

Wielowymiarowe modele w analizie danych biologicznych

Monika Mokrzycka

Instytut Genetyki Roślin PAN w Poznaniu

Warsztaty, Politechnika Warszawska

- Wstęp
- Jednowymiarowy model liniowy
- Wielowymiarowy model liniowy
- Zestawy danych
- Identyfikacja macierzy kowariancji
- Struktury kowariancyjne
- Aproksymacja
- Badania symulacyjne

Główne problemy w analizie danych:

- wyznaczenie wartości nieznanymi parametrów
 - dane są obciążone błędami pomiarów,
 - dane zawierają niewyjaśnione szumy, zakłócenia,
 - dlatego można wyznaczyć tylko jakieś oszacowania nieznanymi parametrów populacji
- testowanie hipotez dotyczących wartości nieznanymi parametrów
 - oznacza decydowanie czy dane są zgodne na pewnym poziomie porozumienia z określonym parametrem populacji

Zbierając dane coraz więcej wielkości mierzy się na każdej jednostce próby.

Czasami lepiej byłoby wyizolować każdą cechę w systemie, aby przestudiować ją oddzielnie – prowadzi to do tzw. modelu jednowymiarowego.

Jednak cechy mogą mieć wpływ na siebie nawzajem w takim stopniu, że oddzielna analiza dawałaby słabe informacje o całym systemie - na pierwszy plan wysuwają się modele wielowymiarowe, które badają wszystkie wielkości jednocześnie.

W porównaniu z modelem jednowymiarowym, analiza modeli wielowymiarowych wymaga trudniejszych technik wnioskowania statystycznego.

Modele liniowe

- są głównym narzędziem statystyki stosowanej
- są intensywnie studiowane
- bogata literatura statystyczna dotycząca różnych aspektów modeli liniowych i związanych z nimi metod statystycznych
- zyskały popularność dzięki własnościom oraz wszechstronnością zastosowania

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

$\mathbf{y} = (y_1, \dots, y_n)'$ wektor obserwacji
 $\mathbf{X}_{n \times k}$ znana macierz układu (pełnego rzędu)
 $\boldsymbol{\beta}_{k \times 1}$ wektor nieznanych parametrów (1-ego stopnia), $k < n$
 \mathbf{e}_n wektor błędów losowych

Założenie:

$$\mathbf{e} \sim N_n(\mathbf{0}, \boldsymbol{\Omega})$$

$\boldsymbol{\Omega}_{n \times n}$ macierz nieznanych parametrów (2-ego stopnia)

Podstawowy problem: wnioskowanie o nieznanych parametrach pierwszego stopnia

Jednowymiarowy model liniowy

Liczba nieznanych parametrów:

$$\left. \begin{array}{l} \beta_{k \times 1} \quad k \\ \Omega_{n \times n} \quad \frac{n(n+1)}{2} \end{array} \right\} \text{przeparametryzowany model}$$

Jednowymiarowy model liniowy

Liczba nieznanymi parametrów:

$$\beta_{k \times 1} \quad k$$

$$\Omega_{n \times n} \quad 1$$

Struktura macierzy kowariancji:

$$\Omega = \sigma^2 \mathbf{I}$$

lub

$$\Omega = \sigma^2 \mathbf{V}, \text{ gdzie } \mathbf{V} \text{ znane}$$

σ^2 nieznanymi parametr wariacyjny

Jednowymiarowy model liniowy

Na każdej jednostce próby zwykle mierzona jest jedna charakterystyka.

Uogólnienie:

założmy jednowymiarowy model dla p zmiennych zależnych, czyli więcej niż jedna cecha jest mierzona na każdej jednostce próby - należy wziąć pod uwagę zależność pomiarów od tej samej jednostki próby

Przykłady:

- p wyników z egzaminu (z różnych przedmiotów) dla n różnych studentów
- p charakterystyk mierzonych na n roślinach (np. różnych odmianach)
- p cech pogodowych (np. temperatura, opady, ciśnienie) mierzonych dla n odległych punktów czasowych lub lokalizacji

Wymóg: ta sama macierz układu dla każdej zmiennej zależnej i każda niezależna jednostka próby ma ten sam zestaw zmiennych zależnych

Jednowymiarowy model liniowy

Dla j -tej zmiennej zależnej (cechy) mamy

$$\mathbf{y}_j = \mathbf{X}\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, p$$

$$E(\boldsymbol{\varepsilon}_j) = \mathbf{0}, \quad \text{Var}(\boldsymbol{\varepsilon}_j) = \omega_{jj}^2 \mathbf{I}_n$$

$\mathbf{y}_j = (y_{1j}, \dots, y_{nj})'$ wektor obserwacji dla j -tej zmiennej
 $\mathbf{X}_{n \times k}$ znana macierz układu (pełnego rzędu)
 $\boldsymbol{\beta}_j$ k – wymiarowy wektor nieznanymi parametrów (1-ego stopnia), $k < n$
 $\boldsymbol{\varepsilon}_j$ n – wymiarowy wektor błędów losowych

Modele są powiązane $\frac{p(p+1)}{2}$ kowariancjami

$$\text{Cov}(\mathbf{y}_i, \mathbf{y}_j) = \omega_{ij}^2 \mathbf{I}_n$$

Jednowymiarowy model liniowy

Dla j -tej zmiennej zależnej (cechy) mamy

$$\mathbf{y}_j = \mathbf{X}\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, p$$

$$E(\boldsymbol{\varepsilon}_j) = \mathbf{0}, \quad \text{Var}(\boldsymbol{\varepsilon}_j) = \omega_{jj}^2 \mathbf{I}_n$$

Obserwacje

$$\mathbf{y}_j, \quad j = 1, \dots, p \quad \Rightarrow \quad \mathbf{Y}$$

Nieznane parametry

$$\boldsymbol{\beta}_j, \quad j = 1, \dots, p \quad \Rightarrow \quad \mathbf{B}$$

Błędy losowe

$$\boldsymbol{\varepsilon}_j, \quad j = 1, \dots, p \quad \Rightarrow \quad \mathbf{E}$$

Wielowymiarowy model liniowy

Standardowy model wielowymiarowy

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

$$\mathbf{E}(\mathbf{E}) = \mathbf{0}, \quad \text{Var}(\mathbf{E}) = \mathbf{\Omega} \otimes \mathbf{I}_n$$

$\mathbf{Y}_{n \times p} = (\mathbf{y}_1, \dots, \mathbf{y}_n)'$	macierz obserwacji
$\mathbf{X}_{n \times k}$	znana macierz układu
$\mathbf{B}_{k \times p} = (\beta_1, \dots, \beta_p)$	macierz nieznanych parametrów (1-ego stopnia), $k < n$
$\mathbf{E}_{n \times p} = (\varepsilon_1, \dots, \varepsilon_p)$	macierz błędów losowych
$\mathbf{\Omega}_{p \times p} = (\omega_{ij}^2)_{ij}$	macierz nieznanych parametrów (2-ego stopnia)
\otimes	iloczyn Kroneckera

Założenia:

- $\mathbf{E} \sim N_{n,p}(\mathbf{0}, \mathbf{I}_n, \mathbf{\Omega})$
- $n > p + r(\mathbf{X})$

Wielowymiarowy model liniowy

Model ten można traktować jako próbę n niezależnych obserwacji o p wartościach

- oznaczając jako \mathbf{y}'_i i \mathbf{x}'_i i -ty wiersz macierzy \mathbf{Y} i \mathbf{X} odpowiednio

$$\mathbf{y}_i \sim N_p(\mathbf{B}'\mathbf{x}_i, \mathbf{\Omega}), \quad i = 1, \dots, n$$

Układ (ustawienie) danych

- horyzontalny ("klasyczny") - niezależne p -wartościowe obserwacje są ułożone poziomo jedna pod drugą w wierszach, tworząc $n \times p$ wymiarową macierz obserwacji \mathbf{Y} , zachowując kierunek układania obserwacji jak w jednowymiarowym modelu liniowym

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}'_1 \\ \mathbf{y}'_2 \\ \vdots \\ \mathbf{y}'_n \end{pmatrix}, \quad \text{zatem} \quad \underset{n \times p}{\mathbf{Y}} = \underset{n \times k}{\mathbf{X}} \underset{k \times p}{\mathbf{B}} + \underset{n \times p}{\mathbf{E}}$$

Wielowymiarowy model liniowy

Model ten można traktować jako próbę n niezależnych obserwacji o p wartościach

- oznaczając jako \mathbf{y}'_i i \mathbf{x}'_i i -ty wiersz macierzy \mathbf{Y} i \mathbf{X} odpowiednio

$$\mathbf{y}_i \sim N_p(\mathbf{B}'\mathbf{x}_i, \mathbf{\Omega}), \quad i = 1, \dots, n$$

Układ (ustawienie) danych

- wertykalny - niezależne p -wartościowe obserwacje są ułożone pionowo jedna obok drugiej w kolumnach, tworząc $p \times n$ wymiarową macierz obserwacji \mathbf{Y} - jest to transponowanie poprzedniej aranżacji danych

$$\mathbf{Y}_* = \begin{pmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_n \end{pmatrix}, \quad \text{zatem} \quad \underset{p \times n}{\mathbf{Y}_*} = \underset{p \times k}{\mathbf{B}_*} \underset{k \times n}{\mathbf{X}_*} + \underset{p \times n}{\mathbf{E}_*}$$

Dla $\mathbf{Y}_* = (\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_n)$

- $E(\mathbf{Y}_*) = \mathbf{B}'\mathbf{X}'$, ponieważ $\mathbf{B}_* = \mathbf{B}'$ i $\mathbf{X}_* = \mathbf{X}'$
- rozkład $\text{vec } \mathbf{Y}_*$ ma postać

$$\text{vec } \mathbf{Y}_* = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{pmatrix} \sim N_{pn}(\text{vec}(\mathbf{B}'\mathbf{X}'), \mathbf{I}_n \otimes \mathbf{\Omega})$$

Zatem \mathbf{Y}_* ma macierzowy rozkład normalny

$$\mathbf{Y}_* \sim N_{p,n}(\mathbf{B}'\mathbf{X}', \mathbf{\Omega}, \mathbf{I}_n)$$

Ponieważ $\mathbf{Y} = \mathbf{Y}'_*$ to:

- $E(\mathbf{Y}) = E(\mathbf{Y}'_*) = \mathbf{XB}$
- rozkład $\text{vec } \mathbf{Y}$ ma postać

$$\text{vec } \mathbf{Y} = \mathbf{K}_{p,n} \text{vec } \mathbf{Y}' \sim N_{pn} \left(\mathbf{K}_{p,n} \text{vec}(\mathbf{B}'\mathbf{X}'), \mathbf{K}_{p,n}(\mathbf{I}_n \otimes \boldsymbol{\Omega})\mathbf{K}_{n,p} \right),$$

gdzie $\mathbf{K}_{p,n}$ jest macierzą komutacji.

Ponieważ $\mathbf{Y} = \mathbf{Y}'_*$ to:

- $E(\mathbf{Y}) = E(\mathbf{Y}'_*) = \mathbf{XB}$
- rozkład $\text{vec } \mathbf{Y}$ ma postać

$$\text{vec } \mathbf{Y} = \mathbf{K}_{p,n} \text{vec } \mathbf{Y}' \sim N_{pn} \left(\underbrace{\mathbf{K}_{p,n} \text{vec}(\mathbf{B}'\mathbf{X}')}_{=\text{vec}(\mathbf{XB})}, \underbrace{\mathbf{K}_{p,n}(\mathbf{I}_n \otimes \mathbf{\Omega})\mathbf{K}_{n,p}}_{=\mathbf{\Omega} \otimes \mathbf{I}_n} \right),$$

gdzie $\mathbf{K}_{p,n}$ jest macierzą komutacji.

Ponieważ $\mathbf{Y} = \mathbf{Y}'_*$ to:

- $E(\mathbf{Y}) = E(\mathbf{Y}'_*) = \mathbf{XB}$
- rozkład $\text{vec } \mathbf{Y}$ ma postać

$$\text{vec } \mathbf{Y} = \mathbf{K}_{p,n} \text{vec } \mathbf{Y}' \sim N_{pn} \left(\text{vec}(\mathbf{XB}), \mathbf{\Omega} \otimes \mathbf{I}_n \right),$$

gdzie $\mathbf{K}_{p,n}$ jest macierzą komutacji.

