1} (\mathbf{C}^i + \mathbf{C}^{i'})$ .

*Dowód.* Różniczkując (3.7) ze względu na  $\varrho$  (por. Magnus i Neudecker [24], Fackler [6] lub Wprowadzenie) z zastosowaniem reguły łańcuchowej różniczkowania funkcji złożonej postaci

$$\frac{\partial f_Q}{\partial \varrho} = \frac{\partial f_Q}{\partial \mathbf{Z}^2} \cdot \frac{\partial \mathbf{Z}^2}{\partial \mathbf{Z}} \cdot \frac{\partial \mathbf{Z}}{\partial (\Psi_{AR} \otimes \Sigma)} \cdot \frac{\partial (\Psi_{AR} \otimes \Sigma)}{\partial \Psi_{AR}} \cdot \frac{\partial \Psi_{AR}}{\partial \varrho},$$

gdzie  $\mathbf{Z} = \mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}) - \mathbf{I}_{pq}$ , otrzymujemy

$$\begin{aligned} \frac{\partial f_Q}{\partial \varrho} &= \text{vec}'(\mathbf{I}_{pq}) \cdot \left\{ [(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1} - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [(\mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}) - \mathbf{I}_{pq})] \right\} \\ &\quad \cdot (\mathbf{I}_{pq} \otimes \mathbf{\Omega}^{-1}) \cdot (\mathbf{I}_p \otimes \mathbf{K}_{q,p} \otimes \mathbf{I}_q) (\mathbf{I}_{p^2} \otimes \text{vec } \mathbf{\Sigma}) \cdot \text{vec } \mathbf{F} \\ &= \left\{ \text{vec}' \mathbf{F} (\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{\Sigma}) (\mathbf{I}_p \otimes \mathbf{K}_{p,q} \otimes \mathbf{I}_q) (\mathbf{I}_{pq} \otimes \mathbf{\Omega}^{-1}) \right. \\ &\quad \left. \cdot \left( [(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1} - \mathbf{I}_{pq}] \otimes \mathbf{I}_{pq} + \mathbf{I}_{pq} \otimes [(\mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}) - \mathbf{I}_{pq})] \right)' \cdot \text{vec}(\mathbf{I}_{pq}) \right\}'. \end{aligned}$$

Wyznaczenie pochodnej ze względu na  $\mathbf{\Sigma}$  przebiega analogicznie jak w Rozdziale 3.1, gdzie  $\mathbf{\Psi}$  zostało zastąpione przez  $\mathbf{\Psi}_{\text{AR}}$ , więc zostanie pominięte.

Stosując własności operatora śladów częściowych (Lematy 1.10 – 1.14) oraz własności podane w Lematach 1.2 i 1.6, możemy zapisać następujące przekształcenia pochodnej funkcji  $f_Q$  ze względu na  $\varrho$ :

$$\begin{aligned} \frac{\partial f_Q}{\partial \varrho} &= \left\{ \text{vec}' \mathbf{F} \cdot \text{vec} \left( 2 \text{PTr}_q \left\{ (\mathbf{I}_p \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1} [(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1} - \mathbf{I}_{pq}] \right\} \right) \right\}' \\ &= \text{tr} \left\{ \mathbf{F} \cdot 2 \text{PTr}_q \left\{ (\mathbf{I}_p \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1} [(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1} - \mathbf{I}_{pq}] \right\} \right\} \\ &= 2 \text{tr} \left\{ \text{PTr}_q \left\{ (\mathbf{F} \otimes \mathbf{I}_q) (\mathbf{I}_p \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1} [(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1} - \mathbf{I}_{pq}] \right\} \right\} \\ &= 2 \text{tr} \left\{ (\mathbf{F} \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1} [(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1} - \mathbf{I}_{pq}] \right\}. \end{aligned}$$

Przyrównując powyższą pochodną do zera możemy zapisać

$$\text{tr} [(\mathbf{F} \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1} (\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1}] - \text{tr} [(\mathbf{F} \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1}] = 0.$$

Podsumowując, równania pozwalające wyznaczyć minimum kwadratowej funkcji straty w zbiorze  $\mathcal{S}_{\text{AR}}$  można zapisać w postaci (3.8). ■

Ponownie pierwsze równanie układu (3.8) zależy od  $\mathbf{\Sigma}$ , a drugie równanie zależy od  $\varrho$ , dlatego układ ten można rozwiązać iteracyjnie za pomocą algorytmu flip-flop. Ze względu na złożoność obliczeniową każdego z równań, rozwiązanie układu równań (3.8) jest bardzo czasochłonne.

Warto zauważyć, że algorytm flip-flop rozwiązujący (3.8) jest zbieżny. Wynika to z faktu, że przy ustalonym  $\varrho$  (odpowiednio  $\mathbf{\Sigma}$ ) kwadratowa funkcja straty jest wypukła, a zatem uzyskiwany w każdym kroku algorytmu punkt stacjonarny  $\mathbf{\Sigma}$  (odpowiednio  $\varrho$ ) jest minimum lokalnym.

Podobnie jak w Rozdziale 3.2, badania symulacyjne pokazały, że rozwiązanie układu równań (3.8) niekoniecznie jest określone dodatnio. W związku z tym problem określoności rozwiązania wymaga dalszych badań.

# 4. Badania symulacyjne oraz zastosowanie minimum funkcji rozbieżności do danych rzeczywistych

W tym rozdziale zastosujemy wyniki algebraiczne przedstawione w Rozdziałach 2 i 3 do zagadnień statystycznych.

## 4.1. Wprowadzenie do badań symulacyjnych

W dwóch pierwszych podrozdziałach pokażemy, że zarówno entropijna jak i kwadratowa funkcja straty prawidłowo rozpoznają prawdziwą strukturę kowariancyjną tensora obserwacji oraz zbadamy własności estymatorów otrzymanych poprzez minimalizację rozważanej funkcji rozbieżności. W tym celu wygenerowane zostały dane z rozkładu normalnego z wartością oczekiwaną  $\mathbf{0}$  i macierzą kowariancji  $\Psi_{CS} \otimes \Sigma$  oraz  $\Psi_{AR} \otimes \Sigma$ . W doborze parametrów  $\varrho$  i  $\Sigma$  pomocny będzie następujący lemat.

**Lemat 4.1.** *Niech  $\Omega$  będzie macierzą kowariancji w modelu (1.1). Indeksy rozbieżności  $\zeta_{UN}^{(k)}$ ,  $\zeta_{CS}^{(k)}$  i  $\zeta_{AR}^{(k)}$ ,  $k \in \{E, Q\}$ , mają następujące własności:*

- $\zeta_{UN}^{(k)}$  nie zależy od wyboru  $\Psi$  i  $\Sigma$  gdy  $\Omega = \Psi \otimes \Sigma$  (w szczególności  $\Psi$  może mieć postać CS lub AR(1)),
- $\zeta_{CS}^{(k)}$  nie zależy od wyboru  $\varrho$  i  $\Sigma$  gdy  $\Omega = \Psi_{CS} \otimes \Sigma$  oraz zależy od wyboru  $\varrho$  gdy  $\Omega = \Psi_{AR} \otimes \Sigma$ ,
- $\zeta_{AR}^{(k)}$  zależy od wyboru  $\varrho$  i nie zależy od wyboru  $\Sigma$  gdy  $\Omega = \Psi_{CS} \otimes \Sigma$  lub  $\Omega = \Psi_{AR} \otimes \Sigma$ .

*Dowód.* Wiadomo, że transformacja liniowa macierzy obserwacji,  $\mathbf{Z} = \mathbf{A}\mathbf{Y}$ , ma rozkład normalny z macierzą kowariancji  $\mathbf{A}(\Psi \otimes \Sigma)\mathbf{A}'$ . Jeżeli  $\mathbf{A} \neq \mathbf{I}_{pq}$ , to aby przy takim przekształceniu zachować strukturę macierzy kowariancji (ze zbioru  $\mathcal{S}, \mathcal{S}_{CS}, \mathcal{S}_{AR}$ ) macierz  $\mathbf{A}$  musi być nieosobliwa oraz posiadać strukturę iloczynu Kroneckera,  $\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{A}_2$ , przy czym  $\mathbf{A}_1, \mathbf{A}_2$  są nieosobliwe oraz

(Z1) jeśli  $\Psi = \Psi_{CS}$  to  $\mathbf{A}_1$  jest taka, że  $\mathbf{A}_1\Psi_{CS}\mathbf{A}_1'$  jest macierzą o strukturze CS (w szczególności  $\mathbf{A}_1 = \mathbf{A}_{CS}$ ),

(Z2) jeśli  $\Psi = \Psi_{AR}$  to  $\mathbf{A}_1$  jest taka, że  $\mathbf{A}_1\Psi_{AR}\mathbf{A}_1'$  jest macierzą o strukturze AR(1) (w szczególności  $\mathbf{A}_1 = \mathbf{V}(\Psi_{AR})^{-1/2}$  gdzie  $\mathbf{V}\mathbf{V}' = \mathbf{A}_{AR}$  oraz  $(\Psi_{AR})^{-1/2}(\Psi_{AR})^{-1/2} = (\Psi_{AR})^{-1}$ ).

Niech zatem  $\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{A}_2$  spełnia powyższe warunki. Przypomnijmy, że macierze  $\mathbf{A}(\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})\mathbf{A}'$  oraz  $\mathbf{A}(\boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma})\mathbf{A}'$  generują całe zbiory  $\mathcal{S}$  i  $\mathcal{S}_{\text{CS}}$  odpowiednio; por. Wprowadzenie, Rozdział 1.2. Oznacza to, że spełnione są założenia Lematu 1.1. We wszystkich przypadkach MLE macierzy  $\mathbf{A}(\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})\mathbf{A}'$  jest równy  $\mathbf{A}\mathbf{S}\mathbf{A}'$ , gdzie  $\mathbf{S}$  jest MLE macierzy  $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$  i nie należy do żadnego z rozważanych zbiorów. Zatem warunek (i) Lematu 1.1 nigdy nie jest spełniony. Ponieważ zbiory  $\mathcal{S}$  i  $\mathcal{S}_{\text{CS}}$  są zamknięte ze względu na obustronne mnożenie (por. Wprowadzenie, Rozdział 1.2) spełniony jest warunek (ii). Zatem możemy podsumować, że dla  $k \in \{E, Q\}$

- $\zeta_{\text{UN}}^{(k)}$  nie zależy od wyboru  $\boldsymbol{\Psi}$  i  $\boldsymbol{\Sigma}$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$  (w szczególności  $\boldsymbol{\Psi}$  może mieć postać CS lub AR(1)),
- $\zeta_{\text{CS}}^{(k)}$  nie zależy od wyboru  $\varrho$  i  $\boldsymbol{\Sigma}$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}$ .

Wiadomo, że powyższe rozważania nie dotyczą zbioru  $\mathcal{S}_{\text{AR}}$ . Zajmijmy się jednak osobno czynnikami iloczynu Kroneckera, tzn.  $\mathbf{A}_1\boldsymbol{\Psi}_{\text{AR}}\mathbf{A}'_1$  oraz  $\mathbf{A}_2\boldsymbol{\Sigma}\mathbf{A}'_2$ . Ponieważ zgodnie z założeniami  $\mathbf{A}_2$  jest dowolną macierzą nieosobliwą oraz  $\boldsymbol{\Sigma}$  jest dowolną macierzą określoną dodatnio, to  $\mathbf{A}_2\boldsymbol{\Sigma}\mathbf{A}'_2$  generuje cały zbiór macierzy określonych dodatnio. Spełnione są zatem założenia Lematu 1.1 oraz warunek (i). Własności te nie zachodzą dla  $\mathbf{A}_1\boldsymbol{\Psi}_{\text{AR}}\mathbf{A}'_1$ . Możemy zatem powiedzieć, że

- $\zeta_{\text{AR}}^{(k)}$  nie zależy od wyboru  $\boldsymbol{\Sigma}$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{AR}} \otimes \boldsymbol{\Sigma}$ ,
- $\zeta_{\text{AR}}^{(k)}$  zależy od wyboru  $\varrho$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{AR}} \otimes \boldsymbol{\Sigma}$ .

Stosując takie samo rozumowanie otrzymujemy następujące wnioski:

- $\zeta_{\text{UN}}^{(k)}$  nie zależy od wyboru  $\boldsymbol{\Sigma}$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}$  lub  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{AR}} \otimes \boldsymbol{\Sigma}$ ,
- $\zeta_{\text{CS}}^{(k)}$  nie zależy od wyboru  $\boldsymbol{\Sigma}$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$  lub  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{AR}} \otimes \boldsymbol{\Sigma}$ ,
- $\zeta_{\text{AR}}^{(k)}$  nie zależy od wyboru  $\boldsymbol{\Sigma}$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$  lub  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}$ .

Na zakończenie zauważmy, że jeżeli  $\mathbf{A}_1$  spełnia założenie (Z1) to w ogólności nie spełnia założenia (Z2), a zatem  $\mathbf{A}_1\boldsymbol{\Psi}_{\text{AR}}\mathbf{A}'_1$  nie posiada struktury AR(1). Podobnie, jeżeli  $\mathbf{A}_1$  spełnia założenie (Z2), to w ogólności nie spełnia założenia (Z1), a zatem  $\mathbf{A}_1\boldsymbol{\Psi}_{\text{CS}}\mathbf{A}'_1$  nie posiada struktury CS. W obu tych przypadkach nie są spełnione założenia Lematu 1.1, a więc

- $\zeta_{\text{CS}}^{(k)}$  zależy od wyboru  $\varrho$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{AR}} \otimes \boldsymbol{\Sigma}$ ,
- $\zeta_{\text{AR}}^{(k)}$  zależy od wyboru  $\varrho$  gdy  $\boldsymbol{\Omega} = \boldsymbol{\Psi}_{\text{CS}} \otimes \boldsymbol{\Sigma}$ . ■

Z powyższego lematu wynika, że w przeprowadzonych badaniach symulacyjnych możemy przyjąć  $\boldsymbol{\Sigma} = \mathbf{I}_q$ , natomiast wartości  $\varrho$  w określonej dodatnio macierzy  $\boldsymbol{\Psi}_{\text{CS}}$

i  $\Psi_{AR}$  muszą być różne. Przyjmować będziemy  $\varrho$  ze zbioru  $\{-0.9, -0.5, -0.1, 0, 0.1, 0.5, 0.9\}$  takie, że macierz  $\Psi_{CS}$  lub  $\Psi_{AR}$  jest określona dodatnio. Symulacje przeprowadzone zostały dla  $n = 100$  oraz  $(p, q)$  ze zbioru  $\{(3, 3), (10, 3), (15, 3), (3, 5), (10, 5), (15, 5)\}$ . Ze względu na założenie  $pq < n$  oraz złożoność obliczeniową algorytmów dla kwadratowej funkcji straty nie uwzględniamy większych wartości  $p$ .

Z uwagi na fakt, że  $\mathcal{S}_{CS}$  i  $\mathcal{S}_{AR}$  są podzbiarami  $\mathcal{S}$ , rozbieżność od zbioru struktur rozdzielnych  $\mathcal{S}$  jest zawsze najmniejsza i w związku z tym uzasadnione jest rozważanie struktur kowariancyjnych tylko ze zbiorów  $\mathcal{S}_{CS}$  i  $\mathcal{S}_{AR}$ . Zatem wygenerowano dane

$$\mathbf{Y} \sim N_{n,pq}(\mathbf{0}, \mathbf{I}_n, \Psi_{CS} \otimes \mathbf{I}_q) \quad \text{oraz} \quad \mathbf{Y} \sim N_{n,pq}(\mathbf{0}, \mathbf{I}_n, \Psi_{AR} \otimes \mathbf{I}_q)$$

dla różnych wartości parametrów  $\varrho$ .

Ponieważ w rzeczywistości nie znamy prawdziwej struktury kowariancyjnej, po wygenerowaniu danych uznajemy ją za nieznaną i w dalszych badaniach posługujemy się MLE nieustrukturyzowanej macierzy kowariancji postaci

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

Zatem do wyznaczenia indeksów rozbieżności (1.5) przyjmować będziemy  $\mathbf{\Omega} = \mathbf{S}$ .

Ten tok rozumowania powtarzany jest 1000 razy, dlatego wartości indeksów rozbieżności są uśredniane, a własności estymatorów są analizowane pod kątem średniej i odchylenia standardowego.