Zatem \mathbf{Y} ma macierzowy rozkład normalny

$$\mathbf{Y} \sim N_{n,p}(\mathbf{XB}, \mathbf{I}_n, \mathbf{\Omega})$$

Model

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

$$\mathbf{E} \sim N_{n,p}(\mathbf{0}, \mathbf{I}_n, \mathbf{\Omega})$$

jest bardzo ogólny. Obejmuje:

- wielowymiarowy model regresji

\mathbf{X} jest macierzą stałych regresji, zazwyczaj $\mathbf{1}_n \in \mathcal{C}(\mathbf{X})$

- MANOVA model

\mathbf{X} jest 0-1 macierzą układu

- ogólny model średniej

$\mathbf{X} = \mathbf{1}_n$, czyli $\mathbf{B}' = \mu$ jest p – wymiarowym wektorem

Model

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

$$\mathbf{E} \sim N_{n,p}(\mathbf{0}, \mathbf{I}_n, \mathbf{\Omega})$$

jest bardzo ogólny. Obejmuje:

- uogólniony model MANOVA (nazywany modelem krzywej wzrostu)
 - rozważa $\underset{p \times 1}{\boldsymbol{\beta}}_i = \underset{p \times r}{\mathbf{Z}} \underset{r \times 1}{\mathbf{b}}_i, i = 1, \dots, k$
 - modelem krzywej wzrostu

$$\mathbf{Y} = \mathbf{XBZ}' + \mathbf{E}, \quad \text{gdzie } \mathbf{B} = \begin{pmatrix} \mathbf{b}'_1 \\ \mathbf{b}'_2 \\ \vdots \\ \mathbf{b}'_k \end{pmatrix}$$

- redukuje się do standardowego modelu dla $\mathbf{Z} = \mathbf{I}_p$.

- Nowoczesne techniki badawcze pozwalają na zbieranie bardzo dużych wielowymiarowych zestawów danych.
- Wykonuje się eksperymenty, w których wiele cech mierzonych jest w wielu lokalizacjach, na różnych głębokościach albo pomiary powtarzane są tak często jak to tylko możliwe.
- Dane metabolomiczne - gazowa lub cieczowa chromatografia ze spektrometrią mas
- Często pomiary wykonywane są z użyciem kamer czy dronów

Dane:

- fenotypowe
- genotypowe
- omiczne

Co to jest fenotypowanie: film

Fenotypowanie buraków: film

Dane omiczne: film

Genotyp a fenotyp: film

Zestawy danych

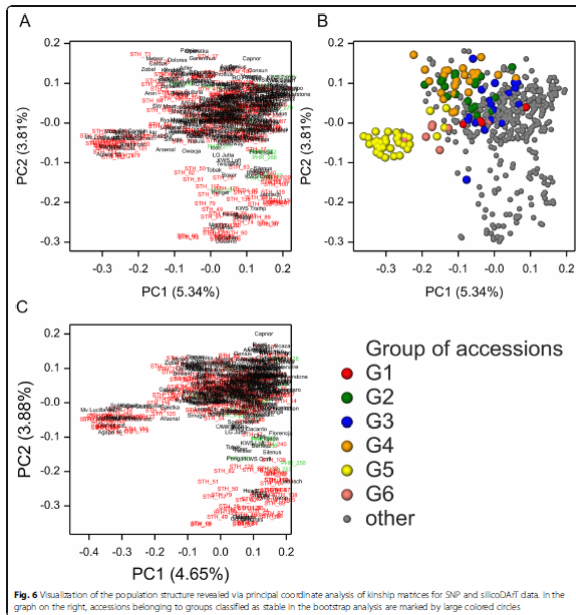
Genotyp	Biomasa całkowita			Biomasa korz			Dł. Korzenia			l. źdźbeł prod			ogólna l. źdźbeł			Wyso
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
1	10,97	21,24	9,25	1,94	2,77	1,39	17	15	15	5	11	4	5	12	4	4
2	13,74	7,64	10,18	2,28	1,59	2,17	10	14	14	5	3	5	5	3	5	5
3	12,04	9,35	16,29	1,88	1,56	2,1	11	15	15	5	6	8	5	6	8	8
4	14,51	6,55	13,2	1,59	1,23	1,91	15	17	19	6	4	6	6	4	6	6
5	11,04	7,15	4,2	1,32	0,6	0,5	17	7	9	5	4	2	5	4	2	2
6	9,05	11,32	11,4	0,32	0,84	1,66	14	13	15	5	6	5	5	8	5	5
7	11,59	12,32	16,85	2,41	2,59	2,76	17	14	13	7	6	10	7	6	10	10
8	20,78	9,35	8,82	3,22	1,7	2,48	22	19	12	8	3	4	9	3	4	4
9	14,08	8,55	10,91	1,44	1,49	1,64	13	13	14	7	5	5	9	5	5	5
10	15,75	9,2	10,55	3,55	2,11	1,3	13,5	14	13	7	4	5	7	5	5	5
11	10,72	9	9,51	1,62	1,2	0,69	17	14	13	4	5	5	4	5	5	5
12	12,66	15,07	12,12	1,65	2,92	3,24	14	15	19	7	7	4	7	7	4	4
13	8,83	7,56	7,69	2,26	0,82	1,34	19	16	13	4	4	4	4	6	4	4
14	9,27	6,85	8,98	1,36	1,22	1,48	11	12	11	4	2	5	4	4	6	6
15	10,27	6,95	7,2	1,61	1,02	0,89	17	19	14	5	3	4	5	3	5	5
16	15,71	14,93	6,38	1,66	2,71	1,78	14	15	13	6	6	2	6	6	3	3
17	8,22	13,37	8,55	1,9	1,54	2,01	15	22	17	3	5	3	3	5	3	3
18	11,01	8,58	5,93	2,65	1,83	0,78	17	16	17	4	5	3	5	6	3	3
19	5,04	9,2	5,63	0,79	1,34	0,71	13	13	14	4	6	3	4	7	3	3
20	10,17	4,65	9,92	1,11	0,8	1,45	19	18	17	5	3	5	5	3	5	5
21	6,65	8,11	10,66	0,69	1,19	1,85	26	26	22	4	5	6	4	5	6	6
22	23,25	8,34	5,58	2,9	1,42	1,2	17	12	16	10	5	3	10	6	3	3
23	9,32	10,23	7,35	1,09	1,4	1,33	13	21	20	6	5	5	6	5	5	5
24	8,12	8,64	5,43	0,98	1,75	0,99	17	15	13	5	4	4	5	4	4	4
25	12,32	8,76	3,3	1,9	1,05	0,55	19	18	15	7	6	3	7	6	3	3
26	8,56	8,4	9,24	1,5	1,5	1,86	13	15	14	5	4	5	5	4	6	6
27	7,55	13,86	11,04	1,59	4,39	2,72	13,5	16	16	4	7	6	6	8	7	7
28	8,04	8,85	9,05	2,64	1,62	2,27	14	15	14,5	4	6	5	4	6	5	5
29	5,47	7,94	7,63	0,69	1,52	1,18	14	19	16	3	4	4	3	4	4	4
30	15,41	12,02	8,36	1,64	2,42	0,98	17	14	13	6	4	4	6	5	4	4
31	8,15	6,16	7,8	0,84	1,14	1,24	18	16	10	5	5	5	8	5	6	6
32	8,89	6,32	5,98	1,94	1,96	1,64	15,5	16	15	4	4	3	4	4	4	4
33	12,26	7,65	9,37	1,46	0,92	0,77	17	18	19	7	6	6	7	7	6	6
34	15,97	12,88	11,61	3,86	2,72	2,27	16	20	19	7	8	7	7	8	7	7

Zestawy danych

Obiekt_1	Plon_2018_a1_w	Plon_2017_a1_w	Plon_2016_a1_w	Plon_2015_a1_w	Plon_2018_a2_w	Plon_2017_a2_w	Plon_2016_a2_w	Plon_2015_a2_w	ZimotrwaA_oA	Choroby_podsta	MA_cznik_praw	Rdza_brunatna	Rdza_A4A3A_ta	Brunatna
Arkadia	96	91	98	94	99	96	103	99	6	7,6	6,4	7,7	5,8	
Arktis			92	93			90	92	4,5	7,5	7,9	7,1	8,7	
Artist	102	101	104	102	103	101	104	103	4	7,8	7,8	7,5	8,5	
Askalon			90	99			92	100	2,5	7,9	7,7	7,9	8,7	
Astoria		95	86	87		96	88	90	3	7,8	7,3	7,6	8,1	
Bamberka		92	88	94		94	90	95	3	7,7	7,9	7,2	8,6	
Bartosz		99	96	103		100	93	103	3,5	8,2	7,7	7,5	8,4	
Belissa	96	97	102	95	98	100	103	99	5	8	7,7	7,4	7,1	
Blyskawica	98	103	96		100	103	101		4	7,9	6,9	8,4	8,7	
Bogatka				92				95	5,5	7,8	7,7	7,5	6,7	
Bonanza	102	96		106	103	97		103	4	8,2	8,3	8,1	8,1	
Comandor	100	102	109		101	102	104		4,5	8,3	7,8	8,3	8,7	
Dakar		98	77	101		96	73	100	2	7,9	8,2	8,3	8,6	
Delawar	97	99	101		97	97	101		4	8,2	8,2	8,2	8,9	
Dolores	99	99	98	101	98	98	96	101	4	7,8	8,1	8,2	8,6	
Estivus			91	98			91	98	3	7,9	7,9	8	8,6	
Euforia	100	103	112		102	102	107		5,5	8	8,1	8	8,6	
Fakir	96	98	98	97	96	96	99	97	4,5	8	8,1	7,9	8,8	
Fidelius		98	97	97		97	99	97	4,5	7,7	7,9	7,7	8,5	
Florencja			88				88		2	7,8	7,4	8	7,8	
Florus		97	88	103		96	87	101	3	7,6	7,9	7,9	8,7	
Formacja	99	104	102	99	99	102	101	98	4,5	7,6	7,8	7,9	8,3	
Forum			100				100		2	7,2	7,5	7,3	8,8	
Franz		98	86	103		97	88	103	3	7,8	8,2	7,9	8,8	
Frisky	101	102	98	104	99	101	98	102	3	8,1	8	8,4	8,4	
Gimantis		99				100			3,5	7,8	8,3	7,8	8,5	
Hondia	98	100	101	96	96	98	101	94	5,5	8,1	8,3	7,9	8,8	
Janosch	96	98	94		99	98	95		3	8,3	7	7,9	8,6	
Jantarka		96	102	96		96	100	98	5,5	7,6	7,8	7,5	8,3	
Kepler				98				98	2,5	8,5	7,7	7,6	8,9	
Kometa		99	63	99		96	64	98	2	8,3	8,2	8,2	8,8	
Kredo			59	103			63	103	2	8,2	7,8	7,3	8,8	
KWS Dacanto			85	106			86	105	2	7,9	7,6	8,1	8,8	
KWS Livius		99	90	98		97	92	99	3	7,9	7,6	7,6	8,9	
KWS Loft			76	101			78	101	2	7,9	8,1	8,1	7,6	
KWS Magic			86	98			86	100	1,5	7,8	7,5	7	7,9	
KWS Malibu			80				80		2	8,2	8,2	8,2	8,7	
KWS Ozon	94	97	96	101	97	97	97	100	4	7,7	7,8	7,5	8,8	
KWS Spencer	98	101	100	101	97	101	100	101	4,5	7,8	7,8	7,9	8,7	
Lavantus		91	90	97		91	89	97	3	7,4	6,9	7,7	7,7	