Przypomnijmy, że zgodnie z Twierdzeniem 2.2 tylko wartości  $\zeta_{CS}^{(E)}$  są wyznaczone jednoznacznie, natomiast aproksymacje w pozostałych zbiorach oraz aproksymacje kwadratową funkcją straty nie gwarantują jednoznaczności. Niemniej jednak w przeprowadzonych badaniach symulacyjnych we wszystkich przypadkach istniało dokładnie jedno minimum. Ponadto, w przypadku kwadratowej funkcji straty minima wyznaczone na podstawie Twierdzeń 3.1–3.3 niekoniecznie są określone dodatnio. W niniejszych symulacjach sytuacja taka miała miejsce dla  $(p, q) = (15, 5)$ , gdzie spośród 1000 symulowanych danych z rozkładu normalnego z macierzą kowariancji należąca do  $\mathcal{S}_{CS}$ , dla  $\varrho = 0$  w 40 przypadkach stwierdzono nieokreśloność rozwiązania, natomiast dla  $\varrho = 0.1$ ,  $\varrho = 0.5$  i  $\varrho = 0.9$  odpowiednio w 36, 40 i 45 przypadkach. Dla danych generowanych z rozkładu normalnego z macierzą kowariancji należąca do  $\mathcal{S}_{AR}$  nieokreślone rozwiązanie otrzymano czterokrotnie dla  $\varrho = -0.9$  i pięciokrotnie dla  $\varrho = 0.9$ . Wyniki te jako niespełniające założenia o dodatniej określoności macierzy kowariancji, zostały usunięte z dalszych rozważań.

Rezultaty dotyczące identyfikacji struktury przedstawia Rozdział 4.2, natomiast wyniki związane z estymacją omówione są w Rozdziale 4.3.

W pracy Filipiak i in. [12, 13] testowane były hipotezy dotyczące struktury za pomocą testu ilorazu wiarygodności (ang. *likelihood ratio test*, LRT) oraz testu wynikowego Rao (ang. *Rao score test*, RST). W pracy Filipiak i in. [7] wykorzystano normę Frobeniusa (1.2) jako miarę rozbieżności między hipotezami o różnych strukturach rozdzielnych do zbadania mocy LRT. Zauważmy jednak, że norma Frobeniusa jest interpretowana jako algebraiczna odległość między dwiema macierzami, podczas gdy entropijna funkcja straty i kwadratowa funkcja straty reprezentują rozbieżności między dwoma rozkładami prawdopodobieństwa o różnych macierzach kowariancji. Dlatego celem naszych badań jest zweryfikowanie czy entropijna i kwadratowa funkcja straty również mogą być użyte jako miary rozbieżności w kontekście mocy testów.

Wiadomo, że statystyki testowe mają następujące postaci:

$$\begin{aligned} \text{LRT} &= n \operatorname{tr} \left[ \hat{\Omega} \left( \hat{\Psi} \otimes \hat{\Sigma} \right)^{-1} \right] - n \ln \left| \hat{\Omega} \left( \hat{\Psi} \otimes \hat{\Sigma} \right)^{-1} \right| - npq, \\ \text{RST} &= \frac{n}{2} \operatorname{tr} \left\{ \left[ \hat{\Omega} \left( \hat{\Psi} \otimes \hat{\Sigma} \right)^{-1} - \mathbf{I}_{pq} \right]^2 \right\}; \end{aligned}$$

por. Filipiak i Klein [8]. Zauważmy, że statystyka LRT ma, z dokładnością do stałej, postać entropijnej funkcji straty z zamienioną rolą macierzy bez struktury i macierzy o strukturze rozdzielnej. Podobnie statystyka RST ma, z dokładnością do stałej, postać kwadratowej funkcji straty również z zamienionymi rolami macierzy. W związku z tym nasuwa się przypuszczenie, że w badaniu mocy testów entropijna i kwadratowa funkcja straty mogą odegrać istotną rolę. W zależności od wyboru funkcji rozbieżności między zbiorami struktur w analizowanych hipotezach, entropijna funkcja straty może wskazywać na większą moc testu LRT, natomiast kwadratowa funkcja straty na większą moc testu RST. W Rozdziale 4.4 zweryfikujemy to przypuszczenie.

Rozdział 4.5 ilustruje metodę identyfikacji struktury kowariancji poprzez entropijną i kwadratową funkcję straty oraz normę Frobeniusa na danych rzeczywistych.

## 4.2. Identyfikacja struktury

W tym rozdziale przedstawione wcześniej metody aproksymacji zostaną zastosowane do sprawdzenia, czy odpowiednia funkcja rozbieżności prawidłowo rozpoznaje prawdziwą strukturę kowariancji.

Jeśli uśrednione indeksy rozbieżności są mniejsze (lub równe) dla odpowiednich struktur, to jest

$$\bar{\zeta}_{\text{CS}}^{(k)} \leq \bar{\zeta}_{\text{AR}}^{(k)}, \text{ gdy } \Omega \in \mathcal{S}_{\text{CS}} \quad \text{oraz} \quad \bar{\zeta}_{\text{AR}}^{(k)} \leq \bar{\zeta}_{\text{CS}}^{(k)}, \text{ gdy } \Omega \in \mathcal{S}_{\text{AR}}, \quad k \in \{E, Q\},$$

wnioskujemy, że funkcja prawidłowo rozpoznaje strukturę kowariancyjną; por. Filipiak i in [11]. Jako dodatkowy wskaźnik wyznaczamy proporcję poprawnie zidentyfikowanych struktur.

Algorytm postępowania wygląda następująco:

1. generowanie danych z rozkładu  $N_{n,qp}(\mathbf{0}, \mathbf{I}_n, \mathbf{\Omega})$ , gdzie  $\mathbf{\Omega} \in \mathcal{S}_{CS}$  lub  $\mathbf{\Omega} \in \mathcal{S}_{AR}$ ;
2. wyliczenie  $\mathbf{S}$ ;
3. obliczenie rozbieżności

$$\zeta_{UN}^{(k)} = \min_{\Psi, \Sigma} L_S^{(k)}(\Psi \otimes \Sigma), \quad \zeta_{CS}^{(k)} = \min_{\varrho, \Sigma} L_S^{(k)}(\Psi_{CS} \otimes \Sigma), \quad \zeta_{AR}^{(k)} = \min_{\varrho, \Sigma} L_S^{(k)}(\Psi_{AR} \otimes \Sigma);$$

4. powtórzenie 1000 razy kroków 1–3;
5. uśrednienie  $\zeta_{UN}^{(k)}$ ,  $\zeta_{CS}^{(k)}$  and  $\zeta_{AR}^{(k)}$ ;
6. zweryfikowanie czy

$$\begin{aligned} \mathbf{\Omega} \in \mathcal{S}_{CS} &\implies \bar{\zeta}_{CS}^{(k)} \leq \bar{\zeta}_{AR}^{(k)}, \\ \mathbf{\Omega} \in \mathcal{S}_{AR} &\implies \bar{\zeta}_{AR}^{(k)} \leq \bar{\zeta}_{CS}^{(k)}; \end{aligned}$$

7. wyznaczenie proporcji poprawnie zidentyfikowanych struktur

$$\pi^{(k)} = \frac{\text{liczba poprawnie zidentyfikowanych struktur}}{1000} \cdot 100\%.$$

Ponieważ  $\mathcal{S}_{CS}, \mathcal{S}_{AR} \subset \mathcal{S}$  to spełnione muszą być nierówności  $\bar{\zeta}_{UN}^{(k)} \leq \bar{\zeta}_{CS}^{(k)}$  oraz  $\bar{\zeta}_{UN}^{(k)} \leq \bar{\zeta}_{AR}^{(k)}$ . Sama wartość  $\zeta_{UN}^{(k)}$  nie jest więc interesująca dla identyfikacji struktury i pełni rolę kontrolną dla algorytmu.

W Tabelach 4.1–4.4 zaprezentowane zostały uśrednione indeksy rozbieżności dla każdej z rozważanych funkcji rozbieżności oraz proporcje poprawnie zidentyfikowanych struktur. Zgodnie z oczekiwaniami, dla każdego zestawu parametrów, zarówno dla entropijnej jak i kwadratowej funkcji straty uśredniona rozbieżność  $\bar{\zeta}_{UN}^{(k)}$  jest najmniejsza (wśród  $\bar{\zeta}_{UN}^{(k)}$ ,  $\bar{\zeta}_{CS}^{(k)}$  i  $\bar{\zeta}_{AR}^{(k)}$ ,  $k \in \{E, Q\}$ ). Ponadto, jeśli  $\varrho = 0$  to  $\mathcal{S}_{CS} = \mathcal{S}_{AR}$ , a zatem  $\bar{\zeta}_{CS}^{(k)}$  i  $\bar{\zeta}_{AR}^{(k)}$  są równe. Przypomnijmy, że wszystkie wartości zamieszczone w Tabelach 4.1–4.4 są wynikiem skomplikowanych algorytmów numerycznych, których rozwiązanie wymagało użycia metod iteracyjnych. Z tego powodu wartości, które dla  $\varrho = 0$  lub na mocy Lematu 4.1 powinny być równe, mogą wykazywać nieznaczne różnice.

Tabela 4.1. Uśrednione rozbieżności oraz proporcja poprawnie zidentyfikowanych struktur dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{CS} \otimes \mathbf{I}_3)$ .

$p$	$\varrho$	Entropijna funkcja straty				Kwadratowa funkcja straty			
		$\bar{\zeta}_{UN}^{(E)}$	$\bar{\zeta}_{CS}^{(E)}$	$\bar{\zeta}_{AR}^{(E)}$	$\pi^{(E)}$	$\bar{\zeta}_{UN}^{(Q)}$	$\bar{\zeta}_{CS}^{(Q)}$	$\bar{\zeta}_{AR}^{(Q)}$	$\pi^{(Q)}$
3	-0.1	0.3673	0.4101	0.4468	83.3	0.6656	0.7384	0.7916	82.3
3	0	0.3673	0.4101	0.4104	49.5	0.6680	0.7457	0.7455	49.4
3	0.1	0.3673	0.4101	0.4352	79.0	0.6685	0.7446	0.7823	77.3
3	0.5	0.3673	0.4101	0.7680	99.7	0.6648	0.7405	1.2324	99.7
3	0.9	0.3673	0.4101	1.1689	100.0	0.6697	0.7467	1.6168	100.0
10	-0.1	5.1192	5.8713	14.6891	100.0	7.5544	9.2058	16.0191	100.0
10	0	5.1192	5.8713	5.8705	51.3	7.5186	9.1562	9.1570	50.4
10	0.1	5.1192	5.8713	6.3907	100.0	7.5397	9.1748	9.5268	98.4
10	0.5	5.1192	5.8713	9.8416	100.0	7.5396	9.1722	10.3823	100.0
10	0.9	5.1192	5.8713	15.8255	100.0	7.5224	9.1755	20.6869	100.0
15	0	13.2331	15.3668	15.3673	48.1	16.1694	20.6584	20.6600	50.9
15	0.1	13.2331	15.3668	16.2939	100.0	16.1766	20.7239	21.0452	95.4
15	0.5	13.2331	15.3668	20.5370	100.0	16.1810	20.6695	21.9300	99.7
15	0.9	13.2331	15.3668	26.7463	100.0	16.1766	20.7239	35.7882	100.0

Tabela 4.2. Uśrednione rozbieżności oraz proporcja poprawnie zidentyfikowanych struktur dla danych generowanych z rozkładu  $N_{100,5p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{CS} \otimes \mathbf{I}_5)$ .

$p$	$\varrho$	Entropijna funkcja straty				Kwadratowa funkcja straty			
		$\bar{\zeta}_{UN}^{(E)}$	$\bar{\zeta}_{CS}^{(E)}$	$\bar{\zeta}_{AR}^{(E)}$	$\pi^{(E)}$	$\bar{\zeta}_{UN}^{(Q)}$	$\bar{\zeta}_{CS}^{(Q)}$	$\bar{\zeta}_{AR}^{(Q)}$	$\pi^{(Q)}$
3	-0.1	1.1187	1.1626	1.2237	88.3	1.9417	2.0125	2.0847	85.9
3	0	1.1187	1.1626	1.1628	50.5	1.9373	2.0102	2.0096	48.0
3	0.1	1.1187	1.1626	1.2037	83.5	1.9365	2.0080	2.0544	78.2
3	0.5	1.1187	1.1626	1.7556	100.0	1.9528	2.0258	2.7004	99.8
3	0.9	1.1187	1.1626	2.4229	100.0	1.9466	2.0187	3.2529	100.0
10	-0.1	18.8921	19.9532	34.4452	100.0	22.6178	24.9413	30.9468	100.0
10	0	18.8921	19.9532	19.9526	47.7	22.6429	24.9688	24.9664	48.7
10	0.1	18.8921	19.9532	20.8259	100.0	22.5908	24.9008	25.1874	92.7
10	0.5	18.8921	19.9532	26.5757	100.0	22.5610	24.8555	25.9044	99.7
10	0.9	18.8921	19.9532	36.5468	100.0	22.6213	24.9214	36.1259	100.0
15	0	62.1249	66.8870	66.8921	53.2	47.8388	55.7290	57.5923	49.6
15	0.1	62.1249	66.8870	68.4512	100.0	47.8383	55.7367	57.6013	66.6
15	0.5	62.1249	66.8870	75.5178	100.0	47.8401	55.7347	57.6007	79.9
15	0.9	62.1249	66.8870	85.8640	100.0	47.8380	55.7302	57.5979	95.0

Tabela 4.3. Uśrednione rozbieżności oraz proporcja poprawnie zidentyfikowanych struktur dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{AR} \otimes \mathbf{I}_3)$ .

		Entropijna funkcja straty				Kwadratowa funkcja straty			
$p$	$\varrho$	$\bar{\zeta}_{UN}^{(E)}$	$\bar{\zeta}_{CS}^{(E)}$	$\bar{\zeta}_{AR}^{(E)}$	$\pi^{(E)}$	$\bar{\zeta}_{UN}^{(Q)}$	$\bar{\zeta}_{CS}^{(Q)}$	$\bar{\zeta}_{AR}^{(Q)}$	$\pi^{(Q)}$
3	-0.9	0.3673	5.6723	0.4102	100.0	0.6656	2.9835	0.7385	100.0
3	-0.5	0.3673	1.2159	0.4105	100.0	0.6680	1.5982	0.7458	100.0
3	-0.1	0.3673	0.4315	0.4105	77.1	0.6685	0.7764	0.7444	75.9
3	0	0.3673	0.4101	0.4104	50.5	0.6688	0.7436	0.7441	48.6
3	0.1	0.3673	0.4283	0.4104	73.1	0.6648	0.7676	0.7406	71.3
3	0.5	0.3673	0.7327	0.4101	99.8	0.6635	1.1854	0.7424	99.4
3	0.9	0.3673	1.1618	0.4099	100.0	0.6697	1.6283	0.7465	99.9
10	-0.9	5.1192	24.7209	5.8707	100.0	7.5544	14.7257	9.2062	100.0
10	-0.5	5.1192	11.5063	5.8705	100.0	7.5186	12.5526	9.1550	100.0
10	-0.1	5.1192	6.0907	5.8704	98.8	7.5397	9.3764	9.1783	91.7
10	0	5.1192	5.8713	5.8705	48.7	7.5262	9.1647	9.1627	50.5
10	0.1	5.1192	6.0799	5.8706	98.0	7.5396	9.3499	9.1703	89.7
10	0.5	5.1192	10.1883	5.8714	100.0	7.5441	11.8149	9.1835	100.0
10	0.9	5.1192	16.3827	5.8715	100.0	7.5224	13.3154	9.1743	100.0
15	-0.9	13.2331	43.2050	15.3688	100.0	16.1694	25.9321	20.6622	100.0
15	-0.5	13.2331	24.3583	15.3678	100.0	16.1766	24.0034	20.7219	100.0
15	-0.1	13.2331	15.7331	15.3673	99.2	16.1458	20.8070	20.6064	86.9
15	0	13.2331	15.3668	15.3673	51.9	16.1810	20.6695	20.6709	49.5
15	0.1	13.2331	15.7220	15.3674	99.3	16.1384	20.7976	20.6146	86.4
15	0.5	13.2331	22.8995	15.3679	100.0	16.1350	23.3561	20.5539	100.0
15	0.9	13.2331	34.0544	15.3669	100.0	16.1736	24.9547	20.6650	100.0

Tabela 4.4. Uśrednione rozbieżności oraz proporcja poprawnie zidentyfikowanych struktur dla danych generowanych z rozkładu  $N_{100,5p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{\text{AR}} \otimes \mathbf{I}_5)$ .

		Entropijna funkcja straty				Kwadratowa funkcja straty			
$p$	$\varrho$	$\bar{\zeta}_{\text{UN}}^{(E)}$	$\bar{\zeta}_{\text{CS}}^{(E)}$	$\bar{\zeta}_{\text{AR}}^{(E)}$	$\pi^{(E)}$	$\bar{\zeta}_{\text{UN}}^{(Q)}$	$\bar{\zeta}_{\text{CS}}^{(Q)}$	$\bar{\zeta}_{\text{AR}}^{(Q)}$	$\pi^{(Q)}$
3	-0.9	1.1187	9.9190	1.1629	100.0	1.9417	5.3227	2.0136	100.0
3	-0.5	1.1187	2.5031	1.1630	100.0	1.9373	3.2242	2.0092	100.0
3	-0.1	1.1187	1.1989	1.1629	82.1	1.9365	2.0480	2.0074	77.1
3	0	1.1187	1.1626	1.1628	49.5	1.9485	2.0187	2.0193	50.6
3	0.1	1.1187	1.1919	1.1628	78.9	1.9528	2.0634	2.0268	76.2
3	0.5	1.1187	1.6923	1.1626	100.0	1.9351	2.6372	2.0048	100.0
3	0.9	1.1187	2.4013	1.1626	100.0	1.9466	3.2567	2.0187	100.0
10	-0.9	18.8921	51.2986	19.9531	100.0	22.6178	29.9053	24.9458	100.0
10	-0.5	18.8921	29.3081	19.9530	100.0	22.6429	27.8461	24.9641	100.0
10	-0.1	18.8921	20.3183	19.9527	99.7	22.5908	25.0613	24.9009	83.7
10	0	18.8921	19.9532	19.9526	52.3	22.5982	24.9289	24.9234	51.4
10	0.1	18.8921	20.2975	19.9525	99.1	22.5610	25.0013	24.8571	81.3
10	0.5	18.8921	27.1100	19.9524	100.0	22.5945	27.1032	24.8945	100.0
10	0.9	18.8921	37.4108	19.9530	100.0	22.6213	28.5374	24.9214	100.0
15	-0.9	62.1249	113.0665	66.8903	100.0	47.8414	56.4771	55.7352	98.8
15	-0.5	62.1249	81.7438	66.8917	100.0	47.8262	56.7270	55.6779	97.2
15	-0.1	62.1249	67.4898	66.8921	99.2	47.8274	55.7968	55.7438	58.0
15	0	62.1249	66.8870	66.8922	46.8	47.9225	55.8428	55.8495	49.0
15	0.1	62.1249	67.4737	66.8923	99.8	47.8262	55.7413	55.6759	61.6
15	0.5	62.1249	79.3455	66.8929	100.0	47.8274	56.6595	55.7428	94.3
15	0.9	62.1249	97.8834	66.8900	100.0	47.8427	56.3616	55.7456	97.1

Z Tabel 4.1 i 4.2 można zauważyć, że dla wszystkich zestawów parametrów (nie uwzględniając  $\varrho = 0$ ) oraz dla obu funkcji rozbieżności spełniona jest nierówność  $\bar{\zeta}_{\text{CS}}^{(k)} < \bar{\zeta}_{\text{AR}}^{(k)}$ . Zatem możemy wywnioskować, że zarówno entropijna jak i kwadratowa funkcja straty dobrze rozpoznają strukturę kowariancyjną ze zbioru  $\mathcal{S}_{\text{CS}}$ . Wniosek ten jest potwierdzony proporcją poprawnie rozpoznanych struktur. Proporcja rzędu 50% dla  $\varrho = 0$  wynika z równości  $\mathcal{S}_{\text{CS}} = \mathcal{S}_{\text{AR}}$ . Wraz ze wzrostem  $\varrho$  różnica między strukturą CS i AR(1) jest bardziej dostrzegalna, stąd większe wartości  $\pi^{(k)}$  dla  $|\varrho| > 0.1$ ,  $k \in \{E, Q\}$ . Podobnie, wraz ze wzrostem  $p$  różnice między strukturami CS i AR(1) są coraz większe, stąd nawet dla małych wartości  $\varrho$  struktura CS powinna być lepiej rozpoznawana. Tę obserwację potwierdzają rosnące wartości  $\pi^{(k)}$  dla ustalonego  $\varrho$  i rosnącego  $p$ . Wyjątek stanowią wartości otrzymane dla  $p = 15$  dla kwadratowej funkcji straty. Może być to wynikiem numerycznej złożoności algorytmów wyznaczających  $\zeta_{\text{CS}}^{(Q)}$  i  $\zeta_{\text{AR}}^{(Q)}$ , jednak weryfikacja tego spostrzeżenia

wymagałaby dodatkowych badań nad optymalizacją algorytmów. Zwróćmy uwagę, że zastosowanie algorytmów wykorzystywanych w pracy do rzeczywistych danych nie stanowi problemu, nawet jeżeli zwiększymy ich dokładność, jednakże dla badań symulacyjnych proces ten jest bardzo czasochłonny.

Z Tabel 4.3 i 4.4 można dostrzec, że dla wszystkich zestawów parametrów (nie uwzględniając  $\varrho = 0$ ) oraz dla obu funkcji rozbieżności spełniona jest nierówność  $\bar{\zeta}_{\text{AR}}^{(k)} < \bar{\zeta}_{\text{CS}}^{(k)}$ . Proporcja rzędu 50% dla  $\varrho = 0$  wynika z równości zbiorów  $\mathcal{S}_{\text{CS}}$  i  $\mathcal{S}_{\text{AR}}$ . Wraz ze wzrostem  $|\varrho|$  różnica między strukturą CS i AR(1) jest bardziej widoczna, stąd większe wartości  $\pi^{(k)}$  dla  $|\varrho| > 0.1$ . Podobnie, wraz ze wzrostem  $p$  różnice między strukturami CS i AR(1) są coraz większe, stąd nawet dla małych wartości  $\varrho$  struktura AR(1) jest zazwyczaj lepiej rozpoznawana. Jednak porównując  $\pi^{(Q)}$  dla  $p = 15$  można zauważyć spadek proporcji poprawnie zidentyfikowanych struktur. Podobnie jak w przypadku  $\Psi_{\text{CS}} \otimes \mathbf{I}_q$  różnice mogą wynikać z kumulacji błędów numerycznych przy zastosowaniu algorytmu wyznaczającego najlepsze przybliżenie kwadratową funkcją straty.

Podsumowując, zarówno dla danych generowanych z rozkładu o strukturze kowariancji ze zbioru  $\mathcal{S}_{\text{CS}}$  jak i  $\mathcal{S}_{\text{AR}}$ , obie funkcje rozbieżności prawidłowo identyfikują strukturę kowariancyjną.

### 4.3. Estymacja

Macierze  $\Psi$  i  $\Sigma$  otrzymane w procesie minimalizacji entropijnej lub kwadratowej funkcji straty mogą być traktowane jako estymatory nieznanych czynników struktury rozdzielnej. Przypomnijmy, że estymatory otrzymane poprzez minimalizację entropijnej funkcji straty (odpowiednio kwadratowej funkcji straty) oznaczamy przez ELE (odpowiednio QLE). W tym rozdziale porównamy własności ELE i QLE z powszechnie stosowanymi estymatorami największej wiarygodności (MLE).

W Rozdziale 2.2 pokazane zostało, że algorytmy wyznaczania MLE i ELE struktury  $\Psi_{\text{CS}} \otimes \Sigma$  odpowiednio z (2.18) i (2.17) są takie same, jednak różnią się danymi wejściowymi. W związku z tym wyznaczone estymatory mają różne własności statystyczne. W przypadku estymatorów QLE struktury  $\Psi_{\text{CS}} \otimes \Sigma$  oraz estymatorów ELE, QLE i MLE struktury  $\Psi_{\text{AR}} \otimes \Sigma$  wszystkie algorytmy są różne.

Przypomnijmy, że zgodnie z Twierdzeniem 2.2 wartości ELE dla  $\Omega = \Psi_{\text{CS}} \otimes \Sigma$  są wyznaczone jednoznacznie. W pozostałych przypadkach jednoznaczność rozwiązania nie została udowodniona analitycznie, jednak we wszystkich badaniach symulacyjnych istniało dokładnie jedno minimum.

Stosować będziemy analogiczny algorytm jak w Rozdziale 4.2, przy czym interesować nas będą uśrednione wartości estymatorów ze względu na 1000 powtórzeń i ich odchylenia standardowe. Zatem algorytm postępowania wygląda następująco:

1. generowanie danych z rozkładu  $N_{n,qp}(\mathbf{0}, \mathbf{I}_n, \mathbf{\Omega})$ , gdzie  $\mathbf{\Omega} \in \mathcal{S}_{CS}$  lub  $\mathbf{\Omega} \in \mathcal{S}_{AR}$ ;
2. wyznaczenie  $\mathbf{S}$ ;
3. wyznaczenie estymatorów

$$\begin{aligned} (\check{\varrho}, \check{\mathbf{\Sigma}}) &= \operatorname{argmin}_{\varrho, \mathbf{\Sigma}} L_S^{(E)}(\mathbf{\Psi}_{CS} \otimes \mathbf{\Sigma}), & (\check{\varrho}, \check{\mathbf{\Sigma}}) &= \operatorname{argmin}_{\varrho, \mathbf{\Sigma}} L_S^{(E)}(\mathbf{\Psi}_{AR} \otimes \mathbf{\Sigma}), \\ (\tilde{\varrho}, \tilde{\mathbf{\Sigma}}) &= \operatorname{argmin}_{\varrho, \mathbf{\Sigma}} L_S^{(Q)}(\mathbf{\Psi}_{CS} \otimes \mathbf{\Sigma}), & (\tilde{\varrho}, \tilde{\mathbf{\Sigma}}) &= \operatorname{argmin}_{\varrho, \mathbf{\Sigma}} L_S^{(Q)}(\mathbf{\Psi}_{AR} \otimes \mathbf{\Sigma}); \end{aligned}$$

4. powtórzenie 1000 razy kroków 1–3 dla każdego ze zbiorów;
5. uśrednienie  $\check{\varrho}, \check{\mathbf{\Sigma}}, \tilde{\varrho}, \tilde{\mathbf{\Sigma}}$  dla każdego ze zbiorów;
6. obliczenie odchyłeń standardowych estymatorów dla każdego ze zbiorów.

Uśrednione wartości estymatorów i ich odchylenia standardowe (lub normy Frobeniusa tych wartości) dla danych generowanych z rozkładu  $N_{100,pq}(\mathbf{0}, \mathbf{I}_{100}, \mathbf{\Psi}_{CS} \otimes \mathbf{I}_q)$  dla różnych wartości parametrów  $p$  i  $q$  oraz  $\varrho \in \{0, 0.5, 0.9\}$  zaprezentowane są w Tabelach 4.5–4.9.

Tabela 4.5. Uśrednione wartości (av) i odchylenia standardowe (std) estymatorów  $\varrho$  dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \mathbf{\Psi}_{CS} \otimes \mathbf{I}_3)$ .

		MLE		ELE		QLE	
$p$	$\varrho$	av( $\hat{\varrho}$ )	std( $\hat{\varrho}$ )	av( $\check{\varrho}$ )	std( $\check{\varrho}$ )	av( $\tilde{\varrho}$ )	std( $\tilde{\varrho}$ )
3	0	0.0010	0.0331	0.0022	0.0342	0.0030	0.0381
3	0.5	0.4999	0.0331	0.5010	0.0342	0.5014	0.0375
3	0.9	0.8997	0.0093	0.9000	0.0096	0.8997	0.0105
10	0	-0.0001	0.0085	0.0006	0.0103	0.0012	0.0168
10	0.5	0.4988	0.0235	0.5004	0.0282	0.5014	0.0466
10	0.9	0.8994	0.0079	0.8998	0.0094	0.9001	0.0158
15	0	-0.0002	0.0057	0.0004	0.0079	0.0032	0.0175
15	0.5	0.4985	0.0228	0.5000	0.0315	0.5000	0.0714
15	0.9	0.8992	0.0079	0.8995	0.0109	0.8982	0.0277

Tabela 4.6. Uśrednione wartości (av) i odchylenia standardowe (std) estymatorów  $\varrho$  dla danych generowanych z rozkładu  $N_{100,5p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{CS} \otimes \mathbf{I}_5)$ .

		MLE		ELE		QLE	
$p$	$\varrho$	av( $\hat{\varrho}$ )	std( $\hat{\varrho}$ )	av( $\check{\varrho}$ )	std( $\check{\varrho}$ )	av( $\tilde{\varrho}$ )	std( $\tilde{\varrho}$ )
3	0	-0.0004	0.0266	0.0004	0.0276	-0.0015	0.0301
3	0.5	0.4989	0.0266	0.4996	0.0276	0.4992	0.0320
3	0.9	0.8995	0.0075	0.8997	0.0078	0.8999	0.0091
10	0	0.0004	0.0066	0.0009	0.0095	0.0009	0.0224
10	0.5	0.5007	0.0181	0.5014	0.0257	0.4973	0.0646
10	0.9	0.9001	0.0060	0.9002	0.0085	0.8972	0.0229
15	0	0.0003	0.0044	0.0014	0.0093	0.0124	0.0469
15	0.5	0.5006	0.0176	0.5030	0.0365	0.4910	0.1750
15	0.9	0.9001	0.0060	0.9004	0.0125	0.8669	0.1309

Z Tabel 4.5 i 4.6 można dostrzec, że dla wszystkich zestawów parametrów obciążenie każdego z estymatorów jest niewielkie i nie zależy od parametrów  $p$  i  $q$ . Odchylenia standardowe MLE i ELE są porównywalne, podczas gdy dla estymatorów QLE są nieznacznie większe.

Tabela 4.7. Normy Frobeniusa z uśrednionych wartości (av) i odchylen standardowych (std) estymatorów  $\Sigma$  dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{CS} \otimes \mathbf{I}_3)$ .

		MLE		ELE		QLE	
$p$	$\varrho$	$\ \text{av}(\hat{\Sigma})\ _F^2$	$\ \text{std}(\hat{\Sigma})\ _F^2$	$\ \text{av}(\check{\Sigma})\ _F^2$	$\ \text{std}(\check{\Sigma})\ _F^2$	$\ \text{av}(\tilde{\Sigma})\ _F^2$	$\ \text{std}(\tilde{\Sigma})\ _F^2$
3	0	2.942	0.039	2.463	0.036	2.070	0.039
3	0.5	2.948	0.042	2.473	0.039	2.098	0.044
3	0.9	2.954	0.050	2.479	0.046	2.098	0.052
10	0	2.944	0.012	1.412	0.009	0.687	0.012
10	0.5	2.942	0.016	1.420	0.012	0.693	0.016
10	0.9	2.938	0.025	1.426	0.019	0.710	0.024
15	0	2.942	0.008	0.857	0.005	0.252	0.007
15	0.5	2.936	0.012	0.862	0.007	0.259	0.010
15	0.9	2.930	0.023	0.866	0.013	0.268	0.018

Tabela 4.8. Normy Frobeniusa z uśrednionych wartości (av) i odchyłeń standardowych (std) estymatorów  $\Sigma$  dla danych generowanych z rozkładu  $N_{100,5p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{CS} \otimes \mathbf{I}_5)$ .

		MLE		ELE		QLE	
$p$	$\varrho$	$\ \text{av}(\widehat{\Sigma})\ _F^2$	$\ \text{std}(\widehat{\Sigma})\ _F^2$	$\ \text{av}(\widetilde{\Sigma})\ _F^2$	$\ \text{std}(\widetilde{\Sigma})\ _F^2$	$\ \text{av}(\widetilde{\Sigma})\ _F^2$	$\ \text{std}(\widetilde{\Sigma})\ _F^2$
3	0	4.907	0.100	3.636	0.087	2.733	0.095
3	0.5	4.905	0.104	3.640	0.091	2.729	0.098
3	0.9	4.901	0.112	3.640	0.097	2.752	0.105
10	0	4.901	0.029	1.184	0.015	0.305	0.022
10	0.5	4.920	0.033	1.194	0.018	0.313	0.024
10	0.9	4.936	0.043	1.201	0.023	0.313	0.032
15	0	4.903	0.020	0.278	0.006	0.020	0.006
15	0.5	4.922	0.024	0.282	0.007	0.022	0.009
15	0.9	4.940	0.035	0.287	0.010	0.023	0.016

Tabela 4.9. Uśrednione wartości (av) oraz odchylenia standardowe (std) estymatorów macierzy  $\Sigma$  dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{CS} \otimes \mathbf{I}_3)$  gdy  $\varrho = 0.5$ .