Zestawy danych

rs	alleles	chrom	pos	strand	assembly	center	STH_1	STH_2	STH_3	STH_4	STH_5	STH_6	STH_7	STH_8	STH_9	STH_10	STH_11	STH_12	STH_13	STH_14
5325144 F 0-9 GC	G/C	NA	NA	NA	Wheat_ChineseSpring10	DarT	CC	NN	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
1033260 F 0-26 TC	T/C	chrUn	330012453	NA	Wheat_ChineseSpring10	DarT	TT	TC	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT
1033260 F 0-44 AG	A/G	chrUn	330012453	NA	Wheat_ChineseSpring10	DarT	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
1306353 F 0-8 TC	T/C	chr5D	326383801	NA	Wheat_ChineseSpring10	DarT	TC	TT	TC	TC	TC	TT	TT	TT	TC	TC	TC	TT	TC	TT
7344639 F 0-14 CA	C/A	NA	NA	NA	Wheat_ChineseSpring10	DarT	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
3955947 F 0-15 CT	C/T	chrUn	328699296	NA	Wheat_ChineseSpring10	DarT	CT	CT	CT	CT	CT	CC	CC	CC	CC	CC	CT	CT	CT	CT
4404786 F 0-21 GT	G/T	NA	NA	NA	Wheat_ChineseSpring10	DarT	TT	GT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT
3936841 F 0-22 GC	G/C	chrUn	361989072	NA	Wheat_ChineseSpring10	DarT	GG	GG	GC	GG	GG	GC	GC	GC	GG	CC	GC	CC	CC	GG
2275611 F 0-25 AG	A/G	chr3B	767917788	NA	Wheat_ChineseSpring10	DarT	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	AA	GG	GG
1007201 F 0-42 AG	A/G	chr1A	10132814	NA	Wheat_ChineseSpring10	DarT	GG	GG	GG	GG	AA	AA	AA	AG	GG	GG	AA	GG	GG	AA
1270155 F 0-38 AG	A/G	chr7A	690860873	NA	Wheat_ChineseSpring10	DarT	AG	AA	AA	AA	GG	AG	GG	NN	AA	AA	AA	AA	GG	AA
4262251 F 0-11 CG	C/G	chr5D	53028568	NA	Wheat_ChineseSpring10	DarT	CG	CG	CG	CG	CG	CG	CG	CC	CC	CC	CG	CG	CG	CC
1016512 F 0-28 GA	G/A	chr1B	19772673	NA	Wheat_ChineseSpring10	DarT	GA	NN	AA	GG	GG	GG	GG	GG	GG	GG	GG	GG	AA	GG
1016148 F 0-25 CG	C/G	chr1D	302264148	NA	Wheat_ChineseSpring10	DarT	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
1019083 F 0-32 GC	G/C	chr5A	674938241	NA	Wheat_ChineseSpring10	DarT	GG	CC	CC	CC	CC	CC	CC	GG	GG	CC	CC	CC	CC	GG
1033127 F 0-8 GC	G/C	NA	NA	NA	Wheat_ChineseSpring10	DarT	GG	NN	GG	CC	CC	GG	GG	GG	GG	GG	CC	CC	GG	CC
3935125 F 0-13 AG	A/G	NA	NA	NA	Wheat_ChineseSpring10	DarT	AA	AA	AA	AA	AA	AA	GG	AA	GG	AA	AG	GG	GG	AA
1049252 F 0-13 GA	G/A	chrUn	42802772	NA	Wheat_ChineseSpring10	DarT	GG	GG	GG	GG	GG	GG	GG	GA	GG	GG	GG	AA	GG	GG
1018953 F 0-6 TC	T/C	chr7D	418916798	NA	Wheat_ChineseSpring10	DarT	CC	CC	TC	TC	CC	CC	CC	CC	CC	TC	TC	CC	TC	TC
1033160 F 0-35 CT	C/T	chr7B	747308026	NA	Wheat_ChineseSpring10	DarT	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
1034464 F 0-58 AG	A/G	chr2D	44609696	NA	Wheat_ChineseSpring10	DarT	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AG	GG	AA	AA
1162266 F 0-7 GA	G/A	chrUn	334142195	NA	Wheat_ChineseSpring10	DarT	GA	GA	GA	GG	GG	GA	GA	GA	GA	GA	GA	GA	GA	GG
4004790 F 0-9 GA	G/A	NA	NA	NA	Wheat_ChineseSpring10	DarT	NN	GG	NN	AA	AA	NN	NN	AA	NN	NN	GG	AA	AA	NN
3026027 F 0-5 AC	A/C	chr5D	380079708	NA	Wheat_ChineseSpring10	DarT	AA	AA	AC	AC	AC	AC	AC	AC	AC	AA	AA	AC	AA	AA
1088758 F 0-60 GC	G/C	NA	NA	NA	Wheat_ChineseSpring10	DarT	GG	CC	GG	GG	GG	GG	GG	GG	GG	NN	NN	GG	GG	GG
1017157 F 0-13 AG	A/G	chr2D	593851647	NA	Wheat_ChineseSpring10	DarT	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	GG	AA	AA
1015856 F 0-37 CG	C/G	chr3B	17613543	NA	Wheat_ChineseSpring10	DarT	CC	NN	NN	NN	CC	CC	CC	CC	NN	CC	GG	GG	NN	NN
1017121 F 0-62 CT	C/T	chrUn	316013593	NA	Wheat_ChineseSpring10	DarT	TT	CC	TT	TT	CC	CC	NN	NN	NN	CC	CC	NN	TT	TT
4261936 F 0-14 TC	T/C	chrUn	384628356	NA	Wheat_ChineseSpring10	DarT	CC	CC	CC	CC	CC	NN	NN	TT	CC	CC	CC	TT	CC	CC
3939551 F 0-15 CG	C/G	chr3B	373696	NA	Wheat_ChineseSpring10	DarT	CC	CG	GG	GG	CC	GG	GG	CG	CC	CC	CG	GG	CC	GG
1041292 F 0-13 CG	C/G	chr1D	416610556	NA	Wheat_ChineseSpring10	DarT	CG	CG	CG	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
5367542 F 0-10 CG	C/G	chr3D	426440396	NA	Wheat_ChineseSpring10	DarT	CC	CC	CC	CC	CC	CC	CC	CG	CC	CC	CC	CC	CC	CC
3953491 F 0-8 CG	C/G	chr3D	560942822	NA	Wheat_ChineseSpring10	DarT	CC	CG	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
4993437 F 0-7 GT	G/T	NA	NA	NA	Wheat_ChineseSpring10	DarT	GG	GG	GG	GG	GG	GG	GG	GG	GT	GT	GT	GG	GG	GT
4003263 F 0-19 CT	C/T	NA	NA	NA	Wheat_ChineseSpring10	DarT	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
3934928 F 0-23 GT	G/T	chr2B	15696431	NA	Wheat_ChineseSpring10	DarT	TT	GG	GG	GG	GG	TT	TT	GT	TT	GT	GT	GG	GG	GG
4261523 F 0-10 CG	C/G	chr3B	3912019	NA	Wheat_ChineseSpring10	DarT	CC	CC	CG	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
1034124 F 0-38 GC	G/C	chr7A	112775432	NA	Wheat_ChineseSpring10	DarT	GG	GG	GG	GC	GC	GG	GG	GG	GG	GG	GG	GG	GG	GG
993225 F 0-27 CT	C/T	chr6B	679285878	NA	Wheat_ChineseSpring10	DarT	CC	TT	CC	CC	CC	TT	TT	CC	CC	CC	CC	CC	CC	CC