		MLE			ELE			QLE		
$p$		$\widehat{\Sigma}$			$\widetilde{\Sigma}$			$\widetilde{\Sigma}$		
3	av	0.992	-0.001	0.000	0.909	-0.002	-0.001	0.835	0.000	-0.001
		-0.001	0.993	-0.001	-0.002	0.909	0.000	0.000	0.838	-0.001
		0.000	-0.001	0.989	-0.001	0.000	0.906	-0.001	-0.001	0.836
	std	0.090	0.057	0.057	0.087	0.055	0.055	0.091	0.055	0.058
		0.057	0.088	0.057	0.055	0.084	0.055	0.055	0.090	0.056
		0.057	0.057	0.084	0.055	0.055	0.082	0.058	0.056	0.091
10	av	0.990	0.000	0.001	0.688	0.000	0.001	0.481	-0.001	-0.002
		0.000	0.991	0.000	0.000	0.688	0.000	-0.001	0.481	0.000
		0.001	0.000	0.990	0.001	0.000	0.688	-0.002	0.000	0.480
	std	0.057	0.032	0.032	0.050	0.026	0.026	0.059	0.029	0.031
		0.032	0.057	0.031	0.026	0.050	0.025	0.029	0.060	0.030
		0.032	0.031	0.058	0.026	0.025	0.051	0.031	0.030	0.060
15	av	0.990	0.000	0.001	0.537	0.000	0.001	0.292	0.000	0.000
		0.000	0.989	0.000	0.000	0.536	0.000	0.000	0.295	-0.002
		0.001	0.000	0.989	0.001	0.000	0.535	0.000	-0.002	0.295
	std	0.052	0.025	0.027	0.041	0.019	0.019	0.049	0.022	0.023
		0.025	0.053	0.025	0.019	0.043	0.019	0.022	0.049	0.023
		0.027	0.025	0.054	0.019	0.019	0.042	0.023	0.023	0.049

Z Tabel 4.7–4.9 można zauważyć, że dla  $(p, q) = (3, 5)$  estymatory MLE i ELE macierzy  $\Sigma$  są nieznacznie obciążone, podczas gdy obciążenie QLE jest większe. Wraz ze wzrostem  $p$  zarówno ELE jak i QLE są znacznie bardziej obciążone niż MLE. Wielkości normy Frobeniusa odchyłeń standardowych są porównywalne dla wszystkich estymatorów.

Uśrednione wartości estymatorów i ich odchylenia standardowe (lub normy Frobeniusa tych wartości) dla danych generowanych z rozkładu  $N_{100,pq}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{AR} \otimes \mathbf{I}_q)$  dla różnych wartości parametrów  $p$  i  $q$  oraz  $\varrho \in \{-0.9, -0.5, 0, 0.5, 0.9\}$  zaprezentowane zostały w Tabelach 4.10–4.13.

Z Tabel 4.10–4.11 możemy zauważyć, że dla wszystkich zestawów parametrów obciążenie każdego z estymatorów jest niewielkie i nie zależy od parametrów  $p$  i  $q$ . Odchylenia standardowe MLE i ELE są porównywalne, podczas gdy dla estymatorów QLE są nieznacznie większe.

Tabela 4.10. Uśrednione wartości (av) i odchylenia standardowe (std) estymatorów  $\varrho$  dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{AR} \otimes \mathbf{I}_3)$ .

$p$	$\varrho$	MLE		ELE		QLE	
		av( $\hat{\varrho}$ )	std( $\hat{\varrho}$ )	av( $\check{\varrho}$ )	std( $\check{\varrho}$ )	av( $\tilde{\varrho}$ )	std( $\tilde{\varrho}$ )
3	-0.9	-0.8995	0.0091	-0.8996	0.0093	-0.9000	0.0103
3	-0.5	-0.4989	0.0316	-0.4992	0.0323	-0.4993	0.0366
3	0	0.0008	0.0402	0.0012	0.0412	-0.0020	0.0474
3	0.5	0.5001	0.0318	0.5008	0.0328	0.5004	0.0368
3	0.9	0.8998	0.0091	0.9001	0.0094	0.8998	0.0103
10	-0.9	-0.8995	0.0063	-0.8997	0.0075	-0.9001	0.0123
10	-0.5	-0.4995	0.0163	-0.5000	0.0198	-0.5025	0.0319
10	0	-0.0002	0.0191	-0.0002	0.0236	0.0010	0.0385
10	0.5	0.4994	0.0155	0.4998	0.0192	0.4998	0.0324
10	0.9	0.8996	0.0060	0.8999	0.0073	0.9002	0.0125
15	-0.9	-0.8995	0.0055	-0.9000	0.0072	-0.9009	0.0165
15	-0.5	-0.4994	0.0129	-0.5000	0.0180	-0.5013	0.0403
15	0	-0.0001	0.0150	-0.0003	0.0215	0.0004	0.0472
15	0.5	0.4996	0.0128	0.4996	0.0179	0.4997	0.0389
15	0.9	0.8997	0.0055	0.8999	0.0075	0.9003	0.0169

Tabela 4.11. Uśrednione wartości (av) i odchylenia standardowe (std) estymatorów  $\varrho$  dla danych generowanych z rozkładu  $N_{100,5p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{AR} \otimes \mathbf{I}_5)$ .

$p$	$\varrho$	MLE		ELE		QLE	
		av( $\hat{\varrho}$ )	std( $\hat{\varrho}$ )	av( $\check{\varrho}$ )	std( $\check{\varrho}$ )	av( $\tilde{\varrho}$ )	std( $\tilde{\varrho}$ )
3	-0.9	-0.8998	0.0072	-0.9000	0.0074	-0.9000	0.0088
3	-0.5	-0.4999	0.0253	-0.5005	0.0260	-0.5007	0.0299
3	0	-0.0005	0.0326	-0.0007	0.0335	-0.0002	0.0390
3	0.5	0.4991	0.0257	0.4995	0.0266	0.5006	0.0296
3	0.9	0.8996	0.0073	0.8998	0.0076	0.9000	0.0088
10	-0.9	-0.8997	0.0048	-0.8997	0.0068	-0.8993	0.0167
10	-0.5	-0.4998	0.0127	-0.4997	0.0177	-0.5005	0.0440
10	0	-0.0002	0.0154	0.0001	0.0213	0.0002	0.0515
10	0.5	0.4999	0.0130	0.5001	0.0180	0.4992	0.0422
10	0.9	0.9001	0.0049	0.9002	0.0068	0.8984	0.0174
15	-0.9	-0.8998	0.0042	-0.9003	0.0084	-0.8855	0.0770
15	-0.5	-0.4998	0.0103	-0.5001	0.0201	-0.4815	0.1209
15	0	0.0000	0.0123	0.0005	0.0236	0.0004	0.1345
15	0.5	0.5000	0.0103	0.5006	0.0198	0.4966	0.1192
15	0.9	0.9000	0.0042	0.9004	0.0084	0.8822	0.0861

Tabela 4.12. Normy Frobeniusa z uśrednionych wartości (av) i odchyłeń standardowych (std) estymatorów  $\Sigma$  dla danych generowanych z rozkładu  $N_{100,pq}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{AR} \otimes \mathbf{I}_q)$ .

		MLE		ELE		QLE	
$p$	$\varrho$	$\ \text{av}(\widehat{\Sigma})\ _F^2$	$\ \text{std}(\widehat{\Sigma})\ _F^2$	$\ \text{av}(\widetilde{\Sigma})\ _F^2$	$\ \text{std}(\widetilde{\Sigma})\ _F^2$	$\ \text{av}(\widetilde{\Sigma})\ _F^2$	$\ \text{std}(\widetilde{\Sigma})\ _F^2$
$q = 3$							
3	-0.9	2.943	0.048	2.467	0.045	2.090	0.047
3	-0.5	2.941	0.041	2.462	0.038	2.072	0.041
3	0	2.942	0.039	2.462	0.036	2.077	0.039
3	0.5	2.948	0.041	2.470	0.038	2.093	0.043
3	0.9	2.954	0.049	2.481	0.045	2.098	0.052
10	-0.9	2.934	0.020	1.417	0.015	0.693	0.019
10	-0.5	2.943	0.013	1.413	0.009	0.694	0.013
10	0	2.945	0.012	1.412	0.009	0.685	0.012
10	0.5	2.943	0.013	1.412	0.009	0.682	0.013
10	0.9	2.938	0.019	1.420	0.014	0.700	0.019
15	-0.9	2.932	0.014	0.863	0.009	0.266	0.012
15	-0.5	2.941	0.008	0.857	0.005	0.253	0.008
15	0	2.943	0.008	0.856	0.005	0.252	0.006
15	0.5	2.942	0.008	0.856	0.005	0.256	0.007
15	0.9	2.940	0.014	0.862	0.009	0.265	0.013
$q = 5$							
3	-0.9	4.914	0.111	3.654	0.097	2.741	0.103
3	-0.5	4.910	0.102	3.644	0.090	2.742	0.097
3	0	4.906	0.100	3.637	0.087	2.744	0.097
3	0.5	4.904	0.103	3.637	0.089	2.756	0.102
3	0.9	4.902	0.111	3.643	0.096	2.754	0.104
10	-0.9	4.892	0.037	1.187	0.019	0.314	0.027
10	-0.5	4.901	0.030	1.184	0.016	0.308	0.023
10	0	4.901	0.029	1.184	0.015	0.306	0.021
10	0.5	4.902	0.030	1.186	0.016	0.308	0.021
10	0.9	4.923	0.038	1.196	0.019	0.310	0.027
15	-0.9	4.898	0.026	0.282	0.007	0.022	0.011
15	-0.5	4.904	0.021	0.278	0.006	0.020	0.007
15	0	4.904	0.020	0.277	0.006	0.020	0.006
15	0.5	4.906	0.020	0.278	0.006	0.021	0.007
15	0.9	4.921	0.027	0.282	0.007	0.022	0.011

Tabela 4.13. Uśrednione wartości (av) oraz odchylenia standardowe (std) estymatorów macierzy  $\Sigma$  dla danych generowanych z rozkładu  $N_{100,3p}(\mathbf{0}, \mathbf{I}_{100}, \Psi_{AR} \otimes \mathbf{I}_3)$  gdy  $\varrho = 0.5$ .

$p$		MLE			ELE			QLE		
		$\hat{\Sigma}$			$\check{\Sigma}$			$\tilde{\Sigma}$		
3	av	0.992	-0.001	0.000	0.908	-0.002	-0.001	0.833	-0.002	0.001
		-0.001	0.993	-0.001	-0.002	0.908	0.000	-0.002	0.839	-0.001
		0.000	-0.001	0.989	-0.001	0.000	0.905	0.001	-0.001	0.834
	std	0.088	0.057	0.057	0.085	0.055	0.054	0.089	0.058	0.055
		0.057	0.086	0.057	0.055	0.083	0.054	0.058	0.089	0.057
		0.057	0.057	0.082	0.054	0.054	0.080	0.055	0.057	0.087
10	av	0.990	0.000	0.001	0.686	0.000	0.001	0.475	-0.001	0.000
		0.000	0.991	0.000	0.000	0.687	0.000	-0.001	0.477	-0.003
		0.001	0.000	0.990	0.001	0.000	0.686	0.000	-0.003	0.478
	std	0.047	0.032	0.032	0.042	0.026	0.026	0.050	0.030	0.030
		0.032	0.048	0.031	0.026	0.042	0.025	0.030	0.050	0.030
		0.032	0.031	0.047	0.026	0.025	0.042	0.030	0.030	0.049
15	av	0.991	0.000	0.001	0.535	0.000	0.001	0.292	0.001	0.000
		0.000	0.990	0.000	0.000	0.534	0.000	0.001	0.293	0.000
		0.001	0.000	0.990	0.001	0.000	0.533	0.000	0.000	0.292
	std	0.038	0.025	0.027	0.031	0.018	0.019	0.037	0.022	0.022
		0.025	0.040	0.025	0.018	0.033	0.019	0.022	0.037	0.022
		0.027	0.025	0.039	0.019	0.019	0.032	0.022	0.022	0.037

Z Tabel 4.12–4.13 można zaobserwować, że dla  $(p, q) = (3, 5)$  estymatory MLE i ELE macierzy  $\Sigma$  są nieznacznie obciążone, podczas gdy obciążenie QLE jest większe. Wraz ze wzrostem  $p$  zarówno ELE jak i QLE są znacznie bardziej obciążone niż MLE. Wielkości normy Frobeniusa odchyłeń standardowych są porównywalne dla wszystkich estymatorów.

Z przeprowadzonych badań symulacyjnych można wywnioskować, że estymatory największej wiarygodności są zazwyczaj lepsze (w sensie obciążenia) niż estymatory wyznaczone w wyniku minimalizacji entropijnej funkcji straty czy kwadratowej funkcji straty. Zatem, aby używać ELE i QLE warto byłoby poprawić ich własności, na przykład poprzez modyfikację macierzy  $\mathbf{S}$  (por. Stein [39]) lub uwzględnienie funkcji kary w procesie minimalizacji.

#### 4.4. Testowanie hipotez

W tym rozdziale użyjemy indeksów  $\zeta_{\text{UN}}^{(k)}$ ,  $\zeta_{\text{CS}}^{(k)}$  i  $\zeta_{\text{AR}}^{(k)}$ ,  $k \in \{E, Q\}$ , w kontekście badania mocy testów, do pomiaru rozbieżności między zbiorami struktur kowariancyjnych rozważanych w procesie testowania hipotez za pomocą LRT i RST. Podobnie jak w pracach Lu i Zimmerman [22, 23] oraz Filipiak i in. [12, 13], rozważane są następujące hipotezy:

$$H_{01} : \mathbf{\Omega} \in \mathcal{S}, \quad H_{02} : \mathbf{\Omega} = \mathbf{\Psi}_{\text{CS}} \otimes \mathbf{\Sigma}, \quad H_{03} : \mathbf{\Omega} = \mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}$$

przeciwko  $H_A : \mathbf{\Omega}$  bez struktury (określona dodatnio).

Przypomnijmy, że moc testu to prawdopodobieństwo odrzucenia hipotezy zerowej, gdy prawdziwa jest hipoteza alternatywna, to jest

$$P(\text{odrzućenie } H_0 \mid H_1 \text{ jest prawdziwa}) = P_{\Omega}(T > \kappa_{\alpha}),$$

gdzie  $T$  to statystyka LRT lub RST i  $\kappa_{\alpha}$  jest  $(1 - \alpha)100\%$  kwantylem rozkładu statystyki testowej przy prawdziwości hipotezy zerowej (ang. *empirical null distribution*); por. Filipiak i in. [12, 13]. We wszystkich symulacjach przyjęto poziom istotności  $\alpha = 0.05$ .

Ponieważ rozkład dokładny statystyk LRT i RST nie jest znany, natomiast rozkład asymptotyczny  $\chi_{\nu}^2$ , gdzie  $\nu$  jest różnicą między liczbą nieznanymi parametrów w hipotezie alternatywnej i hipotezie zerowej (por. Rao [30]), jest słabym przybliżeniem zwłaszcza dla małych wielkości prób (por. Filipiak i in. [12, 13]), kwantyl  $\kappa_{\alpha}$  zostanie zastąpiony przez kwantyl odpowiedniego rozkładu empirycznego wyznaczonego dla hipotez  $H_{02}$  i  $H_{03}$  w pracach Filipiak i in. [12, 13] na podstawie 50000 prób dla różnych wielkości  $p, q, n$ . Moc empiryczna obliczana jest jako proporcja liczby odrzuconych hipotez zerowych do liczby wszystkich wykonanych testów (50000), gdy  $H_A$  jest prawdziwa. Należy zatem dobrać macierz  $\mathbf{\Omega}$  bez struktury z uwzględnieniem jej rozbieżności (w sensie entropijnej lub kwadratowej funkcji straty) od zbioru  $\mathcal{S}$ ,  $\mathcal{S}_{\text{CS}}$  lub  $\mathcal{S}_{\text{AR}}$ .