Zestawy danych

Gatunek!	D/K!	Record No	P. Light	P. Dur.[sec]	Light	ABS/RC	Dio/RC	Tro/RC	Eto/RC	Rea/RC	ABS/Cso	Dio/Cso	Tro/Cso	Eto/Cso	Reo/Cso	ABS/Csm	Dio/Csm	Tro/Csm	Eto/Csm	Reo/Csm	gamma[RC]	phi[Po]	1-1 psi[Et]	1-1 p Pi abs	dRo[1-1 dRo Pi total	DF abs	DF Total	kP / ABS * 1 kN / A		
Ce	K	1	0	0	1500	1.9698	0.3656	1.8010	0.8975	0.435	211.0000	39.2256	171.7744	96.2978	46.6817	1123.0001	211.0000	924.0000	518.0000	251.0000	0.5085	4.379	1.2759	2.841	0.9401	2.6708	0.4535	0.4286	0.0039	0.0000
Ce	K	2	0	0	1500	1.7058	0.3309	1.3749	0.8117	0.447	245.0000	47.5257	197.4743	116.5835	64.2082	1263.000	245.0000	1018.0000	601.0000	331.0000	0.5862	4.155	1.4412	3.151	1.2259	4.3039	0.5454	0.6339	0.0033	0.0000
Ce	K	3	0	0	1500	1.9977	0.3964	1.6012	0.8581	0.458	231.0000	45.8428	185.1572	99.2268	52.9871	1164.000	231.0000	933.0000	500.0000	267.0000	0.5006	4.039	1.1547	2.335	1.1459	2.6754	0.3682	0.4274	0.0035	0.0000
Ce	K	4	0	0	1500	2.1222	0.6737	1.4685	0.6235	0.179	295.0000	52.7772	202.2228	85.8582	24.2164	956.000	295.0000	643.0000	273.0000	97.0000	0.4608	2.180	0.7378	0.751	0.3929	0.2949	-0.1245	-0.5033	0.0023	0.0010
Ce	K	5	0	0	1500	1.755	0.3500	1.4050	0.8057	0.446	212.0000	42.2603	169.7197	97.3246	53.8476	1063.000	212.0000	851.0000	488.0000	270.0000	0.5099	4.514	1.3444	3.075	1.2385	3.8084	0.4878	0.5907	0.0038	0.0000
Ce	K	6	0	0	1500	1.832	0.3713	1.2448	0.7612	0.439	237.0000	44.4376	193.5625	117.5500	67.8750	1264.000	237.0000	1027.0000	628.0000	362.0000	0.5637	3.333	1.5739	4.455	1.2609	6.0585	0.6485	0.7834	0.0034	0.0000
Ce	K	7	0	0	1500	2.2576	0.7310	1.5266	0.6423	0.244	305.0000	98.7527	206.2473	96.7768	33.0255	1423.000	305.0000	837.0000	268.0000	102.0000	0.4429	2.069	0.7263	0.672	0.6145	0.4128	-0.1727	-0.4824	0.0022	0.0010
Ce	K	8	0	0	1500	1.4918	0.2772	1.2146	0.7859	0.451	236.0000	43.8551	192.1449	124.3181	71.3575	1270.000	236.0000	1034.0000	609.0000	384.0000	0.6703	4.381	1.8329	3.583	1.3474	2.7528	0.7310	0.8805	0.0034	0.0000
Ce	K	9	0	0	1500	1.5134	0.2634	1.2500	0.7694	0.392	228.0000	39.6824	188.3176	116.9145	59.0015	1310.000	228.0000	1082.0000	606.0000	339.0000	0.6068	4.746	1.6010	3.020	1.0367	5.2044	0.7007	0.7164	0.0036	0.0000
Ce	K	10	0	0	1500	1.4651	0.2823	1.1828	0.7569	0.448	226.0000	43.5431	182.4569	116.7570	69.1679	1173.000	226.0000	947.0000	606.0000	359.0000	0.6826	4.190	1.7771	3.503	1.4534	3.7875	0.7061	0.8685	0.0036	0.0000
Ce	K	11	0	0	1500	2.2293	0.6202	1.7072	0.8930	0.419	217.0000	50.8333	166.3667	84.0000	40.8333	930.000	217.0000	713.0000	360.0000	175.0000	0.4486	3.286	1.0198	1.503	0.9459	1.4218	0.1770	0.1528	0.0035	0.0010
Ce	K	12	0	0	1500	1.7956	0.3884	1.4072	0.9504	0.375	284.0000	41.2288	222.5712	150.2375	59.2558	1310.000	284.0000	1029.0000	695.0000	274.0000	0.5569	3.623	2.0808	4.199	0.6508	2.7328	0.6231	0.4366	0.0028	0.0000
Ce	K	13	0	0	1500	1.9756	0.3989	1.5767	0.9136	0.511	234.0000	47.2442	186.7558	108.2174	60.5595	1159.000	234.0000	925.0000	536.0000	300.0000	0.5062	3.953	1.3779	2.757	1.2712	3.5048	0.4405	0.5447	0.0034	0.0000
Ce	K	14	0	0	1500	1.5598	0.3074	1.2522	0.8035	0.465	259.0000	51.0510	207.9490	133.4422	77.2664	1314.000	259.0000	1055.0000	677.0000	392.0000	0.6412	4.073	1.7910	4.678	1.3754	6.4339	0.6700	0.8085	0.0031	0.0000
Ce	K	15	0	0	1500	1.5166	0.2875	1.2281	0.7485	0.432	262.0000	49.7060	212.2940	129.3874	74.7487	1381.000	262.0000	1119.0000	682.0000	394.0000	0.6598	4.271	1.5606	3.498	1.3681	6.0166	0.6432	0.7793	0.0031	0.0000
Ce	K	16	0	0	1500	1.7671	0.3464	1.4208	0.7940	0.454	236.0000	48.2591	189.7409	106.0432	60.5681	1304.000	236.0000	968.0000	541.0000	309.0000	0.5659	4.102	1.2670	2.941	1.3319	3.9168	0.4685	0.6929	0.0034	0.0000
D	15	0	0	0	1500	1.4475	0.2829	1.1647	0.7200	0.383	264.0000	51.5885	212.4115	131.3161	69.7617	1261.000	264.0000	1087.0000	672.0000	357.0000	0.6908	4.117	1.6193	4.606	1.1333	5.2201	0.6633	0.7177	0.0030	0.0000
D	16	0	0	0	1500	1.7562	0.3318	1.4244	0.8289	0.458	253.0000	47.8036	205.1964	119.4145	65.9425	1339.000	253.0000	1086.0000	632.0000	349.0000	0.5694	4.292	1.3921	3.402	1.2332	4.1960	0.5318	0.6228	0.0032	0.0000
D	17	0	0	0	1500	1.6014	0.2984	1.3030	0.8072	0.460	251.0000	46.7713	204.2287	126.5249	72.1136	1347.000	251.0000	1096.0000	679.0000	387.0000	0.6245	4.367	1.6233	4.440	1.3253	5.8845	0.6474	0.7697	0.0032	0.0000
D	18	0	0	0	1500	1.7523	0.3634	1.3889	0.7976	0.442	236.0000	48.9420	187.0580	107.4236	59.5185	1137.000	236.0000	902.0000	518.0000	287.0000	0.5707	3.822	1.3490	2.942	1.2424	3.6566	0.4687	0.5630	0.0034	0.0000
D	19	0	0	0	1500	1.9046	0.4033	1.5013	0.8543	0.453	245.0000	51.8799	193.1201	109.9006	58.2325	1168.000	245.0000	912.0000	519.0000	275.0000	0.5201	3.722	1.3206	2.581	1.1270	2.9090	0.4118	0.4337	0.0032	0.0000
D	20	0	0	0	1500	1.9384	0.4185	1.5199	0.8353	0.456	236.0000	50.9570	185.0430	101.6981	55.4913	1093.000	236.0000	857.0000	471.0000	257.0000	0.5159	3.631	1.2202	2.286	1.2009	2.7452	0.3591	0.4386	0.0033	0.0000
D	21	0	0	0	1500	INF	INF	INF	INF	INF	3840.0000	3840.0000	0.0000	0.0000	0.0000	3840.000	3840.0000	0.0000	0.0000	0.0000	0.0000	INF	INF	INF	INF	INF	INF	INF	0.0000	0.0000
D	22	0	0	0	1500	1.5267	0.2895	1.2371	0.7508	0.441	231.0000	43.8103	187.1897	113.8534	66.7586	1218.000	231.0000	987.0000	599.0000	352.0000	0.6550	4.273	1.5438	4.321	1.4251	6.1575	0.6356	0.7894	0.0035	0.0000
D	23	0	0	0	1500	1.5590	0.2972	1.2624	0.7599	0.429	238.0000	46.5510	193.4490	116.4508	65.5630	1254.000	238.0000	1015.0000	611.0000	344.0000	0.6412	4.247	1.5124	4.118	1.2884	5.3058	0.6147	0.7840	0.0034	0.0000
D	24	0	0	0	1500	2.0807	0.4547	1.6260	1.0023	0.330	259.0000	58.7233	212.2177	125.6033	42.6117	1231.000	259.0000	962.0000	593.0000	195.0000	0.4800	4.573	1.6079	2.763	0.4959	1.3533	0.4412	0.1314	0.0275	0.0000
D	25	0	0	0	1500	1.6085	0.3130	1.2955	0.7946	0.470	275.0000	53.5209	221.4791	136.9457	80.3786	1413.000	275.0000	1138.0000	698.0000	413.0000	0.6217	4.138	1.5864	4.081	1.4491	5.9142	0.6108	0.7719	0.0029	0.0000
D	26	0	0	0	1500	1.8803	0.4701	1.4103	0.7753	0.317	77.0000	19.2500	57.7500	31.7500	13.0000	30.000	77.0000	231.0000	127.0000	62.0000	0.5318	3.000	1.2212	1.948	0.6933	1.3508	0.2897	0.1306	0.0097	0.0030
D	27	0	0	0	1500	2.6216	0.9238	1.6979	-0.0584	-0.983	253.0000	89.1490	163.8510	-6.5379	-95.8440	1310.000	253.0000	465.0000	-16.0000	-272.0000	0.3814	1.838	-0.0333	-0.023	-1.0625	0.0248	INF	-1.6059	0.0026	0.0010
D	28	0	0	0	1500	1.6153	0.2993	1.3016	0.8136	0.445	252.0000	46.6941	205.3059	126.9255	69.4853	1360.000	252.0000	1108.0000	685.0000	376.0000	0.6191	4.397	1.6194	4.408	1.2097	5.3322	0.6442	0.7229	0.0032	0.0000
D	29	0	0	0	1500	1.8	0.3650	1.4350	0.8136	0.454	262.0000	53.1300	208.8700	118.4272	66.1084	1292.000	262.0000	1030.0000	584.0000	326.0000	0.5556	3.931	1.3094	2.860	1.2636	3.6136	0.4563	0.5978	0.0030	0.0000
D	30	0	0	0	1500	2.0518	0.4107	1.6411	0.8782	0.462	246.0000	49.2400	196.7600	105.2856	55.4451	1229.000	246.0000	983.0000	526.0000	277.0000	0.4784	3.996	1.5150	2.242	1.1124	2.4936	0.3505	0.9698	0.0033	0.0000
K	1	0	0	0	1500	1.6477	0.3001	1.3477	0.8124	0.436	287.0000	52.2646	234.7354	141.4968	75.9385	1576.000	287.0000	1289.0000	777.0000	417.0000	0.6089	4.491	1.5176	4.137	1.1563	4.7915	0.6166	0.8005	0.0028	0.0000
K	2	0	0	0	1500	1.7594	0.3653	1.4031	0.7980	0.465	242.0000	49.0075	192.9625	109.7607	63.9933	1195.000	242.0000	953.0000	542.0000	316.0000	0.5864	3.938	1.3187	2.952	1.3982	4.1272	0.4701	0.6157	0.0033	0.0000
K	3	0	0	0	1500	2.2784	0.5145	1.7639	0.9671	0.523	298.0000	67.5234	231.4766	126.9169	65.8526	1324.000	298.0000	1025.0000	562.0000	304.0000	0.4389	3.428	1.2138	1.826	1.1783	2.1520	0.2616	0.3328	0.0026	0.0000
K	4	0	0	0	1500	1.7148	0.3073	1.4076	0.8111	0.403	222.0000	39.7772	182.2228	104.9976	52.1404	1239.000	222.0000	1017.0000	586.0000	291.0000	0.5831	4.581	1.3596	3.622	0.9804	3.5829	0.5602	0.5442	0.0037	0.0000
K	5	0	0	0	1500	1.7312	0.3478	1.3834	0.7888	0.423	265.0000	53.2411	211.7																	

Wielowymiarowe eksperymenty przeprowadzane są:

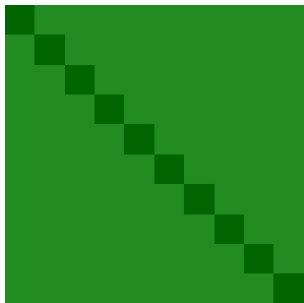
- w biologii
- w genetyce
- w rolnictwie
- w obszarach biomedycznych
- w medycynie
- w obszarach związanych z ochroną środowiska
- w naukach inżynierskich
- w naukach ekonomicznych

Identyfikacja macierzy kowariancji:

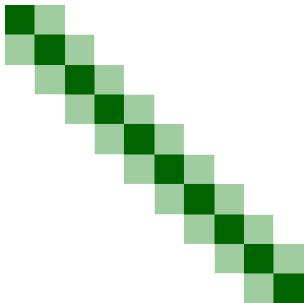
- pozwala poszerzyć wiedzę o zachowaniu zmiennych
- umożliwia analizę danych przy użyciu precyzyjnego modelu statystycznego o mniejszej liczbie parametrów kowariancyjnych, pozostawiając więcej swobody na estymację interesujących parametrów modelu
- zmniejszenie liczby parametrów umożliwia wnioskowanie statystyczne w sytuacji, gdy pojawia się problem wysokiej wielowymiarowości (ang. *high-dimensionality*)
- znajomość zależności kowariancyjnych jest istotna w wielu modelach analizy statystycznej, między innymi w analizie składowych głównych, liniowej i kwadratowej analizie dyskryminacyjnej, a analizie regresji czy analizie niezależności zmiennych

Struktury kowariancyjne

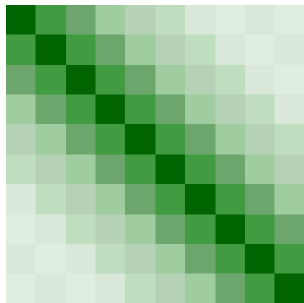
CS



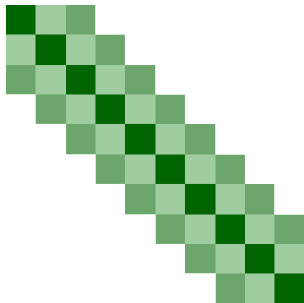
T1



AR

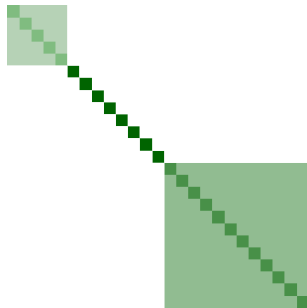


T2

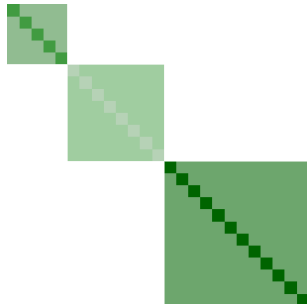


Struktury kowariancyjne

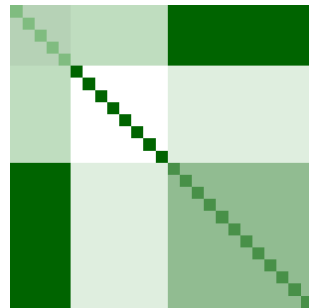
blokowo – diagonalna



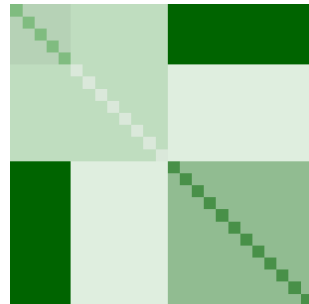
blokowo – diagonalna



blokowa

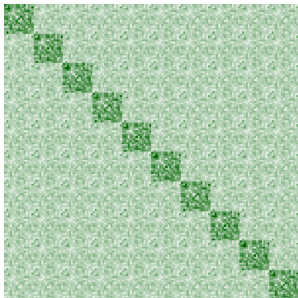


blokowa

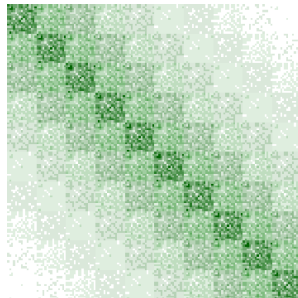


Struktury kowariancyjne

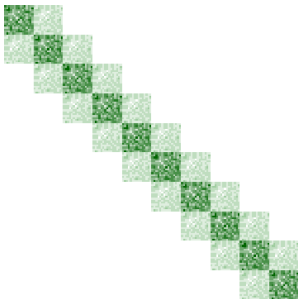
$CS \otimes UN$



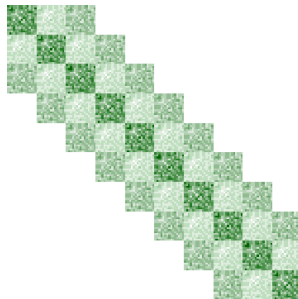
$AR \otimes UN$



$T1 \otimes UN$



$T2 \otimes UN$



compound symmetry (CS)