Przypomnijmy, że zarówno entropijna jak i kwadratowa funkcja straty są nieograniczone z góry i są nieporównywalne. Dlatego użyte zostaną skorygowane rozbieżności  $\zeta^{(k)}$ ,  $k \in \{E, Q\}$ , o wartościach z przedziału  $[0; 1)$ . W szczególności  $\xi^{(k)} = 1 - 1/[\zeta^{(k)} + 1]$ , gdzie  $\zeta^{(k)}$  jest odpowiednio równa  $\zeta_{\text{UN}}^{(k)}$ ,  $\zeta_{\text{CS}}^{(k)}$  lub  $\zeta_{\text{AR}}^{(k)}$ . W rozdziale tym macierze  $\mathbf{\Omega}$  zostały wygenerowane losowo tak, aby odpowiednie rozbieżności  $\xi^{(k)}$  były równe około 0.1 i 0.5. Wartości te będziemy dalej nazywać poziomami

rozbieżności, a macierz  $\mathbf{\Omega}$  oznaczać będziemy  $\mathbf{\Omega}_g$ . Ponadto sprawdzimy, czy dla dwóch hipotez alternatywnych o takim samym poziomie rozbieżności moce testu są zbliżone. W tym celu dobrana została druga macierz, nieposiadająca struktury,  $\mathbf{\Omega}_t$ , konstruowana na podstawie macierzy o strukturze zadanej hipotezą zerową w taki sposób, że zmodyfikowane są co najwyżej dwa elementy, aby uzyskać zadany poziom rozbieżności (0.1 lub 0.5).

W niniejszym rozdziale sprawdzimy również jak w opisanej sytuacji zachowuje się najbardziej intuicyjna miara rozbieżności między macierzami, czyli skorygowana do przedziału  $[0, 1)$  norma Frobeniusa:

$$\xi^{(F)} = \zeta^{(F)} / \|\mathbf{S}\|_F,$$

gdzie

$$\zeta^{(F)} = \min_{\mathbf{\Gamma}} f_F(\mathbf{S}, \mathbf{\Gamma}), \quad \mathbf{\Gamma} \in \mathcal{S}, \mathbf{\Gamma} \in \mathcal{S}_{CS} \text{ lub } \mathbf{\Gamma} \in \mathcal{S}_{AR};$$

por. Filipiak i Klein [7].

Algorytm postępowania wygląda następująco:

1. generowanie danych z rozkładu  $N_{n,qp}(\mathbf{0}, \mathbf{I}_n, \mathbf{\Omega})$ , gdzie  $\mathbf{\Omega} = \mathbf{\Omega}_g$  lub  $\mathbf{\Omega} = \mathbf{\Omega}_t$  jest macierzą nieposiadającą struktury o zadanym poziomie rozbieżności;
2. wyznaczenie  $\mathbf{S}$ ;
3. przeprowadzenie testów LRT i RST;
4. powtórzenie 50000 razy kroków 1–3 dla każdego ze zbiorów;
5. wyznaczenie mocy każdego testu jako proporcji

$$\frac{\text{liczba odrzuconych hipotez zerowych}}{50000}.$$

Symulacje przeprowadzone zostały dla różnych rozmiarów prób:

$$n \in \{10, 15, 25, 50, 100, 150, 200\}.$$

Aby przeprowadzić test LRT spełniony musi być warunek  $n > pq$ , natomiast dla testu RST wystarczy warunek  $n > \max\{p, q\}$  (por. Filipiak i in. [13]). W celu porównania mocy testów wybrane zostały parametry gwarantujące możliwość przeprowadzenia obu testów, czyli pary  $(p, q)$  równe  $(3, 2)$ ,  $(3, 3)$ ,  $(5, 2)$  lub  $(5, 5)$ .

Wybrane moce empiryczne testów LRT i RST w zależności od poziomu rozbieżności bazującej na entropijnej funkcji straty zaprezentowane zostały w Tabelach 4.14–4.15, na kwadratowej funkcji straty w Tabelach 4.16–4.17 oraz dla parametrów  $(p, q) = (3, 2)$  na Rysunku 4.1. Moce testów oparte na normie Frobeniusa przedstawiają Tabele 4.18–4.19.

Tabela 4.14. Moce empiryczne testu LRT dla testowania struktury kowariancyjnej ze względu na skorygowane rozbieżności entropijnej funkcji straty pomiędzy  $\Omega$  bez struktury i  $H_{0i}$ ,  $i = 1, 2, 3$ , na podstawie 50000 symulacji.

$\xi^{(E)}$		0.1						0.5					
$(p, q)$	$n$	$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$		$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$	
		$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$
(3, 2)	10	0.063	0.064	0.060	0.064	0.061	0.063	0.239	0.244	0.232	0.277	0.261	0.288
	25	0.117	0.119	0.109	0.120	0.108	0.118	0.885	0.905	0.866	0.919	0.906	0.927
	50	0.231	0.239	0.207	0.242	0.206	0.241	0.999	1.000	0.999	1.000	1.000	1.000
	100	0.515	0.530	0.462	0.532	0.461	0.530	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.744	0.758	0.690	0.757	0.694	0.757	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.884	0.897	0.846	0.897	0.846	0.894	1.000	1.000	1.000	1.000	1.000	1.000
(3, 3)	10	0.053	0.053	0.053	0.053	0.053	0.053	0.082	0.091	0.082	0.093	0.082	0.096
	25	0.084	0.088	0.083	0.088	0.083	0.088	0.643	0.743	0.629	0.757	0.626	0.775
	50	0.146	0.157	0.144	0.161	0.145	0.161	0.985	0.995	0.983	0.996	0.982	0.997
	100	0.323	0.350	0.319	0.360	0.319	0.363	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.510	0.556	0.509	0.569	0.511	0.572	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.690	0.734	0.687	0.745	0.688	0.747	1.000	1.000	1.000	1.000	1.000	1.000
(5, 2)	15	0.061	0.061	0.060	0.063	0.060	0.062	0.196	0.218	0.200	0.352	0.201	0.367
	25	0.082	0.081	0.080	0.087	0.081	0.088	0.530	0.609	0.526	0.793	0.531	0.808
	50	0.141	0.140	0.132	0.156	0.133	0.157	0.962	0.986	0.959	0.997	0.960	0.998
	100	0.304	0.307	0.270	0.342	0.278	0.345	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.493	0.498	0.443	0.538	0.451	0.539	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.668	0.682	0.612	0.716	0.623	0.717	1.000	1.000	1.000	1.000	1.000	1.000
(5, 5)	50	0.071	0.074	0.071	0.077	0.071	0.077	0.453	0.667	0.454	0.734	0.454	0.743
	100	0.109	0.122	0.109	0.127	0.109	0.127	0.951	0.996	0.950	0.999	0.950	0.999
	150	0.161	0.184	0.165	0.196	0.166	0.197	0.999	1.000	0.999	1.000	0.999	1.000
	200	0.215	0.255	0.217	0.275	0.217	0.273	1.000	1.000	1.000	1.000	1.000	1.000

Tabela 4.15. Moce empiryczne testu RST dla testowania struktury kowariancyjnej ze względu na skorygowane rozbieżności entropijnej funkcji straty pomiędzy  $\Omega$  bez struktury i  $H_{0i}$ ,  $i = 1, 2, 3$ , na podstawie 50000 symulacji.

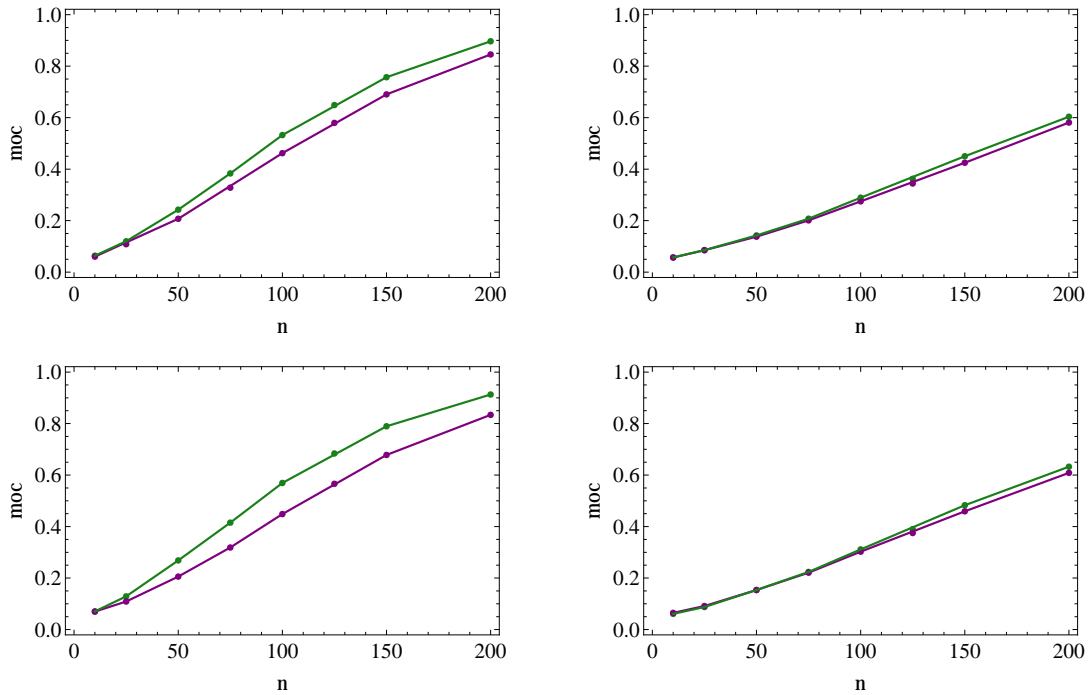
$\xi^{(E)}$		0.1						0.5					
$(p, q)$	$n$	$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$		$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$	
		$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$
(3, 2)	10	0.067	0.067	0.070	0.070	0.070	0.070	0.274	0.230	0.320	0.398	0.367	0.426
	25	0.116	0.119	0.109	0.129	0.110	0.131	0.850	0.873	0.842	0.937	0.897	0.952
	50	0.222	0.236	0.206	0.269	0.203	0.271	0.999	0.999	0.998	1.000	0.999	1.000
	100	0.495	0.521	0.449	0.569	0.444	0.572	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.729	0.757	0.678	0.790	0.677	0.792	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.873	0.894	0.834	0.913	0.834	0.913	1.000	1.000	1.000	1.000	1.000	1.000
(3, 3)	10	0.062	0.064	0.064	0.064	0.063	0.064	0.197	0.252	0.220	0.252	0.214	0.266
	25	0.095	0.099	0.093	0.097	0.091	0.098	0.654	0.784	0.653	0.822	0.642	0.850
	50	0.156	0.173	0.159	0.185	0.159	0.188	0.980	0.997	0.978	0.998	0.975	0.999
	100	0.332	0.374	0.331	0.400	0.332	0.403	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.517	0.576	0.520	0.617	0.523	0.622	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.694	0.753	0.693	0.784	0.694	0.787	1.000	1.000	1.000	1.000	1.000	1.000
(5, 2)	15	0.069	0.065	0.073	0.078	0.072	0.080	0.249	0.232	0.290	0.643	0.281	0.678
	25	0.090	0.083	0.089	0.109	0.090	0.113	0.528	0.523	0.548	0.918	0.547	0.931
	50	0.156	0.141	0.150	0.205	0.150	0.209	0.953	0.973	0.949	0.999	0.951	1.000
	100	0.323	0.298	0.290	0.424	0.295	0.431	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.511	0.483	0.466	0.632	0.472	0.638	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.684	0.671	0.634	0.797	0.644	0.802	1.000	1.000	1.000	1.000	1.000	1.000
(5, 5)	50	0.080	0.090	0.079	0.096	0.080	0.096	0.569	0.881	0.579	0.946	0.578	0.951
	100	0.116	0.144	0.119	0.165	0.119	0.166	0.962	1.000	0.965	1.000	0.965	1.000
	150	0.171	0.226	0.175	0.262	0.176	0.264	0.999	1.000	0.999	1.000	0.999	1.000
	200	0.229	0.313	0.234	0.365	0.235	0.367	1.000	1.000	1.000	1.000	1.000	1.000

Tabela 4.16. Moce empiryczne testu LRT dla testowania struktury kowariancyjnej ze względu na skorygowane rozbieżności kwadratowej funkcji straty pomiędzy  $\Omega$  bez struktury i  $H_{0i}$ ,  $i = 1, 2, 3$ , na podstawie 50000 symulacji.

$\xi^{(Q)}$		0.1						0.5					
$(p, q)$	$n$	$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$		$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$	
		$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$
(3, 2)	10	0.058	0.057	0.057	0.057	0.057	0.057	0.119	0.192	0.138	0.230	0.120	0.249
	25	0.086	0.082	0.085	0.085	0.084	0.086	0.493	0.806	0.590	0.859	0.509	0.882
	50	0.140	0.132	0.138	0.142	0.134	0.142	0.915	0.996	0.960	0.998	0.924	0.999
	100	0.285	0.269	0.275	0.289	0.268	0.294	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.436	0.416	0.425	0.450	0.421	0.461	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.590	0.564	0.581	0.604	0.569	0.611	1.000	1.000	1.000	1.000	1.000	1.000
(3, 3)	10	0.052	0.051	0.052	0.051	0.052	0.052	0.080	0.081	0.081	0.087	0.082	0.089
	25	0.067	0.068	0.066	0.069	0.066	0.069	0.585	0.612	0.588	0.675	0.589	0.698
	50	0.095	0.099	0.093	0.104	0.092	0.104	0.960	0.977	0.962	0.988	0.964	0.991
	100	0.170	0.180	0.166	0.193	0.165	0.193	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.259	0.277	0.253	0.299	0.255	0.303	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.367	0.390	0.357	0.421	0.359	0.424	1.000	1.000	1.000	1.000	1.000	1.000
(5, 2)	15	0.056	0.056	0.055	0.057	0.055	0.057	0.180	0.174	0.195	0.328	0.190	0.349
	25	0.065	0.066	0.066	0.070	0.066	0.070	0.468	0.481	0.489	0.761	0.476	0.787
	50	0.093	0.092	0.089	0.103	0.088	0.103	0.921	0.943	0.927	0.995	0.920	0.997
	100	0.163	0.158	0.148	0.190	0.149	0.194	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.249	0.240	0.225	0.294	0.224	0.297	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.354	0.343	0.315	0.415	0.314	0.417	1.000	1.000	1.000	1.000	1.000	1.000
(5, 5)	50	0.061	0.063	0.061	0.064	0.061	0.064	0.276	0.585	0.280	0.697	0.279	0.712
	100	0.080	0.085	0.080	0.088	0.079	0.089	0.741	0.988	0.745	0.997	0.745	0.998
	150	0.103	0.112	0.104	0.121	0.104	0.121	0.959	1.000	0.961	1.000	0.961	1.000
	200	0.126	0.141	0.126	0.153	0.126	0.154	1.000	1.000	1.000	1.000	1.000	1.000

Tabela 4.17. Moce empiryczne testu RST dla testowania struktury kowariancyjnej ze względu na skorygowane rozbieżności kwadratowej funkcji straty pomiędzy  $\Omega$  bez struktury i  $H_{0i}$ ,  $i = 1, 2, 3$ , na podstawie 50000 symulacji.

$\xi^{(Q)}$		0.1						0.5					
$(p, q)$	$n$	$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$		$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$	
		$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$
(3, 2)	10	0.061	0.059	0.065	0.061	0.064	0.060	0.119	0.193	0.167	0.334	0.150	0.371
	25	0.092	0.082	0.091	0.088	0.090	0.089	0.371	0.763	0.525	0.884	0.455	0.916
	50	0.148	0.130	0.154	0.153	0.150	0.156	0.825	0.995	0.933	0.999	0.884	0.999
	100	0.298	0.264	0.303	0.311	0.296	0.320	0.999	1.000	1.000	1.000	0.999	1.000
	150	0.455	0.414	0.459	0.483	0.452	0.496	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.605	0.561	0.608	0.632	0.597	0.646	1.000	1.000	1.000	1.000	1.000	1.000
(3, 3)	10	0.055	0.057	0.057	0.058	0.056	0.057	0.256	0.206	0.263	0.223	0.262	0.235
	25	0.072	0.075	0.071	0.075	0.070	0.074	0.712	0.666	0.721	0.757	0.722	0.787
	50	0.101	0.106	0.102	0.115	0.102	0.117	0.978	0.982	0.979	0.994	0.980	0.996
	100	0.181	0.191	0.175	0.213	0.174	0.213	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.271	0.288	0.267	0.329	0.267	0.335	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.381	0.406	0.373	0.457	0.375	0.464	1.000	1.000	1.000	1.000	1.000	1.000
(5, 2)	15	0.059	0.057	0.061	0.065	0.061	0.066	0.295	0.195	0.364	0.611	0.345	0.655
	25	0.071	0.066	0.071	0.080	0.071	0.082	0.580	0.424	0.634	0.898	0.617	0.919
	50	0.099	0.093	0.097	0.125	0.097	0.126	0.949	0.915	0.961	0.999	0.955	0.999
	100	0.171	0.157	0.158	0.233	0.160	0.238	1.000	1.000	1.000	1.000	1.000	1.000
	150	0.263	0.234	0.238	0.353	0.240	0.359	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.370	0.337	0.335	0.490	0.335	0.496	1.000	1.000	1.000	1.000	1.000	1.000
(5, 5)	50	0.067	0.070	0.068	0.073	0.068	0.073	0.419	0.818	0.428	0.928	0.429	0.937
	100	0.086	0.094	0.088	0.104	0.087	0.104	0.844	0.998	0.856	1.000	0.856	1.000
	150	0.114	0.126	0.115	0.144	0.115	0.146	0.983	1.000	0.984	1.000	0.984	1.000
	200	0.143	0.164	0.144	0.191	0.145	0.192	1.000	1.000	1.000	1.000	1.000	1.000



Rysunek 4.1. Moce empiryczne dla  $H_{02}$  i  $\xi^{(k)} = 0.1$ ,  $k \in \{E, Q\}$ , dla entropijnej funkcji straty (lewy panel) oraz kwadratowej funkcji straty (prawy panel), odpowiednio dla testu LRT (pierwszy wiersz) oraz RST (drugi wiersz) dla  $(p, q) = (3, 2)$ . Fioletowy kolor reprezentuje  $\Omega_g$ , a zielony  $\Omega_t$ .

Z Tabel 4.14–4.17 oraz Rysunku 4.1 można zaobserwować, że zgodnie z oczekiwaniami, moc testu LRT i RST zarówno dla entropijnej jak i kwadratowej funkcji straty wzrasta wraz z rozmiarem próby. Rośnie również wraz ze wzrostem rozbieżności między macierzą  $\Omega$  a zbiorami  $\mathcal{S}$ ,  $\mathcal{S}_{CS}$  lub  $\mathcal{S}_{AR}$ . Co więcej, w większości przypadków wartości mocy są porównywalne dla  $\Omega_g$  i  $\Omega_t$  o takich samych rozbieżnościach. Może się jednak zdarzyć, że moc testu dla macierzy  $\Omega_g$  różni się istotnie od mocy testu dla macierzy  $\Omega_t$ . Przykładowo dla poziomej rozbieżności 0.5, dla  $(p, q) = (5, 5)$  i  $n = 50$ , zarówno entropijna jak i kwadratowa funkcja straty wskazują znacząco różne moce dla obu testów. Sytuacja ta nie występuje dla rozbieżności  $\xi^{(k)} = 0.1$ ,  $k \in \{E, Q\}$ .

Tabela 4.18. Moce empiryczne testu LRT dla testowania struktury kowariancyjnej ze względu na skorygowane rozbieżności normy Frobeniusa pomiędzy  $\Omega$  bez struktury i  $H_{0i}$ ,  $i = 1, 2, 3$ , na podstawie 50000 symulacji.

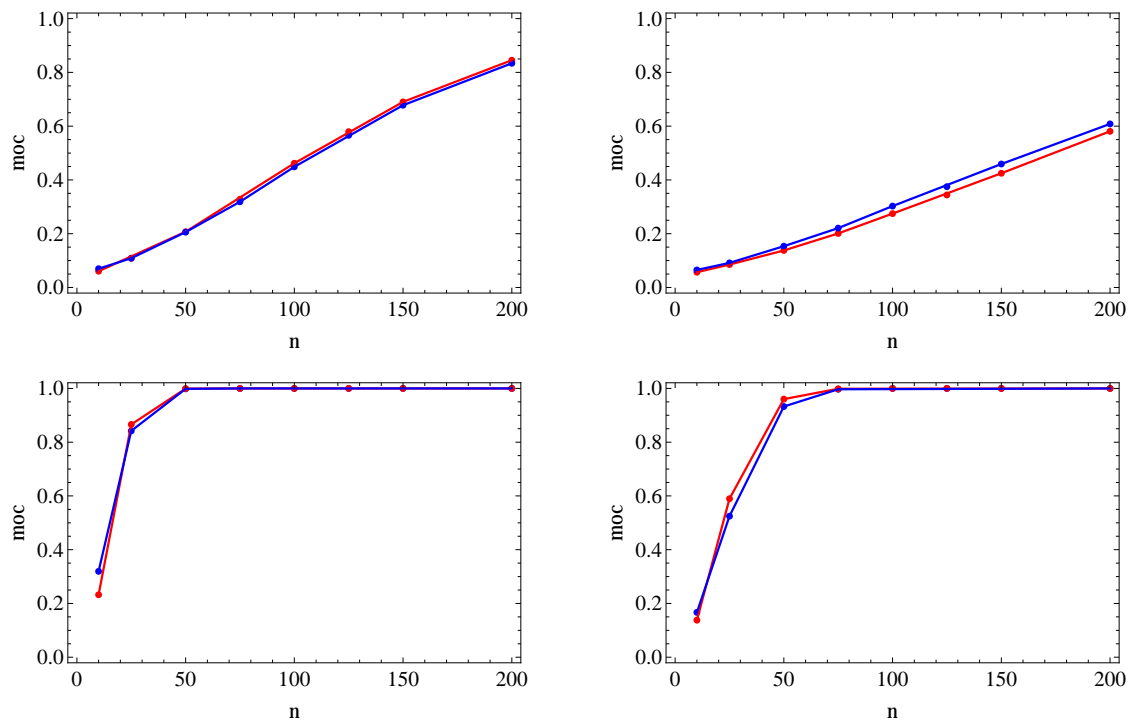
$\xi^{(F)}$		0.1						0.5					
$(p, q)$	$n$	$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$		$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$	
		$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$
(3, 2)	10	0.449	0.058	0.410	0.057	0.545	0.055	0.887	0.214	0.983	0.337	0.981	0.288
	25	0.995	0.085	0.992	0.082	0.998	0.076	1.000	0.854	1.000	0.963	1.000	0.927
	50	1.000	0.138	1.000	0.132	1.000	0.116	1.000	0.999	1.000	1.000	1.000	1.000
	100	1.000	0.286	1.000	0.263	1.000	0.221	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.441	1.000	0.408	1.000	0.343	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.594	1.000	0.551	1.000	0.466	1.000	1.000	1.000	1.000	1.000	1.000
(3, 3)	10	0.278	0.052	0.275	0.052	0.281	0.052	0.136	0.143	0.140	0.116	0.138	0.110
	25	1.000	0.076	1.000	0.074	1.000	0.070	0.980	0.979	0.983	0.915	0.980	0.880
	50	1.000	0.121	1.000	0.119	1.000	0.108	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.244	1.000	0.239	1.000	0.204	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.388	1.000	0.381	1.000	0.323	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.541	1.000	0.530	1.000	0.451	1.000	1.000	1.000	1.000	1.000	1.000
(5, 2)	15	0.636	0.062	0.639	0.065	0.950	0.058	0.477	0.477	0.524	0.693	0.535	0.531
	25	0.992	0.085	0.990	0.095	1.000	0.075	0.941	0.953	0.959	0.985	0.962	0.934
	50	1.000	0.155	1.000	0.181	1.000	0.116	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.349	1.000	0.410	1.000	0.231	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.560	1.000	0.636	1.000	0.361	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.747	1.000	0.809	1.000	0.504	1.000	1.000	1.000	1.000	1.000	1.000
(5, 5)	50	1.000	0.116	1.000	0.127	1.000	0.091	1.000	0.999	1.000	0.999	1.000	0.984
	100	1.000	0.265	1.000	0.299	1.000	0.174	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.463	1.000	0.526	1.000	0.292	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.652	1.000	0.720	1.000	0.418	1.000	1.000	1.000	1.000	1.000	1.000

Zauważmy, że moc testu mierzona na podstawie normy Frobeniusa nie zachowuje monotoniczności (por. Tabele 4.18 i 4.19). Przykładowo dla  $(p, q) = (3, 3)$ ,  $n = 10$  i  $\Omega_g$  wraz ze wzrostem rozbieżności spada zarówno moc testu LRT jak i RST dla każdej z testowanych hipotez. Podobna sytuacja występuje dla macierzy  $\Omega_g$  dla  $(p, q) = (3, 3)$ ,  $n = 25$  dla obu testów lub  $(p, q) = (5, 2)$ ,  $n = 15$  lub  $n = 25$  dla testu LRT dla każdej z testowanych hipotez, jak również dla  $(p, q) = (5, 2)$ ,  $n = 15$  lub  $n = 25$  dla testu RST hipotezy  $H_{03}$ . Ponadto, moce testów dla dwóch macierzy o takiej samej rozbieżności  $\xi^{(F)}$  są znacząco różne (zobacz na przykład  $(p, q) = (3, 3)$ ,  $n = 25$  i rozbieżność rzędu 0.1). Zatem norma Frobeniusa nie jest dobrą miarą rozbieżności w kontekście badania mocy testu.

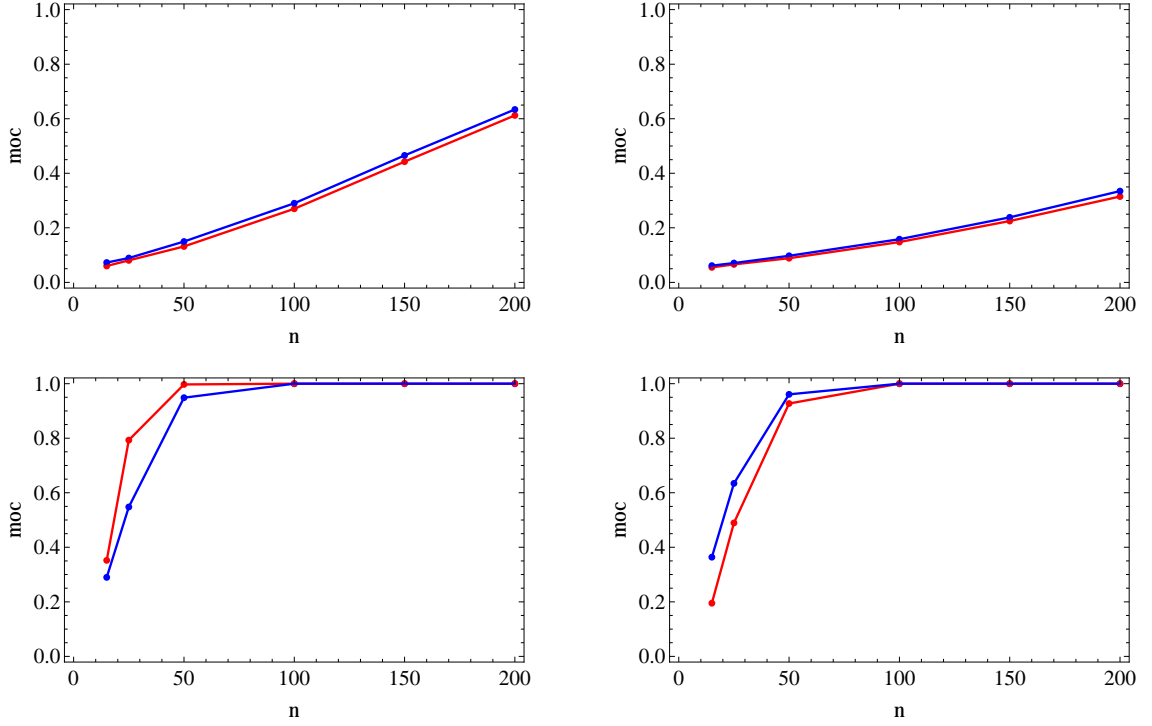
Tabela 4.19. Moce empiryczne testu RST dla testowania struktury kowariancyjnej ze względu na skorygowane rozbieżności normy Frobeniusa pomiędzy  $\Omega$  bez struktury i  $H_{0i}$ ,  $i = 1, 2, 3$ , na podstawie 50000 symulacji.

$\xi^{(F)}$		0.1						0.5					
$(p, q)$	$n$	$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$		$\mathcal{S}$		$\mathcal{S}_{CS}$		$\mathcal{S}_{AR}$	
		$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$	$\Omega_g$	$\Omega_t$
(3, 2)	10	0.396	0.059	0.426	0.060	0.638	0.057	0.785	0.209	0.903	0.469	0.897	0.426
	25	0.991	0.084	0.983	0.083	0.999	0.079	1.000	0.814	1.000	0.972	1.000	0.952
	50	1.000	0.136	1.000	0.141	1.000	0.124	1.000	0.998	1.000	1.000	1.000	1.000
	100	1.000	0.281	1.000	0.282	1.000	0.242	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.440	1.000	0.439	1.000	0.372	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.591	1.000	0.580	1.000	0.497	1.000	1.000	1.000	1.000	1.000	1.000
(3, 3)	10	0.712	0.060	0.710	0.060	0.732	0.058	0.348	0.461	0.358	0.355	0.348	0.327
	25	1.000	0.084	1.000	0.081	1.000	0.076	0.915	0.984	0.914	0.948	0.904	0.929
	50	1.000	0.131	1.000	0.135	1.000	0.121	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.260	1.000	0.266	1.000	0.226	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.404	1.000	0.419	1.000	0.356	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.561	1.000	0.572	1.000	0.494	1.000	1.000	1.000	1.000	1.000	1.000
(5, 2)	15	0.522	0.067	0.573	0.085	0.962	0.069	0.544	0.397	0.577	0.924	0.613	0.835
	25	0.932	0.087	0.932	0.124	1.000	0.090	0.910	0.849	0.902	0.998	0.920	0.986
	50	1.000	0.155	1.000	0.244	1.000	0.146	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.336	1.000	0.508	1.000	0.288	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.543	1.000	0.730	1.000	0.436	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.734	1.000	0.875	1.000	0.594	1.000	1.000	1.000	1.000	1.000	1.000
(5, 5)	50	1.000	0.163	1.000	0.200	1.000	0.124	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.354	1.000	0.452	1.000	0.245	1.000	1.000	1.000	1.000	1.000	1.000
	150	1.000	0.585	1.000	0.706	1.000	0.409	1.000	1.000	1.000	1.000	1.000	1.000
	200	1.000	0.767	1.000	0.866	1.000	0.562	1.000	1.000	1.000	1.000	1.000	1.000

Zweryfikujmy teraz przypuszczenie o wpływie funkcji rozbieżności na moc testu. Ze względu na związki między postacią testu LRT i entropijnej funkcji straty oraz testu RST i kwadratowej funkcji straty chcemy sprawdzić, czy entropijna funkcja straty wskazuje na większą moc testu LRT, a kwadratowa funkcja straty na większą moc testu RST. Zależność między funkcjami  $f_k$ ,  $k \in \{E, Q\}$ , oraz mocami testów LRT i RST dla wybranych parametrów przedstawiają Rysunki 4.2 i 4.3.



Rysunek 4.2. Moce empiryczne dla  $H_{02}$  z wykorzystaniem entropijnej funkcji straty (lewy panel) oraz kwadratowej funkcji straty (prawy panel), odpowiednio dla  $\xi^{(k)} = 0.1$  (pierwszy wiersz) oraz  $\xi^{(k)} = 0.5$  (drugi wiersz),  $k \in \{E, Q\}$ , dla  $(p, q) = (3, 2)$ . Czerwony kolor przedstawia moce LRT, a niebieski RST.



Rysunek 4.3. Moce empiryczne dla  $H_{02}$  z wykorzystaniem entropijnej funkcji straty (lewy panel) oraz kwadratowej funkcji straty (prawy panel), odpowiednio dla  $\xi^{(k)} = 0.1$  (pierwszy wiersz) oraz  $\xi^{(k)} = 0.5$  (drugi wiersz),  $k \in \{E, Q\}$ , dla  $(p, q) = (5, 2)$ . Czerwony kolor przedstawia moce LRT, a niebieski RST.

Powyższe wykresy wskazują, że przypuszczenie o wpływie wyboru funkcji rozbieżności na moc testu okazało się błędne. Z Rysunków 4.2–4.3 wynika, że moce testów LRT i RST są zbliżone i nie można jednoznacznie wskazać testu lepszego.