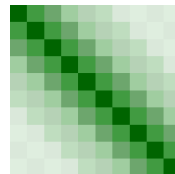
$$\Gamma_{CS} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \dots & 1 \end{pmatrix} = \sigma^2 (\rho \mathbf{J}_m + (1 - \rho) \mathbf{I}_m)$$

Γ_{CS} jest określona dodatnio:

- $\sigma^2 > 0$
- $\rho \in (-\frac{1}{m-1}; 1)$

Macierz autoregresji pierwszego rzędu

autoregression of order one (AR(1))

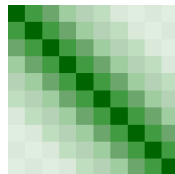


$$\mathbf{\Gamma}_{AR} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{m-1} \\ \rho & 1 & \rho & \dots & \rho^{m-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \dots & 1 \end{pmatrix} = \sigma^2 \sum_{i=0}^{m-1} \rho^i \mathbf{H}_i$$

\mathbf{H}_i - symetryczna macierz stopnia m z elementami na i -tych przekątnych równe 1, pozostałe równe 0

$$\mathbf{H}_0 = \mathbf{I}_m$$

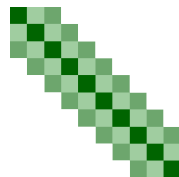
autoregression of order one (AR(1))



$$\Gamma_{AR} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{m-1} \\ \rho & 1 & \rho & \dots & \rho^{m-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \dots & 1 \end{pmatrix} = \sigma^2 \sum_{i=0}^{m-1} \rho^i \mathbf{H}_i$$

Γ_{AR} jest określona dodatnio:

- $\sigma^2 > 0$
- $\rho \in (-1; 1)$



banded symmetric Toeplitz structure (\mathbf{T}_p , $p < m$)

$$\mathbf{\Gamma}_{T_p} = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \dots & \rho_p & 0 & \dots & 0 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_p & \ddots & \vdots \\ \vdots & \rho_1 & 1 & \rho_1 & & \ddots & 0 \\ \rho_p & & \ddots & \ddots & \ddots & & \rho_p \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \rho_1 \\ 0 & \dots & 0 & \rho_p & \dots & \rho_1 & 1 \end{pmatrix} = \sigma^2 (\mathbf{I}_m + \sum_{i=1}^p \rho_i \mathbf{H}_i)$$

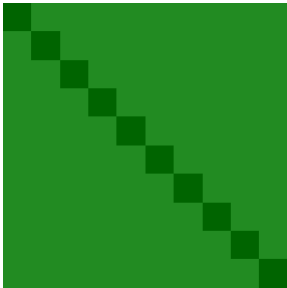
$\mathbf{\Gamma}_{T_1}$ jest określona dodatkowo:

- $\sigma^2 > 0$
- $\rho_1 \in \left(-\frac{1}{2\cos\frac{\pi}{m+1}}; \frac{1}{2\cos\frac{\pi}{m+1}} \right)$

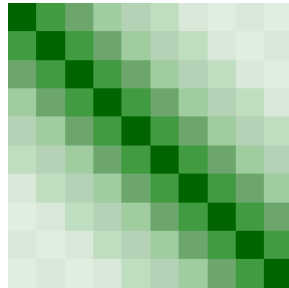
Wstęgowa macierz Toeplitza

$$p = m - 1$$

CS



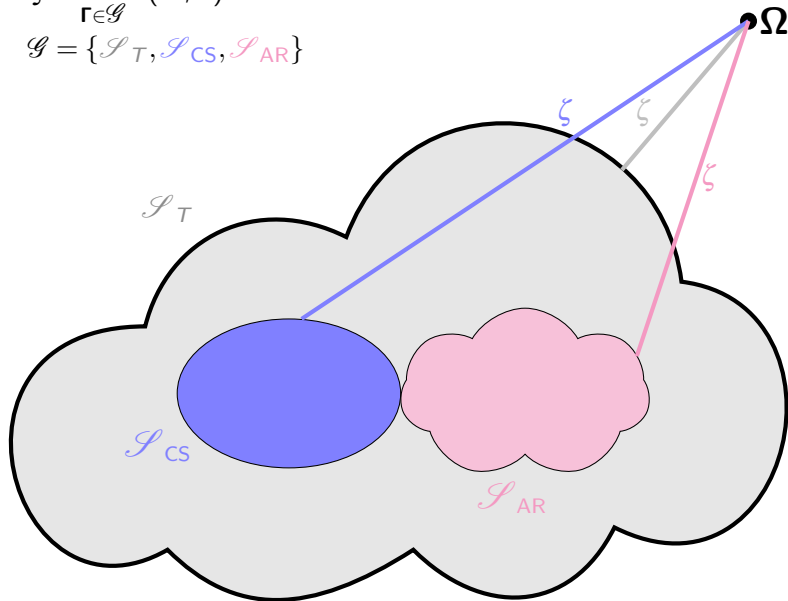
AR



Identyfikacja struktury

$$\zeta = \min_{\Gamma \in \mathcal{G}} d(\Omega; \Gamma)$$

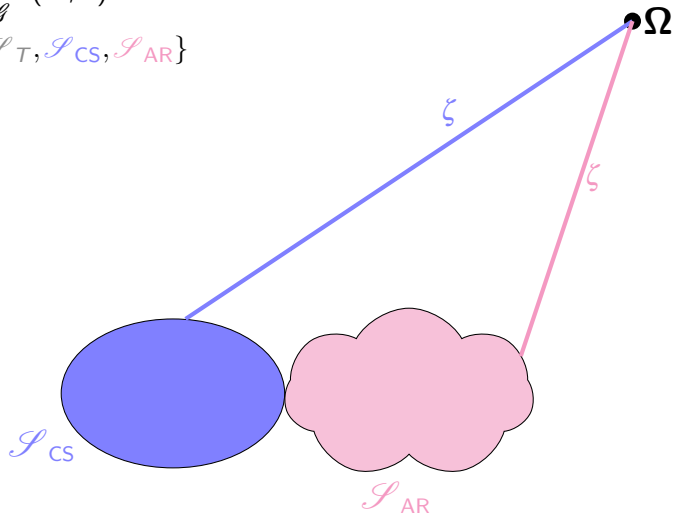
$$\mathcal{G} = \{\mathcal{S}_T, \mathcal{S}_{CS}, \mathcal{S}_{AR}\}$$



Identyfikacja struktury

$$\zeta = \min_{\Gamma \in \mathcal{G}} d(\Omega; \Gamma)$$

$$\mathcal{G} = \{\mathcal{S}_T, \mathcal{S}_{CS}, \mathcal{S}_{AR}\}$$



Funkcje rozbieżności:

- norma Frobeniusa

$$f_F(\mathbf{\Omega}, \mathbf{\Gamma}) = \|\mathbf{\Omega} - \mathbf{\Gamma}\|_F^2 = \text{tr}[(\mathbf{\Omega} - \mathbf{\Gamma})(\mathbf{\Omega} - \mathbf{\Gamma})'] \quad (1)$$

- entropijna funkcja straty (cf. Stein (1956), Dey and Srinivasan (1985), James and Stein(1961), Lin et al.(2014))

$$f_E(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr}[\mathbf{\Omega}^{-1}\mathbf{\Gamma}] - \ln|\mathbf{\Omega}^{-1}\mathbf{\Gamma}| - pq \quad (2)$$

- kwadratowa funkcja straty (cf. Anderson(2003))

$$f_Q(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr}\left\{\left[\mathbf{\Omega}^{-1}\mathbf{\Gamma} - \mathbf{I}_{pq}\right]^2\right\} \quad (3)$$

Dla zadanej macierzy $\mathbf{\Omega}$ i $k \in \{F, E, Q\}$ oznaczmy $f_k(\mathbf{\Omega}, \mathbf{\Gamma})$ jako $L_{\Omega}^{(k)}(\mathbf{\Gamma})$

Problem

$$\zeta^{(k)} = \min_{\mathbf{\Gamma} \in \mathcal{S}} L_{\Omega}^{(k)}(\mathbf{\Gamma}) = L_{\Omega}^{(k)}(\hat{\mathbf{\Gamma}})$$

\mathcal{S} - zbiór rozważanych struktur

Problem

$$\min_{\Gamma \in \mathcal{S}} L_{\Omega}^{(k)}(\Gamma), \quad k \in \{F, E, Q\}$$

$$\min f(x) \quad \frac{df}{dx} = 0$$

$$\min L_{\Omega}^{(k)}(\Gamma) \quad \frac{dL}{d\Gamma} = \mathbf{0}$$

Macierz duplikacji

- macierz kowariancji stopnia m
 - jest symetryczna ($\mathbf{A} = \mathbf{A}'$)
 - $m(m+1)/2$ niewiadomych
- powtarzające się elementy można wyeliminować
 - macierz duplikacji \mathbf{D}_m
 - $\mathbf{D}_m : m^2 \times \frac{1}{2}m(m+1)$
 - $\mathbf{D}_m \cdot \text{vech } \mathbf{A} = \text{vec } \mathbf{A}$

$\mathbf{A} : m \times n$

- $\text{vec}\mathbf{A}$ - układa kolumny macierzy \mathbf{A} jedna pod drugą

$$\mathbf{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \Rightarrow \text{vec}\mathbf{A} = \begin{pmatrix} a \\ d \\ g \\ b \\ e \\ h \\ c \\ f \\ i \end{pmatrix}$$

$\mathbf{A} \in \text{Sym}(m)$

- $\text{vech}\mathbf{A}$ - układa kolumny macierzy \mathbf{A} w wektor bez górnego trójkąta

$$\mathbf{A} = \begin{pmatrix} a & b & c \\ b & e & f \\ c & f & i \end{pmatrix} \Rightarrow \text{vech}\mathbf{A} = \begin{pmatrix} a \\ b \\ c \\ e \\ f \\ i \end{pmatrix}$$

- $\mathbf{D}_m : m^2 \times \frac{1}{2}m(m+1)$ - duplication matrix: $\mathbf{D}_m \text{vech}\mathbf{A} = \text{vec}\mathbf{A}$

Lemat

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

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

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

Fackler(2005)

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

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

Reguła łańcuchowa

$$f_Q(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr} \left\{ \left[\mathbf{\Omega}^{-1} \mathbf{\Gamma} - \mathbf{I}_m \right]^2 \right\}$$

różniczkujemy względem $\mathbf{\Gamma}$

$$\mathbf{Z} = \mathbf{\Omega}^{-1} \mathbf{\Gamma} - \mathbf{I}_m$$

Stosując reguły różniczkowania dla macierzy symetrycznych

$$\frac{\partial f_Q}{\partial \mathbf{\Gamma}} = \frac{\partial f_Q}{\partial \mathbf{Z}^2} \cdot \frac{\partial \mathbf{Z}^2}{\partial \mathbf{Z}} \cdot \frac{\partial \mathbf{Z}}{\partial \mathbf{\Gamma}}$$

Funkcje rozbieżności: norma Frobeniusa

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

$$\frac{\partial f_F}{\partial \mathbf{\Gamma}} =$$

Funkcje rozbieżności: entropijna funkcja straty

$$f_E(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr} \left[\mathbf{\Omega}^{-1} \mathbf{\Gamma} \right] - \ln \left| \mathbf{\Omega}^{-1} \mathbf{\Gamma} \right| - m$$

$$\frac{\partial f_E}{\partial \mathbf{\Gamma}} =$$

Funkcje rozbieżności: kwadratowa funkcja straty

$$f_Q(\mathbf{\Omega}, \mathbf{\Gamma}) = \text{tr} \left\{ \left[\mathbf{\Omega}^{-1} \mathbf{\Gamma} - \mathbf{I}_m \right]^2 \right\}$$

$$\frac{\partial f_Q}{\partial \mathbf{\Gamma}} =$$

Struktura: $\mathbf{\Gamma} = \mathbf{\Gamma}(\sigma^2, \rho)$

$$\frac{\partial f}{\partial \sigma^2} =$$

$$\frac{\partial f}{\partial \rho} =$$

$$\begin{cases} \frac{\partial f}{\partial \sigma^2} = 0 \\ \frac{\partial f}{\partial \rho} = 0 \end{cases}$$

Funkcje rozbieżności: norma Frobeniusa

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

$$\mathbf{\Gamma}_{CS} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \dots & 1 \end{pmatrix} = \sigma^2 (\rho \mathbf{J}_m + (1 - \rho) \mathbf{I}_m)$$

$$\frac{\partial f}{\partial \sigma^2} =$$

$$\frac{\partial f}{\partial \rho} =$$

Funkcje rozbieżności: entropijna funkcja straty

$$f_E(\mathbf{\Omega}, \mathbf{\Gamma}_{CS}) = \text{tr} \left[\mathbf{\Omega}^{-1} \mathbf{\Gamma}_{CS} \right] - \ln \left| \mathbf{\Omega}^{-1} \mathbf{\Gamma}_{CS} \right| - m$$

$$\mathbf{\Gamma}_{CS} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \dots & 1 \end{pmatrix} = \sigma^2 (\rho \mathbf{J}_m + (1 - \rho) \mathbf{I}_m)$$

$$\frac{\partial f}{\partial \sigma^2} =$$

$$\frac{\partial f}{\partial \rho} =$$

Funkcje rozbieżności: kwadratowa funkcja straty

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

$$\mathbf{\Gamma}_{CS} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \dots & 1 \end{pmatrix} = \sigma^2 (\rho \mathbf{J}_m + (1 - \rho) \mathbf{I}_m)$$

$$\frac{\partial f}{\partial \sigma^2} =$$

$$\frac{\partial f}{\partial \rho} =$$

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{CS}}, \quad \mathcal{S}_{\text{CS}} = \{\mathbf{\Gamma}_{\text{CS}} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie [Cui et al. (2016)]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna, określona dodatnio macierz $\mathbf{\Gamma}_{\text{CS}}$, która minimalizuje normę Frobeniusa (1) w zbiorze \mathcal{S}_{CS} i to minimum jest osiągnięte dla ρ , σ^2 spełniających następujący układ równań

$$\begin{cases} \rho &= \frac{\delta}{(m-1)\text{tr}(\mathbf{\Omega})} \\ \sigma^2 &= \frac{\text{tr}(\mathbf{\Omega}) + \rho\delta}{m + m(m-1)\rho^2} \end{cases}$$

gdzie $\delta = \text{tr}[\mathbf{\Omega}(\mathbf{J}_m - \mathbf{I}_m)]$.

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{AR}}, \quad \mathcal{S}_{\text{AR}} = \{\mathbf{\Gamma}_{\text{AR}} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie [Cui et al. (2016)]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna, określona dodatnio macierz $\mathbf{\Gamma}_{\text{AR}}$, która minimalizuje normę Frobeniusa (1) w zbiorze \mathcal{S}_{AR} i to minimum jest osiągnięte dla ρ , σ^2 spełniających następujący układ równań

$$\left\{ \begin{array}{l} -\sum_{i=1}^{m-1} i \rho^{i-1} \text{tr}(\mathbf{\Omega} \mathbf{H}_i) + \frac{2 \sum_{i=0}^{m-1} \rho^i \text{tr}(\mathbf{\Omega} \mathbf{H}_i) \sum_{i=1}^{m-1} (m-i) i \rho^{2i-1}}{m + 2 \sum_{i=1}^{m-1} (m-i) \rho^{2i}} = 0 \\ \sigma^2 = \frac{\sum_{i=0}^{m-1} \rho^i \text{tr}(\mathbf{\Omega} \mathbf{H}_i)}{m + 2 \sum_{i=1}^{m-1} (m-i) \rho^{2i}}. \end{array} \right.$$

Aproksymacja za pomocą normy Frobeniusa

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{T_1}, \quad \mathcal{S}_{T_1} = \{\mathbf{\Gamma}_{T_1} \in \mathbb{R}_m^{\geq}\}$$

Twierdzenie [Cui et al. (2016)]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Gamma}_{T_1}$, która minimalizuje normę Frobeniusa (1) w zbiorze \mathcal{S}_{T_1} i to minimum jest osiągnięte dla ρ , σ^2 spełniających następujący układ równań

$$\begin{cases} \sigma^2 &= \frac{\text{tr}(\mathbf{\Omega})}{m} \\ \rho_1 &= \frac{m \text{tr}(\mathbf{\Omega} \mathbf{H}_1)}{2(m-1) \text{tr}(\mathbf{\Omega})} \end{cases}$$

Filipiak et al. (2018)

$$\hat{\mathbf{\Gamma}}_{T_1} = t \sigma^2 \left(\mathbf{I}_m + \frac{1}{2 \cos \frac{\pi}{m+1}} \mathbf{H}_1 \right) \quad \text{dla } \rho_1 > 0, \quad \hat{\mathbf{\Gamma}}_{T_1} = t \sigma^2 \left(\mathbf{I}_m - \frac{1}{2 \cos \frac{\pi}{m+1}} \mathbf{H}_1 \right) \quad \text{dla } \rho_1 < 0$$

$$\text{gdzie } t = \left(m + \frac{\rho_1(m-1)}{\cos \frac{\pi}{m+1}} \right) / \left(m + \frac{m-1}{2(\cos \frac{\pi}{m+1})^2} \right)$$



$$\mathbf{\Gamma} = \mathbf{\Gamma}_{T_p}, \quad p > 1, \quad \mathcal{S}_{T_p} = \{\mathbf{\Gamma}_{T_p} \in \mathbb{R}_m^{>}\}$$

- $\hat{\mathbf{\Gamma}}_{T_p}$ nie jest dany w postaci jawnej
- algorytm - Filipiak et al. (2018)

ON PROJECTION OF A POSITIVE DEFINITE MATRIX ON A CONE OF NONNEGATIVE DEFINITE TOEPLITZ MATRICES*

KATARZYNA FILIPIAK[†], AUGUSTYN MARKIEWICZ[‡], ADAM MIELDZIOC[‡], AND ANETA SAWIKOWSKA^{‡§}

Abstract. We consider approximation of a given positive definite matrix by nonnegative definite banded Toeplitz matrices. We show that the projection on linear space of Toeplitz matrices does not always preserve nonnegative definiteness. Therefore we characterize a convex cone of nonnegative definite banded Toeplitz matrices which depends on the matrix dimensions, and we show that the condition of positive definiteness given by Parter [*Numer. Math.* 4, 293–295, 1962] characterizes the asymptotic cone. In this paper we give methodology and numerical algorithm of the projection basing on the properties of a cone of nonnegative definite Toeplitz matrices.

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{CS}}, \quad \mathcal{S}_{\text{CS}} = \{\mathbf{\Gamma}_{\text{CS}} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie [Lin et al. (2014)]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna, określona dodatnio macierz $\mathbf{\Gamma}_{\text{CS}}$, która minimalizuje entropijną funkcję straty (86) w zbiorze \mathcal{S}_{CS} i to minimum jest osiągnięte dla ρ , σ^2 spełniających następujący układ równań

$$\begin{cases} \rho &= -\delta / [(m-1)\text{tr}(\mathbf{\Omega}^{-1}) + (m-2)\delta] \\ \sigma^2 &= m / [\text{tr}(\mathbf{\Omega}^{-1}) + \rho\delta] \end{cases}$$

gdzie $\delta = \text{tr}[\mathbf{\Omega}^{-1}(\mathbf{J}_m - \mathbf{I}_m)]$.

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{AR}}, \quad \mathcal{S}_{\text{AR}} = \{\mathbf{\Gamma}_{\text{AR}} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie [Lin et al. (2014)]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna, określona dodatnio macierz $\mathbf{\Gamma}_{\text{AR}}$, która minimalizuje entropijną funkcję straty (86) w zbiorze \mathcal{S}_{AR} i to minimum jest osiągnięte dla ρ , σ^2 spełniających następujący układ równań

$$\begin{cases} \frac{m \sum_{i=1}^{m-1} i \rho^{i-1} \text{tr}(\mathbf{\Omega}^{-1} \mathbf{H}_i)}{\sum_{i=0}^{m-1} \rho^i \text{tr}(\mathbf{\Omega}^{-1} \mathbf{H}_i)} + \frac{2(m-1)\rho}{1-\rho^2} = 0 \\ \sigma^2 = m / \sum_{i=0}^{m-1} \rho^i \text{tr}(\mathbf{\Omega}^{-1} \mathbf{H}_i) \end{cases}$$

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{T_1}, \quad \mathcal{S}_{T_1} = \{\mathbf{\Gamma}_{T_1} \in \mathbb{R}_m^{\geq}\}$$

Twierdzenie [Lin et al. (2014)]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Gamma}_{T_1}$, która minimalizuje entropijną funkcję straty (86) w zbiorze \mathcal{S}_{T_1} i to minimum jest osiągnięte dla ρ , σ^2 spełniających następujący układ równań

$$\begin{cases} \sigma^2 = \sum_{j=1}^m \frac{2s_j}{1+2\rho_1 s_j} / \text{tr}(\mathbf{\Omega}^{-1} \mathbf{H}_1) \\ \sum_{j=1}^m \frac{2s_j}{1+2\rho_1 s_j} - \frac{m \text{tr}(\mathbf{\Omega}^{-1} \mathbf{H}_1)}{\text{tr}(\mathbf{s}^{-1}) + \rho_1 \text{tr}(\mathbf{\Omega}^{-1} \mathbf{H}_1)} = 0 \end{cases}$$

gdzie $s_j = \cos(\pi j / (m+1))$.



$$\Gamma = \Gamma_{T_p}, \quad p > 1, \quad \mathcal{S}_{T_p} = \{\Gamma_{T_p} \in \mathbb{R}_m^{\geq}\}$$

- $\hat{\Gamma}_{T_p}$ nie jest dany w postaci jawnej
- algorytm - Lin et al. (2018)



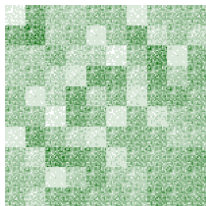
Covariance structure regularization via entropy loss function

Lijing Lin, Nicholas J. Higham, Jianxin Pan*

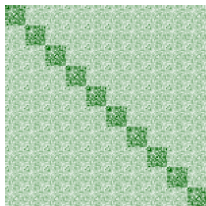
School of Mathematics, University of Manchester, Manchester M13 9PL, UK



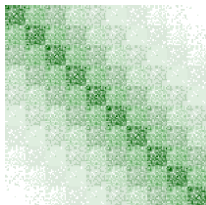
Struktura iloczynu Kroneckera



$$\Psi \otimes \Sigma$$



$$\Psi_{CS} \otimes \Sigma$$



$$\Psi_{AR} \otimes \Sigma$$

Struktura iloczynu Kroneckera

$$\mathcal{I}_{\otimes} = \{ \Psi \otimes \Sigma \in \mathbb{R}_{pq}^+ : \Psi \in \mathbb{R}_p^+, \Sigma \in \mathbb{R}_q^+ \}$$

Struktura iloczynu Kroneckera

$$\mathcal{S}_{\otimes} = \{\Psi \otimes \Sigma \in \mathbb{R}_{pq}^> : \Psi \in \mathbb{R}_p^>, \Sigma \in \mathbb{R}_q^>\}$$

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

$$\mathcal{S}_{\text{AR}\otimes} = \{\Psi_{\text{AR}} \otimes \Sigma \in \mathbb{R}_{pq}^> : \Psi_{\text{AR}} \in \mathbb{R}_p^>, \Sigma \in \mathbb{R}_q^>\}$$

gdzie

$$\Psi_{\text{CS}} = (1 - \rho)\mathbf{I}_p + \rho\mathbf{1}_p\mathbf{1}_p'$$

$$\Psi_{\text{AR}} = \mathbf{I}_p + \sum_{i=1}^{p-1} \rho^i \mathbf{H}_i$$

dla $\mathbf{H} = (h_{ij})$, $h_{ij} = 1$ dla $|j - i| = 1$ i 0 w przeciwnym wypadku, $1 \leq i, j \leq p$.

$$\rho_{\text{CS}} \in \left(-\frac{1}{p-1}, 1\right) \quad \rho_{\text{AR}} \in (-1, 1)$$