Na zakończenie odniesiemy się do wyników opublikowanych w pracy Lu i Zimmerman [22] dla testowania struktury rozdzielnej  $H_{01}$  za pomocą testu LRT. Wykorzystamy entropijną i kwadratową funkcję straty do pomiaru rozbieżności między hipotezami dla macierzy bez struktury proponowanych intuicyjnie przez Lu i Zimmermana i przyporządkowanych do czterech typów bazujących na następującej strukturze rozdzielnej

$$\Sigma_{\varrho} = \begin{pmatrix} 1.0 & \varrho \\ \varrho & 1.0 \end{pmatrix} \otimes \begin{pmatrix} 1.0 & \varrho \\ \varrho & 1.0 \end{pmatrix} = \begin{pmatrix} 1.0 & \varrho & \varrho & \varrho^2 \\ \varrho & 1.0 & \varrho^2 & \varrho \\ \varrho & \varrho^2 & 1.0 & \varrho \\ \varrho^2 & \varrho & \varrho & 1.0 \end{pmatrix}.$$

Rozważano następujące typy macierzy kowariancji bez struktury:

- macierz typu I jest macierzą kowariancji o tych samych korelacjach co  $\Sigma_\rho$  z możliwą jedną zmianą wariancji,  $\sigma_{33} = k$ ,
- macierz typu II jest macierzą kowariancji identyczną z  $\Sigma_\rho$  z wyjątkiem  $\sigma_{34} = k$  (i symetrycznie  $\sigma_{43}$ ),
- macierz typu III jest macierzą kowariancji identyczną z  $\Sigma_\rho$  z wyjątkiem  $\sigma_{23} = k$  (i symetrycznie  $\sigma_{32}$ ),
- macierz typu IV jest macierzą kowariancji identyczną z  $\Sigma_\rho$  z wyjątkiem  $\sigma_{34} = k$  i  $\sigma_{23} = k$  (oraz symetrycznie  $\sigma_{43}$  i  $\sigma_{23}$ ).

W Tabeli 4.20 pokazano moce empiryczne testu LRT dla wymienionych typów macierzy bez struktury, które zostały oryginalnie wyliczone w pracy Lu i Zimmerman [22, Tabela 3]. Można zauważyć, że moc LRT wzrasta wraz z rozmiarem próby oraz wraz ze wzrostem zmiany między strukturą zerową i alternatywną mierzoną zmianą odpowiednich elementów zgodnie z wybranym typem. Warto zwrócić uwagę, że dla pierwszych trzech typów macierzy bez struktury poziom zmiany jest intuicyjny, ponieważ zmienia się tylko jedna wartość (druga zmienia się symetrycznie), podczas gdy dla macierzy typu IV ta miara zmiany nie jest już oczywista. Ponadto, wyniki są nieporównywalne dla różnych wartości  $\rho$  zarówno w obrębie jednego typu jak i między różnymi typami. Rzeczywiście, na przykład dla struktury typu II rozbieżność rzędu  $-0.4$ , którą można zaobserwować dla  $\rho = 0.2$  i  $k = -0.2$  wskazuje moc równą  $0.123$  dla wielkości próby równej  $20$ , podczas gdy ta sama rozbieżność zaobserwowana dla  $\rho = 0.5$  i  $k = 0.1$  wskazuje moc  $0.213$ . Dlatego naszym celem jest wykorzystanie entropijnej i kwadratowej funkcji straty mierzącej rozbieżność między hipotezami alternatywnymi wskazanymi przez Lu i Zimmermana a całym zbiorem rozważanych struktur rozdzielnych do weryfikacji przedstawionych w Tabeli 4.20 mocy. Pokażemy, że moc wzrasta wraz z rozbieżnością niezależnie od typu hipotezy alternatywnej i wartości  $\rho$ .

Zauważmy, że moce testu LRT dla macierzy typu IV zostały wyznaczone w pracy Lu i Zimmerman [22] niepoprawnie; przykładowo dla  $\rho = 0.5$ , macierz typu IV z  $k = 0.5$  i macierz typu III z  $k = 0.5$  reprezentują tę samą hipotezę alternatywną, ale wyliczone moce w tych przypadkach są różne. Ponadto, moc dla macierzy typu IV z  $\rho = 0.5$  i  $k = 0.25$  jest zbliżona do poziomu istotności  $0.05$ , co wskazuje strukturę kowariancyjną założoną w hipotezie zerowej, a to nie ma miejsca. Dlatego ponownie przeliczyliśmy moce dla macierzy typu IV. W Tabeli 4.21 moce z Lu i Zimmerman [22, Tabela 3] dla typów I–III zostały skopiowane oraz przed-

stawione są przeliczone moce dla hipotezy alternatywnej typu IV. Ponadto, dla każdego przypadku przedstawiliśmy w Tabeli 4.21 obliczone skorygowane indeksy  $\zeta_{UN}^{(k)}$ ,  $k \in \{E, Q\}$ . Oczywiście zerowa wartość odpowiedniego  $\xi^{(k)}$  odpowiada braku zmian w alternatywie, więc moc powinna być taka sama jak zakładany poziom istotności 0.05.

Tabela 4.20. Moce empiryczne testu LRT na poziomie istotności 0.05 dla testowania struktury rozdzielnej z hipotezami alternatywnymi typu I-IV (Lu i Zimmerman [22, Tabela 3]).

		Typy I i II				Typy III i IV					
Typ	$\rho$	$k$	$N = 20$	$N = 50$	Typ	$\rho$	$k$	$N = 20$	$N = 50$		
I	0.2	1	.049	.047	III	0.2	-0.40	.321	.795		
		2	.097	.218			-0.25	.143	.343		
		4	.312	.770			0.04	.049	.047		
		0.5	1	.054			.048	0.25	.090	.176	
			2	.144			.367	0.50	.333	.809	
			4	.528			.954	0.75	.890	.999	
	0.8	1	.051	.048		0.5	-0.05	.453	.933		
		2	.650	.987			0.05	.163	.425		
		4	.994	1.000			0.15	.068	.108		
		II	0.2	-0.6			.540	.963	0.25	.054	.048
				-0.4			.264	.677	0.50	.179	.485
				-0.2			.123	.293	0.75	.789	.999
-0.1	.089			.168	0.8	0.58	.407	.885			
0.2	.049			.047		0.60	.122	.302			
0.5	.099			.219		0.64	.051	.048			
0.7	.293		.754	0.75		.275	.708				
0.8	.551		.967	0.80		.529	.964				
0.9	.929		1.000	0.90		.970	1.000				
0.5	-0.1		.449	.917	IV	0.2	-0.40	.727	.997		
	0.1		.213	.553			-0.25	.276	.712		
	0.3		.089	.167			0.04	.049	.047		
	0.5	.054	.048	0.25			.134	.344			
	0.7	.132	.354	0.50			.655	.988			
	0.8	.506	.947	0.75			.997	1.000			
	0.8	0.4	.916	1.000		0.5	0.05	.678	.993		
		0.6	.437	.905			0.15	.113	.268		
		0.7	.169	.445			0.25	.054	.048		
		0.8	.051	.048			0.50	.383	.872		
		0.85	.237	.627			0.75	.988	1.000		
		0.87	.994	1.000			0.8	0.61	.367	.838	
			0.62	.111	.269						
			0.64	.051	.048						
			0.70	.253	.664						
			0.75	.629	.988						
			0.85	.994	1.000						

Z Tabeli 4.21 można zaobserwować, że entropijna funkcja straty jest dobrą miarą rozbieżności między hipotezami, ponieważ wraz ze wzrostem jej wartości wzrasta

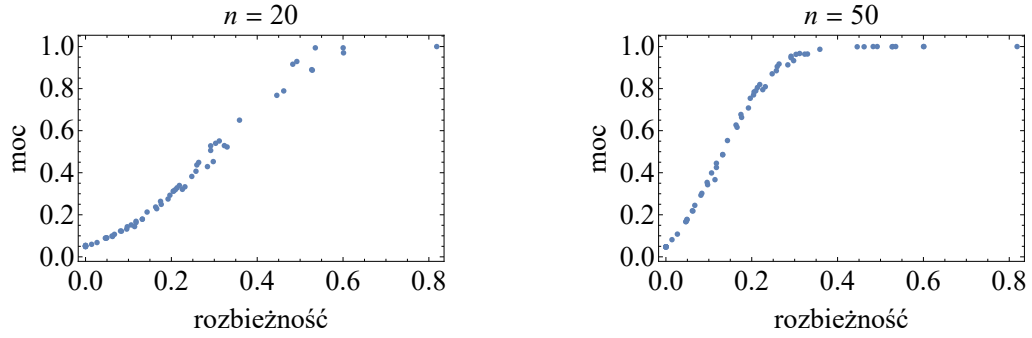
również moc testu LRT. Zjawisko to potwierdza Rysunek 4.4. Ponadto, moce testów można porównywać teraz zarówno w obrębie tego samego typu jak i między różnymi typami.

Z Tabeli 4.21 można również zauważyć, że kwadratowa funkcja straty także może służyć do badania mocy testu, wydaje się być jednak nieco słabsza niż entropijna funkcja straty, zwłaszcza dla  $n = 20$ , co potwierdza Rysunek 4.5.

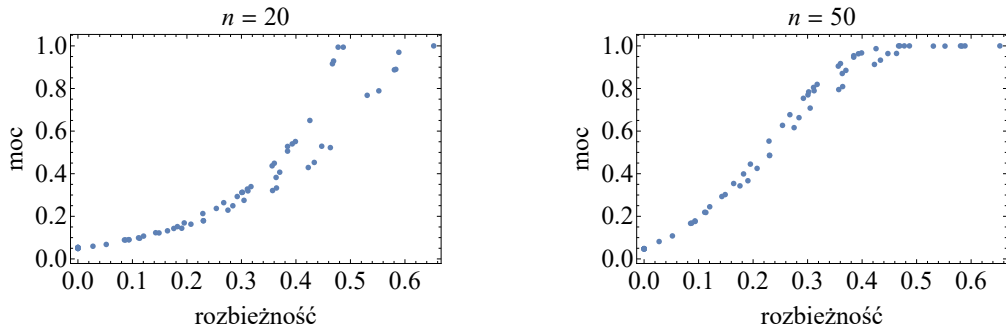
Tabela 4.21. Moce empiryczne testu LRT na poziomie istotności 0.05 dla testowania struktury rozdzielnej z hipotezami alternatywnymi typu I-III (Lu i Zimmerman [22, Tabela 3]), przeliczone moce w ramach hipotezy alternatywnej typu IV oraz indeksy rozbieżności  $\xi^{(k)}$

(skorygowane wartości  $\zeta_{\text{UN}}^{(k)}$ ),  $k \in \{E, Q\}$ .

$\rho$	$k$	$N = 20$	$N = 50$	$\zeta_{\text{UN}}^{(E)}$	$\zeta_{\text{UN}}^{(Q)}$	$\rho$	$k$	$N = 20$	$N = 50$	$\zeta_{\text{UN}}^{(E)}$	$\zeta_{\text{UN}}^{(Q)}$
Typ I						Typ III					
0.2	1	0.050	0.047	0.0000	0.0000	0.2	-0.40	0.321	0.795	0.2257	0.3571
	2	0.100	0.218	0.0629	0.1133		-0.25	0.143	0.343	0.0972	0.1759
	4	0.310	0.770	0.2042	0.3006		0.04	0.049	0.047	0.0000	0.0000
0.5	1	0.050	0.048	0.0000	0.0000	0.25	0.090	0.176	0.0492	0.0937	
	2	0.140	0.367	0.1143	0.1905	0.50	0.333	0.809	0.2316	0.3643	
	4	0.530	0.954	0.2916	0.3846	0.75	0.890	0.999	0.5270	0.5832	
0.8	1	0.050	0.048	0.0000	0.0000	0.5	-0.05	0.453	0.933	0.2976	0.4336
	2	0.650	0.987	0.3587	0.4256		0.05	0.163	0.425	0.1177	0.2073
	4	0.990	1.000	0.5354	0.4774		0.15	0.068	0.108	0.0267	0.0520
Typ II						0.25	0.055	0.048	0.0000	0.0000	
0.2	-0.6	0.540	0.963	0.3031	0.3932	0.50	0.179	0.485	0.1323	0.2302	
	-0.4	0.260	0.677	0.1747	0.2675	0.75	0.789	0.999	0.4618	0.5521	
	-0.2	0.120	0.293	0.0815	0.1427	0.8	0.58	0.407	0.885	0.2573	0.3703
	-0.1	0.090	0.168	0.0473	0.0873		0.60	0.122	0.302	0.0837	0.1488
	0.2	0.050	0.047	0.0000	0.0000	0.64	0.051	0.048	0.0000	0.0000	
	0.5	0.100	0.219	0.0616	0.1112	0.75	0.275	0.708	0.1923	0.3047	
	0.7	0.290	0.754	0.1966	0.2923	0.80	0.529	0.964	0.3235	0.4473	
	0.8	0.550	0.967	0.3116	0.3990	0.90	0.970	1.000	0.6011	0.5888	
	0.9	0.930	1.000	0.4923	0.4691	Typ IV					
0.5	-0.1	0.450	0.917	0.2636	0.3603	0.2	-0.40	0.523	0.965	0.3299	0.4632
	0.1	0.210	0.553	0.1434	0.2292	-0.25	0.229	0.616	0.1658	0.2752	
	0.3	0.090	0.167	0.0461	0.0853	0.04	0.060	0.082	0.0141	0.0276	
	0.5	0.050	0.048	0.0000	0.0000	0.25	0.090	0.178	0.0490	0.0933	
	0.7	0.130	0.354	0.0960	0.1643	0.50	0.429	0.913	0.2841	0.4226	
	0.8	0.510	0.947	0.2916	0.3846	0.75	1.000	1.000	0.8183	0.6528	
0.8	0.4	0.920	1.000	0.4829	0.4670	0.5	0.05	0.250	0.663	0.1764	0.2842
	0.6	0.440	0.905	0.2593	0.3563		0.15	0.152	0.399	0.1064	0.1825
	0.7	0.170	0.445	0.1176	0.1950	0.25	0.107	0.245	0.0673	0.1204	
	0.8	0.050	0.048	0.0000	0.0000	0.50	0.179	0.487	0.1323	0.2302	
	0.85	0.240	0.627	0.1633	0.2539	0.75	0.888	1.000	0.5284	0.5806	
	0.87	0.990	1.000	0.6002	0.4868	0.8	0.61	0.339	0.819	0.2187	0.3175
							0.62	0.328	0.805	0.2132	0.3109
							0.64	0.313	0.784	0.2054	0.3020
							0.70	0.319	0.789	0.2094	0.3119
							0.75	0.382	0.870	0.2476	0.3636
							0.85	0.768	0.999	0.4457	0.5307



Rysunek 4.4. Moce empiryczne testu LRT dla prób wielkości 20 i 50, w odniesieniu do rozbieżności między hipotezą zerową i alternatywną w sensie entropijnej funkcji straty, na podstawie 50000 prób.



Rysunek 4.5. Moce empiryczne testu LRT dla prób wielkości 20 i 50, w odniesieniu do rozbieżności między hipotezą zerową i alternatywną w sensie kwadratowej funkcji straty, na podstawie 50000 prób.

## 4.5. Dane rzeczywiste

W rozdziale tym zilustrujemy metodę identyfikacji struktury kowariancji poprzez entropijną i kwadratową funkcję straty oraz normę Frobeniusa na danych medycznych. Otrzymane wyniki porównamy z decyzją podjętą na podstawie testów LRT i RST.

W pracy McKiernan i in. [25] przedstawiono wielowymiarowe dane obejmujące masę ciała ( $c$ ), masę tkanki tłuszczowej ( $t$ ) i szacunkową masę mięśniową górnej części nogi ( $m$ ) mierzoną u 12 dorosłych samców makaków królewskich w wieku od 16 do 22 lat obserwowane po 6-tym, 9-tym i 12-tym roku badań. Badania te miały na

celu analizę wczesnego stadium sarkopenii (utruty masy mięśniowej) u starzejących się makaków. W przykładzie tym mierzono więc  $q = 3$  cech w  $p = 3$  punktach czasowych dla  $n = 12$  obiektów. Oryginalne dane dostępne są w pracy McKiernan i in. [25, Tabela 1].

Celem niniejszego rozdziału jest zweryfikowanie czy opisane w pracy metody identyfikacji struktury zależności wskazują tę samą zależność co test LRT i RST na poziomie istotności 0.05. W Tabeli 4.22 przedstawione zostały rozbieżności  $\xi^{(k)}$ ,  $k \in \{E, Q, F\}$ , oraz  $p$ -wartości wyznaczone przez rozkłady empiryczne statystyk LRT i RST przy prawdziwości hipotezy zerowej zaczerpnięte z pracy Filipiak i in. [13]. Ponieważ analiza dwóch z trzech obserwowanych cech pozwala na wnioskowanie o szczególnych przypadkach sarkopenii, przeprowadzone zostało porównanie wyników identyfikacji struktury zależności i testowania również dla par cech.

Rozbieżności oparte na entropijnej funkcji straty znaleźć można również w pracy Filipiak i in. [9], natomiast rozbieżności oparte na normie Frobeniusa w pracy Filipiak i Klein [7].