Twierdzenie [Van Loan i Pitsianis, 1992]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi} \otimes \mathbf{\Sigma}$, która minimalizuje normę Frobeniusa (1) w zbiorze \mathcal{S}_{\otimes} oraz minimum to jest osiągnięte dla $\mathbf{\Psi}, \mathbf{\Sigma}$ spełniających następujący układ równań

$$\begin{cases} \text{vec } \mathbf{\Psi} = \sqrt{\delta_1} \mathbf{u}_1 \\ \text{vec } \mathbf{\Sigma} = \sqrt{\delta_1} \mathbf{v}_1 \end{cases}$$

gdzie δ_1 jest największą wartością własną oraz \mathbf{u}_1 i \mathbf{v}_1 są odpowiednio lewym i prawym wektorem własnym $p^2 \times q^2$ -wymiarowej macierzy

$$R(\mathbf{\Omega}) = (\text{vec } \mathbf{\Omega}_{11}, \text{vec } \mathbf{\Omega}_{21}, \dots, \text{vec } \mathbf{\Omega}_{p1}, \dots, \text{vec } \mathbf{\Omega}_{pp})'.$$



Twierdzenie [Filipiak i Klein, 2018]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi}_{CS \otimes} \mathbf{\Sigma}$, która minimalizuje normę Frobeniusa (1) w zbiorze $\mathcal{S}_{CS \otimes}$ oraz minimum to jest osiągnięte dla $\mathbf{\rho}, \mathbf{\Sigma}$ spełniających następujący układ równań

$$\begin{cases} \mathbf{\rho} &= \frac{a - (p-1)c + \sqrt{[(p-1)c - a]^2 + 4(p-1)b^2}}{2(p-1)b} \\ \mathbf{\Sigma} &= \frac{1}{1 + (p-1)\mathbf{\rho}^2} \left[\mathbf{\rho} \tilde{\mathbf{\Omega}}_{11} + (1 - \mathbf{\rho}) \mathbf{H} \right] \end{cases}$$

gdzie

$\mathbf{\Psi}_{CS} = \mathbf{U} \mathbf{G} \mathbf{U}'$, $\mathbf{G} = \text{diag}(1 + (p-1)\mathbf{\rho}, 1 - \mathbf{\rho}, \dots, 1 - \mathbf{\rho})$, \mathbf{U} – orthogonal matrix

$$\tilde{\mathbf{\Omega}} = (\mathbf{U}' \otimes \mathbf{I}_q) \mathbf{\Omega} (\mathbf{U} \otimes \mathbf{I}_q) = (\tilde{\Omega}_{ij})_{1 \leq i, j \leq p} \quad \forall_{i,j} : \tilde{\Omega}_{i,j} \in \mathbb{R}^q$$

$$\mathbf{H} = \frac{1}{p} \sum_{i=1}^p \tilde{\Omega}_{ii}$$

$$a = \text{tr} \left[(\tilde{\Omega}_{11} - \mathbf{H}) (\tilde{\Omega}_{11} - \mathbf{H})' \right], \quad b = \text{tr} \left[(\tilde{\Omega}_{11} - \mathbf{H}) \mathbf{H}' \right], \quad c = \text{tr} [\mathbf{H} \mathbf{H}']$$



Twierdzenie [Filipiak i Klein, 2018]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje symetryczna macierz $\Psi_{AR} \otimes \Sigma$, która minimalizuje normę Frobeniusa (1) w zbiorze $\mathcal{S}_{AR \otimes}$ oraz minimum to jest osiągnięte dla ρ, Σ spełniających następujący układ równań

$$\begin{cases} \text{tr } \Psi_{AR}^2 \cdot \text{tr}[(F \otimes \Sigma'_1)\Omega] - \text{tr}(F\Psi_{AR}) \cdot \text{tr}(\Sigma_1 \Sigma'_1) = 0 \\ \Sigma = [\text{tr}(\Psi_{AR}^2)]^{-1} \Sigma_1 \end{cases}$$

gdzie

$$\begin{aligned} \Sigma_1 &= \text{BTr}_q[(\Psi_{AR} \otimes I_q)\Omega] \\ F &= \sum_{i=1}^{p-1} i \rho^{i-1} H_i \end{aligned}$$

$$\mathbf{A} = (\mathbf{A}_{ij}) : mn \times mn$$

$$\mathbf{A}_{ij} : m \times m$$

Operator śladów częściowych

$$\text{PTr}_m \mathbf{A} = (\text{tr} \mathbf{A}_{ij})_{1 \leq i, j \leq n} : n \times n$$

Operator śladu blokowego

$$\text{BTr}_m \mathbf{A} = \sum_{i=1}^n \mathbf{A}_{ii} : m \times m$$



Twierdzenie [FKMM, 2021]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje symetryczna, określona dodatnio macierz $\Psi \otimes \Sigma$, która minimalizuje entropijną funkcję straty (86) w zbiorze \mathcal{S}_{\otimes} i to minimum jest osiągnięte dla Ψ, Σ spełniających następujący układ równań

$$\begin{cases} \Psi^{-1} = \frac{1}{q} \text{PTr}_q [(I_p \otimes \Sigma) \Omega^{-1}] \\ \Sigma^{-1} = \frac{1}{p} \text{BTr}_q [(\Psi \otimes I_q) \Omega^{-1}]. \end{cases}$$

Twierdzenie [FKM, 2022]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje symetryczna macierz $\Psi \otimes \Sigma$, która minimalizuje kwadratową funkcję straty (86) w zbiorze \mathcal{S}_{\otimes} oraz minimum to jest osiągnięte dla Ψ, Σ spełniających następujący układ równań

$$\begin{cases} \text{PTr}_q [(I_p \otimes \Sigma) \Omega^{-1} (\Psi \otimes \Sigma) \Omega^{-1}] &= \text{PTr}_q [(I_p \otimes \Sigma) \Omega^{-1}] \\ \text{BTr}_q [(\Psi \otimes I_q) \Omega^{-1} (\Psi \otimes \Sigma) \Omega^{-1}] &= \text{BTr}_q [(\Psi \otimes I_q) \Omega^{-1}]. \end{cases}$$



Twierdzenie [FKM, 2018]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje wyznaczona jednoznacznie symetryczna, określona dodatnio macierz $\Psi_{CS} \otimes \Sigma$, która minimalizuje entropijną funkcję straty (86) w zbiorze $\mathcal{S}_{CS\otimes}$, przy czym minimum to jest osiągnięte dla ρ, Σ spełniających następujący układ równań

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

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

Twierdzenie [FKM, 2022]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje symetryczna macierz $\Psi_{CS} \otimes \Sigma$, która minimalizuje kwadratową funkcję straty (86) w zbiorze $\mathcal{S}_{CS\otimes}$ oraz minimum to jest osiągnięte dla Ψ_{CS}, Σ spełniających następujący układ równań

$$\begin{cases} \text{tr}[\{(\mathbf{1}_p\mathbf{1}_p' - \mathbf{I}_p) \otimes \Sigma\}\Omega^{-1}(\Psi_{CS} \otimes \Sigma)\Omega^{-1}] &= \text{tr}[\{(\mathbf{1}_p\mathbf{1}_p' - \mathbf{I}_p) \otimes \Sigma\}\Omega^{-1}] \\ \text{BTr}_q[(\Psi_{CS} \otimes \mathbf{I}_q)\Omega^{-1}(\Psi_{CS} \otimes \Sigma)\Omega^{-1}] &= \text{BTr}_q[(\Psi_{CS} \otimes \mathbf{I}_q)\Omega^{-1}]. \end{cases}$$



Twierdzenie [FKMM, 2021]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje symetryczna, określona dodatnio macierz $\Psi_{AR} \otimes \Sigma$ która minimalizuje entropijną funkcję straty postaci (86) w zbiorze $\mathcal{S}_{AR\otimes}$, przy czym minimum to jest osiągnięte dla ρ, Σ spełniających następujący układ równań

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

gdzie $\mathbf{F} = \sum_{i=1}^{p-1} i\rho^{i-1}\mathbf{H}_i$.

Twierdzenie [FKM, 2022]

Dla danej symetrycznej, określonej dodatnio macierzy Ω , istnieje symetryczna macierz $\Psi_{AR} \otimes \Sigma$, która minimalizuje kwadratową funkcję straty postaci (86) w zbiorze $\mathcal{S}_{AR\otimes}$ oraz minimum to jest osiągnięte dla Ψ_{AR}, Σ spełniających następujący układ równań

$$\begin{cases} \text{tr}[(\mathbf{F} \otimes \Sigma)\Omega^{-1}(\Psi_{AR} \otimes \Sigma)\Omega^{-1}] & = \text{tr}[(\mathbf{F} \otimes \Sigma)\Omega^{-1}] \\ \text{BTr}_q[(\Psi_{AR} \otimes \mathbf{I}_q)\Omega^{-1}(\Psi_{AR} \otimes \Sigma)\Omega^{-1}] & = \text{BTr}_q[(\Psi_{AR} \otimes \mathbf{I}_q)\Omega^{-1}]. \end{cases}$$

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{CS}, \quad \mathcal{S}_{CS} = \{\mathbf{\Gamma}_{CS} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie [FKM, 2022]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi}_{CS} \otimes \mathbf{\Sigma}$, która minimalizuje kwadratową funkcję straty (86) w zbiorze $\mathcal{S}_{CS \otimes}$ oraz minimum to jest osiągnięte dla $\mathbf{\Psi}_{CS}, \mathbf{\Sigma}$ spełniających następujący układ równań

$$\begin{cases} \text{tr} [\{(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \otimes \mathbf{\Sigma}\} \mathbf{\Omega}^{-1} (\mathbf{\Psi}_{CS} \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1}] &= \text{tr} [\{(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \otimes \mathbf{\Sigma}\} \mathbf{\Omega}^{-1}] \\ \text{BTr}_q [(\mathbf{\Psi}_{CS} \otimes \mathbf{I}_q) \mathbf{\Omega}^{-1} (\mathbf{\Psi}_{CS} \otimes \mathbf{\Sigma}) \mathbf{\Omega}^{-1}] &= \text{BTr}_q [(\mathbf{\Psi}_{CS} \otimes \mathbf{I}_q) \mathbf{\Omega}^{-1}]. \end{cases}$$

$$q = 1$$

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{CS}, \quad \mathcal{S}_{CS} = \{\mathbf{\Gamma}_{CS} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi}_{CS} \otimes \sigma^2$, która minimalizuje kwadratową funkcję straty (86) w zbiorze \mathcal{S}_{CS} oraz minimum to jest osiągnięte dla $\mathbf{\Psi}_{CS}, \sigma^2$ spełniających następujący układ równań

$$\begin{cases} \text{tr} [\{(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \otimes \sigma^2\} \mathbf{\Omega}^{-1} (\mathbf{\Psi}_{CS} \otimes \sigma^2) \mathbf{\Omega}^{-1}] &= \text{tr} [\{(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \otimes \sigma^2\} \mathbf{\Omega}^{-1}] \\ \text{BTr}_1 [(\mathbf{\Psi}_{CS} \otimes \mathbf{I}_1) \mathbf{\Omega}^{-1} (\mathbf{\Psi}_{CS} \otimes \sigma^2) \mathbf{\Omega}^{-1}] &= \text{BTr}_1 [(\mathbf{\Psi}_{CS} \otimes \mathbf{I}_1) \mathbf{\Omega}^{-1}]. \end{cases}$$

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{CS}, \quad \mathcal{S}_{CS} = \{\mathbf{\Gamma}_{CS} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi}_{CS} \otimes \sigma^2$, która minimalizuje kwadratową funkcję straty (86) w zbiorze \mathcal{S}_{CS} oraz minimum to jest osiągnięte dla $\mathbf{\Psi}_{CS}, \sigma^2$ spełniających następujący układ równań

$$\begin{cases} \text{tr} [\{(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \sigma^2\} \mathbf{\Omega}^{-1} \mathbf{\Psi}_{CS} \sigma^2 \mathbf{\Omega}^{-1}] &= \text{tr} [\{(\mathbf{1}_p \mathbf{1}_p' - \mathbf{I}_p) \sigma^2\} \mathbf{\Omega}^{-1}] \\ \text{tr} [\mathbf{\Psi}_{CS} \mathbf{\Omega}^{-1} (\mathbf{\Psi}_{CS} \sigma^2) \mathbf{\Omega}^{-1}] &= \text{tr} [\mathbf{\Psi}_{CS} \mathbf{\Omega}^{-1}]. \end{cases}$$

Aproksymacja za pomocą kwadratową funkcji straty

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{AR}}, \quad \mathcal{S}_{\text{AR}} = \{\mathbf{\Gamma}_{\text{AR}} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie [FKM, 2022]