Zgodnie z oczekiwaniami dla wszystkich zestawów zmiennych rozbieżność  $\xi_{\text{UN}}^{(k)}$ ,  $k \in \{E, Q, F\}$ , jest najmniejsza, ponieważ  $\mathcal{S}_{\text{CS}}, \mathcal{S}_{\text{AR}} \subset \mathcal{S}$ . Oznacza to, że największa  $p$ -wartość powinna zostać uzyskana dla hipotezy  $H_{01}$ . Fakt ten jest potwierdzony we wszystkich przypadkach za wyjątkiem testu RST dla wszystkich trzech cech. Zauważmy ponadto, że dla wszystkich hipotez większe wartości  $\xi^{(k)}$ ,  $k \in \{E, Q\}$ , odpowiadają mniejszym  $p$ -wartościom testu LRT. Podobna sytuacja, za wyjątkiem pary  $(c, m)$  dla hipotezy  $H_{01}$  oraz trójki  $(c, t, m)$  dla  $H_{03}$ , zachodzi dla testu RST. Zależność między wielkościami rozbieżności a  $p$ -wartościami nie zachodzi dla normy Frobeniusa.

Skoncentrujmy się teraz na hipotezach  $H_{02}$  i  $H_{03}$ . Dla wszystkich zestawów cech i wszystkich funkcji rozbieżności, w większości przypadków wartości rozbieżności wskazują nieznacznie na ewentualny wybór struktury  $\Psi_{\text{AR}} \otimes \Sigma$ . Niewielkie różnice pomiędzy wartościami rozbieżności dla  $(c, t, m)$  oraz dla  $(c, m)$  i  $(t, m)$  dla  $\xi^{(Q)}$  mogą wynikać z algorytmu numerycznego. Również  $p$ -wartości wskazują na strukturę  $\Psi_{\text{AR}} \otimes \Sigma$ . Na poziomie istotności 0.01 żaden z testów nie odrzuca hipotezy  $H_{03}$ .

Zwróćmy uwagę, że wartości  $\xi^{(Q)}$  dla hipotez  $H_{02}$  i  $H_{03}$  różnią się nieznacznie, podczas gdy  $\xi^{(E)}$  mogą wykazywać większe różnice. Ponadto, we wcześniejszych badaniach symulacyjnych pojawiały się problemy z określonnością QLE oraz złożonością algorytmu obliczeniowego. W związku z tym, entropijna funkcja straty wydaje się najbardziej racjonalną miarą rozbieżności w zagadnieniu identyfikacji struktury rozbieżności.

Tabela 4.22. Rozbieżności wyznaczone przez entropijną funkcję straty, kwadratową funkcję straty i normę Frobeniusa oraz  $p$ -wartości bazujące na empirycznym rozkładzie statystyk LRT i RST przy prawdziwości hipotezy  $H_{0i}$ ,  $i \in \{1, 2, 3\}$ .

cechy		$\mathcal{S}$	$\mathcal{S}_{CS}$	$\mathcal{S}_{AR}$
$(c, t, m)$	$\xi^{(E)}$	0.7052	0.7254	0.7260
	$\xi^{(Q)}$	0.6276	0.6501	0.6502
	$\xi^{(F)}$	0.2535	0.5954	0.5745
	$p$ -wartość LRT	(0.05, 0.10)	< 0.01	(0.01, 0.05)
	$p$ -wartość RST	(0.05, 0.10)	(0.01, 0.05)	(0.10, 0.15)
$(c, t)$	$\xi^{(E)}$	0.3891	0.5443	0.5218
	$\xi^{(Q)}$	0.4362	0.5363	0.5362
	$\xi^{(F)}$	0.0519	0.3612	0.3522
	$p$ -wartość LRT	(0.85, 0.90)	(0.20, 0.25)	(0.60, 0.65)
	$p$ -wartość RST	0.75	(0.05, 0.10)	(0.50, 0.55)
$(c, m)$	$\xi^{(E)}$	0.5335	0.6317	0.6269
	$\xi^{(Q)}$	0.5366	0.5692	0.5693
	$\xi^{(F)}$	0.0877	0.6621	0.6310
	$p$ -wartość LRT	(0.10, 0.15)	< 0.01	(0.01, 0.05)
	$p$ -wartość RST	(0.01, 0.05)	< 0.01	(0.01, 0.05)
$(t, m)$	$\xi^{(E)}$	0.4971	0.6189	0.6176
	$\xi^{(Q)}$	0.5054	0.5775	0.5779
	$\xi^{(F)}$	0.2407	0.6012	0.5798
	$p$ -wartość LRT	(0.45, 0.50)	< 0.01	(0.01, 0.05)
	$p$ -wartość RST	(0.30, 0.35)	< 0.01	(0.01, 0.05)

Opisane spostrzeżenia sugerują, że wszystkie rozważane rozbieżności mogą być przydatne w identyfikacji struktury kowariancji. Identyfikacja struktury zależności nie zastąpi jednak testowania i powinna być raczej traktowana jako sposób na wskazanie najbardziej racjonalnej hipotezy. Zwróćmy uwagę, że mimo modyfikacji  $\zeta$  do przedziału  $[0, 1)$  potrafimy wskazać najbliższą strukturę spośród zadanych, jednakże nie potrafimy dokonać oceny wielkości rozbieżności (stwierdzić, czy ta rozbieżność jest wystarczająco mała aby uznać, że dane są związane określonym typem zależności).



## 5. Podsumowanie

W niniejszej pracy wyprowadzono wzory na najlepsze przybliżenia określonej dodatnio macierzy  $\Omega$  rozdzielną strukturą kowariancyjną  $\Psi \otimes \Sigma$ ,  $\Psi_{CS} \otimes \Sigma$  lub  $\Psi_{AR} \otimes \Sigma$ , z zastosowaniem entropijnej lub kwadratowej funkcji straty jako miar rozbieżności oraz opracowano algorytmy numeryczne wyznaczania odpowiednich rozwiązań. Ponadto, w przypadku struktury  $\Psi_{CS} \otimes \Sigma$  zaproponowano trzy metody poszukiwania najlepszych przybliżeń za pomocą entropijnej funkcji straty i porównano ich tempo zbieżności.

W Rozdziale 4 przeprowadzono badania symulacyjne w celu zastosowania opracowanych algorytmów optymalizacyjnych do rozwiązywania problemów statystycznych.

Najważniejszym wnioskiem jest to, że obie funkcje prawidłowo rozpoznają strukturę kowariancji. Dla małych wartości współczynników korelacji  $\rho$  zbiory  $\mathcal{S}_{CS}$  i  $\mathcal{S}_{AR}$  są do siebie bardzo zbliżone, stąd mniejsza proporcja poprawnie rozpoznanych struktur. Jednakże wraz ze wzrostem  $\rho$  oraz  $p$  wspomniane zbiory różnią się od siebie coraz bardziej, co skutkuje bardzo wysoką rozpoznawalnością struktury kowariancyjnej.

Ponadto pokazano, że minima otrzymane proponowanymi metodami mogą być traktowane jako estymatory struktur kowariancyjnych, jednakże ich własności statystyczne są zazwyczaj słabsze niż powszechnie stosowanych estymatorów największej wiarygodności. Zatem, aby posługiwać się wspomnianymi estymatorami, warto byłoby je zmodyfikować, na przykład poprzez zastosowanie innego estymatora macierzy bez struktury na wstępie procedur, przez modyfikację estymatora  $\mathbf{S}$  (por. Stein [39]) lub przez uwzględnienie funkcji kary w procesie minimalizacji.

W przypadku zastosowania entropijnej i kwadratowej funkcji straty do badania mocy testów pokazano, że funkcje te mogą być użyte jako miary rozbieżności nie tylko ze względu na interpretowalność (jako rozbieżność między dwoma rozkładami prawdopodobieństwa), lecz również ze względu na zachowanie podstawowych własności mocy (monotoniczność względem rozbieżności). Przypuszczenie, że entropijna funkcja straty może wskazywać na większą moc testu LRT, natomiast kwadratowa funkcja straty na większą moc testu RST, okazało się błędne. Moce testów LRT i RST są zbliżone i nie można jednoznacznie wskazać testu lepszego.

Wnioski płynące z analizy danych rzeczywistych sugerują, że obie rozważane funkcje rozbieżności jak i norma Frobeniusa mogą być zastosowane do identyfikacji struktury kowariancji. Identyfikacja struktury zależności nie zastąpi jednak

testowania i powinna być raczej traktowana jako sposób na wskazanie najbardziej racjonalnej hipotezy.

Z przeprowadzonych badań wynika, że rozważane funkcje z powodzeniem mogą być stosowane do rozwiązywania problemów statystycznych. Ponieważ jednak aproksymacja kwadratową funkcją straty nie zawsze daje rozwiązanie określone dodatnio oraz kumulować mogą się błędy numeryczne związane ze złożonością obliczeniową algorytmów, entropijna funkcja straty wydaje się najbardziej racjonalną miarą rozbieżności.

Istotnym założeniem poczynionym w niniejszej pracy jest nieosobliwość estymatora macierzy kowariancji bez struktury. Warunek ten można by jednak osłabić. W pracy Chen i in. [2] zaproponowano metodę identyfikacji macierzy kowariancji za pomocą entropijnej funkcji straty w standardowym modelu wielowymiarowym ze zmodyfikowanym osobliwym estymatorem nieustrukturyzowanej macierzy kowariancji. Podejście to można rozszerzyć również na model podwójnie wielowymiarowy.

# Bibliografia

- [1] T.W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, 2003.
- [2] C. Chen, J. Zhou, J. Pan. Correlation structure regularization via entropy loss function for high-dimension and low-sample-size data. *Comm. Statist. Simulation Comput.* 2019. <https://doi.org/10.1080/03610918.2019.1571607>.
- [3] X. Cui, C. Li, J. Zhao, L. Zeng, D. Zhang, J. Pan. Covariance structure regularization via Frobenius norm discrepancy. *Linear Algebra Appl.* 510, 124–145, 2016.
- [4] E. Devijver, M. Gallopın. Block-diagonal covariance selection for high-dimensional Gaussian graphical models. *J. Am. Stat. Assoc.* 113 (521), 306–314, 2018.
- [5] D.K. Dey, C. Srinivasan. Estimation of a covariance matrix under Stein’s loss. *Ann. Statist.* 13(4), 1581–1591, 1985.
- [6] P.L. Fackler. Notes on matrix calculus. <http://www4.ncsu.edu/~pfackler/MatCalc.pdf>, 2005.
- [7] K. Filipiak, D. Klein. Approximation with a Kronecker product structure with one component as compound symmetry or autoregression. *Linear Algebra Appl.* 559, 11–33, 2018.
- [8] K. Filipiak, D. Klein. Estimation and testing the covariance structure of doubly multivariate data. In: Filipiak, K., Markiewicz, A., von Rosen, D. (Eds.) *Multivariate, Multilinear and Mixed Linear Models*, Springer, 2021 (w druku).
- [9] K. Filipiak, D. Klein, A. Markiewicz, M. Mokrzycka. Approximation with a Kronecker product structure with one component as compound symmetry or autoregression via entropy loss function. *Linear Algebra Appl.* 610, 625–646, 2021.
- [10] K. Filipiak, D. Klein, M. Mokrzycka. Estimators comparison of separable covariance structure with one component as compound symmetry matrix. *Electronic J. Linear Algebra* 33, 83–98, 2018.
- [11] K. Filipiak, D. Klein, M. Mokrzycka. Separable covariance structure identification for doubly multivariate data. In: Filipiak, K., Markiewicz, A., von Rosen, D. (Eds.) *Multivariate, Multilinear and Mixed Linear Models*, Springer, 2021 (w druku).
- [12] 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, 105–124, 2016.
- [13] K. Filipiak, D. Klein, A. Roy. A comparison of likelihood ratio tests and Rao’s score test for three separable covariance matrix structures. *Biom. J.* 59, 192–215, 2017.
- [14] K. Filipiak, D. Klein, E. Vojtková. The properties of partial trace and block trace operators of partitioned matrix. *Electron. J. Linear Algebra* 33, 3–15, 2018.

- [15] K. Filipiak, A. Markiewicz, A. Mieldzioc, A. Sawikowska. On projection of a positive definite matrix on a cone of nonnegative definite Toeplitz matrices. *Electronic J. Linear Algebra* 33, 74–82, 2018.
- [16] M. Gilson, D. Dahmen, R. Moreno-Bote, A. Insabato, M. Helias. The covariance perceptron: A new framework for classification and processing of time series in recurrent neural networks. bioRxiv, 2019. DOI: <https://doi.org/10.1101/562546>.
- [17] L.R. Haff. Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.*, 8(3), 586–597, 1980.
- [18] W. James, C. Stein. Estimation with quadratic loss. In: Neyman, J. Ed. *Proceedings of the Fourth Berkeley Symposium. In: Mathematical Statistics and Probability, vol. 1 (pp. 361–379). The Statistical Laboratory, University of California, June 30–July 30, 1960*. University of California Press, 1961.
- [19] T.G. Kolda, B.W. Bader. Tensor decompositions and applications. *SIAM Review* 51(3), 455–500, 2009.
- [20] T. Kollo, D. von Rosen. *Advanced Multivariate Statistics with Matrices*. Dordrecht, Springer, 2005.
- [21] L. Lin, N. Higham, J. Pan. Covariance structure regularization via entropy loss function. *Comput. Statist. Data Anal.* 72, 315–327, 2014.
- [22] N. Lu, D.L. Zimmerman. On likelihood-based inference for a separable covariance matrix. Technical Report No. 337, Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA, 2004.
- [23] N. Lu, D.L. Zimmerman. The likelihood ratio test for a separable covariance matrix. *Statist. Probab. Lett.* 73, 449–457, 2005.
- [24] J. Magnus, H. Neudecker. Symmetry, 0-1 matrices and Jacobians, a review. *Econom. Theory* 2, 157–190, 1986.
- [25] S.H. McKiernan, R.J. Colman, M. Lopez, T.M. Beasley, R. Weindruch, J.M. Aiken. Longitudinal analysis of early stage Sarcopenia in aging rhesus monkeys. *Experimental Gerontology* 44, 170–176, 2009.
- [26] A.M. Manceur, P. Dutilleul. Unbiased modified likelihood ratio tests for simple and double separability of a variance-covariance structure. *Statist. Probab. Lett.* 83, 631–636, 2013.
- [27] M. Mitchell, M. Genton, M. Gumpertz. A likelihood ratio test for separability of covariances. *J. Multivariate Anal.* 97, 1025–1043, 2006.
- [28] R.J. Muirhead. *Aspects of Multivariate Statistical Theory*. Wiley, Hoboken, 2005.
- [29] J. Pan, K. Fang. *Growth Curve Models and Statistical Diagnostics*. Springer-Verlag, New York, 2002.
- [30] C.R. Rao. Score test: Historical review and recent developments. In: *Advances in Ranking and Selection, Multiple Comparisons, and Reliability* 2005, 3–20.

- [31] A. Roy. A note on testing of Kronecker product covariance structures for doubly multivariate data. In: *Proceedings of American Statistics and Associations Statistical Computing Section*, pp. 2157–2162. Seattle, WA, 2007.
- [32] A. Roy, R. Khattree. On implementation of a test for Kronecker product covariance structure for multivariate repeated measures data. *Stat. Methodol.* 2, 297–306, 2005.
- [33] A. Roy, R. Khattree. Testing the hypothesis of a Kronecker product covariance matrix in multivariate repeated measures data. In: *Proceedings of the 30th Annual SAS Users Group International Conference (SUGI 30)*. Philadelphia, PA, 2005.
- [34] A. Roy, R. Leiva. Likelihood ratio tests for triply multivariate data with structured correlation on spatial repeated measurements. *Statist. Probab. Lett.* 78, 1971–1980, 2008.
- [35] S.L. Simpson. An adjusted likelihood ratio test for separability in unbalanced multivariate repeated measures data. *Stat. Methodol.* 7, 511–519, 2010.
- [36] I. Soloveychik, D. Trushin. Gaussian and robust Kronecker product covariance estimation: Existence and uniqueness. *J. Multivariate Anal.* 149, 92–113, 2016.
- [37] M. Srivastava, R. von Rosen, D. von Rosen. Models with a Kronecker product covariance structure: Estimation and testing. *Math. Methods Statist.* 17, 357–370, 2008.
- [38] C. Stein. *On Some Problems in Multivariate Analysis. Part I*. Department of Statistics, Stanford University, 1956.
- [39] C. Stein, Lectures on the theory of estimation of many parameters. *J. Math. Sci.* 34, 1373–1403, 1986.