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}$, która minimalizuje kwadratową funkcję straty postaci (86) w zbiorze $\mathcal{S}_{\text{AR} \otimes}$ oraz minimum to jest osiągnięte dla $\mathbf{\Psi}_{\text{AR}}, \mathbf{\Sigma}$ spełniających następujący układ równań

$$\begin{cases} \text{tr}[(\mathbf{F} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1}] &= \text{tr}[(\mathbf{F} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1}] \\ \text{BTr}_q[(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{I}_q)\mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma})\mathbf{\Omega}^{-1}] &= \text{BTr}_q[(\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{I}_q)\mathbf{\Omega}^{-1}]. \end{cases}$$

$$q = 1$$

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{\text{AR}}, \quad \mathcal{S}_{\text{AR}} = \{\mathbf{\Gamma}_{\text{AR}} \in \mathbb{R}_m^{\geq}\}$$



Twierdzenie

Dla danej symetrycznej, określonej dodatnio macierzy $\mathbf{\Omega}$, istnieje symetryczna macierz $\mathbf{\Psi}_{\text{AR}} \otimes \mathbf{\Sigma}$, która minimalizuje kwadratową funkcję straty postaci (86) w zbiorze $\mathcal{S}_{\text{AR} \otimes}$ oraz minimum to jest osiągnięte dla $\mathbf{\Psi}_{\text{AR}}, \mathbf{\Sigma}$ spełniających następujący układ równań

$$\begin{cases} \text{tr}[(\mathbf{F}\sigma^2)\mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}}\sigma^2)\mathbf{\Omega}^{-1}] = \text{tr}[(\mathbf{F}\sigma^2)\mathbf{\Omega}^{-1}] \\ \text{tr}[(\mathbf{\Psi}_{\text{AR}})\mathbf{\Omega}^{-1}(\mathbf{\Psi}_{\text{AR}}\sigma^2)\mathbf{\Omega}^{-1}] = \text{tr}[(\mathbf{\Psi}_{\text{AR}})\mathbf{\Omega}^{-1}]. \end{cases}$$

Motywacja

n – liczba obiektów

m – liczba charakterystyk

\mathbf{X} – macierz obserwacji

Model wielowymiarowy

$$\mathbf{X} \sim N_{m,n}(\mu \mathbf{1}'_n, \Omega, \mathbf{I}_n)$$

$\mu \in \mathbb{R}^{m,1}$ – wektor średnich

$\Omega \in \mathbb{R}^m$ – nieznana macierz kowariancji

$n > m$ – macierz Ω jest estymowalna

Motywacja

n – liczba obiektów

m – liczba charakterystyk

\mathbf{X} – macierz obserwacji

Model wielowymiarowy

$$\mathbf{X} \sim N_{m,n}(\mu \mathbf{1}'_n, \Omega, \mathbf{I}_n)$$

$\mu \in \mathbb{R}^{m,1}$ – wektor średnich

$\Omega \in \mathbb{R}^m$ – **nieznana** macierz kowariancji

MLE(Ω)

$$\mathbf{S} = \frac{1}{n} \mathbf{X} \mathbf{Q}_{1_n} \mathbf{X}'$$

$$\mathbf{Q}_{1_n} = \mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}'_n = \mathbf{I}_n - \frac{1}{n} \mathbf{J}_n$$

- $m = 100, \quad n = 1000$
- $\mathbf{\Omega} = \mathbf{\Gamma}_{CS} \text{ lub } \mathbf{\Omega} = \mathbf{\Gamma}_{T_1} \text{ lub } \mathbf{\Omega} = \mathbf{\Gamma}_{T_2} \text{ lub } \mathbf{\Omega} = \mathbf{\Gamma}_{AR}$
- $\mathbf{X} \sim N_{m,n}(\mathbf{0}, \mathbf{\Omega}, \mathbf{I}_n)$
- $\mathbf{S} = \frac{1}{n} \mathbf{X} (\mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n') \mathbf{X}'$

$$\zeta_{CS}^{(k)} = \min_{\mathbf{\Gamma} \in \mathcal{S}_{CS}} L_S^{(k)}(\mathbf{\Gamma}), \quad \zeta_{T_1}^{(k)} = \min_{\mathbf{\Gamma} \in \mathcal{S}_{T_1}} L_S^{(k)}(\mathbf{\Gamma}), \quad \zeta_{T_2}^{(k)} = \min_{\mathbf{\Gamma} \in \mathcal{S}_{T_2}} L_S^{(k)}(\mathbf{\Gamma}), \quad \zeta_{AR}^{(k)} = \min_{\mathbf{\Gamma} \in \mathcal{S}_{AR}} L_S^{(k)}(\mathbf{\Gamma})$$
$$k \in \{F, E\}$$

- $r=100$

$$\zeta_{\text{CS}}^{(k)} = \min_{\Gamma \in \mathcal{S}_{\text{CS}}} L_S^{(k)}(\Gamma), \quad \zeta_{T_1}^{(k)} = \min_{\Gamma \in \mathcal{S}_{T_1}} L_S^{(k)}(\Gamma), \quad \zeta_{T_2}^{(k)} = \min_{\Gamma \in \mathcal{S}_{T_2}} L_S^{(k)}(\Gamma), \quad \zeta_{\text{AR}}^{(k)} = \min_{\Gamma \in \mathcal{S}_{\text{AR}}} L_S^{(k)}(\Gamma)$$

$$k \in \{F, E\}$$

- $\Omega = \Gamma_{\text{CS}} \implies \min\{\zeta_{\text{CS}}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{\text{AR}}^{(k)}\} = \zeta_{\text{CS}}^{(k)}$
- $\Omega = \Gamma_{T_1} \implies \min\{\zeta_{\text{CS}}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{\text{AR}}^{(k)}\} = \zeta_{T_1}^{(k)}$
- $\Omega = \Gamma_{T_2} \implies \min\{\zeta_{\text{CS}}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{\text{AR}}^{(k)}\} = \zeta_{T_2}^{(k)}$
- $\Omega = \Gamma_{\text{AR}} \implies \min\{\zeta_{\text{CS}}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{\text{AR}}^{(k)}\} = \zeta_{\text{AR}}^{(k)}$

Badania symulacyjne - identyfikacja struktury

- $m = 100$
- $\sigma^2 = 2$
- $\rho = \rho_1 = 0.25$
- $\rho_2 = 0.1$

True structure Ω	CS	Set of structures		
		T_1	T_2	AR
CS	$\xi_F = \mathbf{1.78}$	$\xi_F = 49.25$	$\xi_F = 48.75$	$\xi_F = 14.93$
	$\xi_E = \mathbf{0.56}$	$\xi_E = 3.11$	$\xi_E = 3.10$	$\xi_E = 3.11$
T_1	$\xi_F = 6.97$	$\xi_F = \mathbf{0.12}$	$\xi_F = 0.15$	$\xi_F = 1.61$
	$\xi_E = 7.80$	$\xi_E = \mathbf{0.56}$	$\xi_E = 5.63$	$\xi_E = 1.04$
T_2	$\xi_F = 7.45$	$\xi_F = 2.80$	$\xi_F = \mathbf{0.15}$	$\xi_F = 1.08$
	$\xi_E = 6.37$	$\xi_E = 1.70$	$\xi_E = \mathbf{0.56}$	$\xi_E = 0.85$
AR	$\xi_F = 15.74$	$\xi_F = 8.08$	$\xi_F = 4.02$	$\xi_F = \mathbf{0.21}$
	$\xi_E = 21.96$	$\xi_E = 5.38$	$\xi_E = 1.72$	$\xi_E = \mathbf{0.56}$

Motywacja

n – liczba obiektów

p – liczba charakterystyk

q – liczba punktów czasowych

\mathbf{X}_i – macierz obserwacji

Model podwójnie wielowymiarowy

$$\mathbf{x}_i = \text{vec} \mathbf{X}_i \sim N_{pq}(\text{vec} \mathbf{M}, \mathbf{\Omega})$$

$$\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \sim N_{pq,n}(\mathbf{1}'_n \otimes \text{vec} \mathbf{M}, \mathbf{\Omega}, \mathbf{I}_n)$$

$\mathbf{M} \in \mathbb{R}^{p,q}$ – macierz średnich (taka sama dla każdego i)

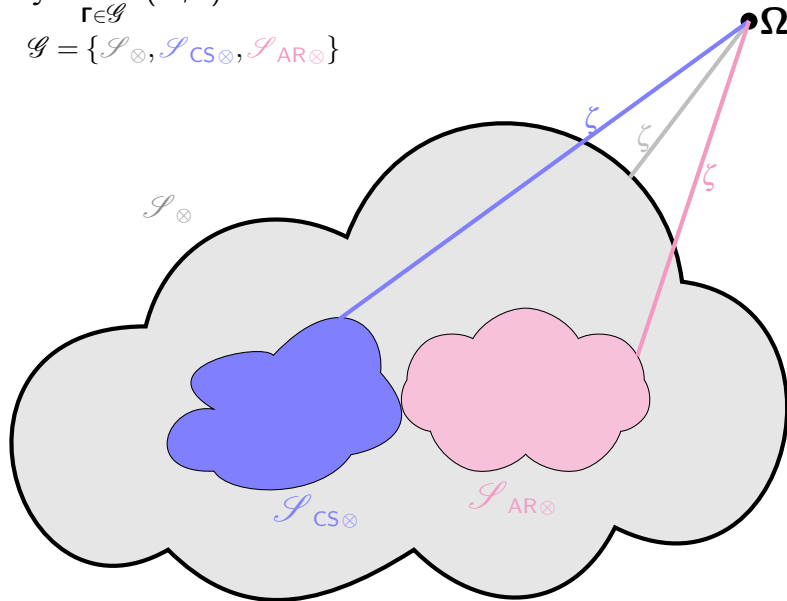
$\mathbf{\Omega} \in \mathbb{R}^{pq}$ – macierz kowariancji, NIEZNANA

$n > pq$ – estymowalna macierz $\mathbf{\Omega}$

Aproksymacja

$$\zeta = \min_{\Gamma \in \mathcal{G}} f(\Omega; \Gamma)$$

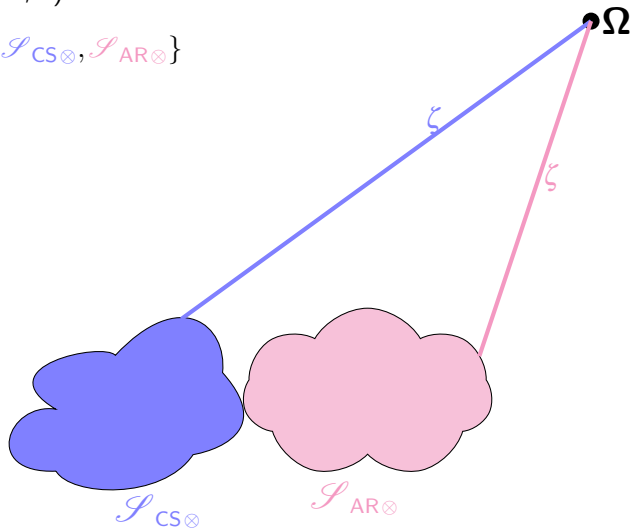
$$\mathcal{G} = \{\mathcal{I}_{\otimes}, \mathcal{I}_{\text{CS}\otimes}, \mathcal{I}_{\text{AR}\otimes}\}$$



Aproksymacja

$$\zeta = \min_{\Gamma \in \mathcal{G}} f(\Omega; \Gamma)$$

$$\mathcal{G} = \{\mathcal{I}_{\otimes}, \mathcal{I}_{\text{CS}\otimes}, \mathcal{I}_{\text{AR}\otimes}\}$$



$$\Gamma: \quad \Psi \otimes \Sigma \quad \Psi_{\text{CS}} \otimes \Sigma \quad \Psi_{\text{AR}} \otimes \Sigma$$

- entropijna funkcja straty:

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

- kwadratowa funkcja straty:

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

Badania symulacyjne - identyfikacja struktury

- $(p, q) \in \{(3, 3), (10, 3), (15, 3), (3, 5), (10, 5), (15, 5)\}, \quad n = 100$

- $\Omega = \Psi_{CS} \otimes I_q \text{ lub } \Omega = \Psi_{AR} \otimes I_q$

- $\mathbf{X} \sim N_{p,q,n}(\mathbf{0}, \Omega, I_n)$

- $\mathbf{S} = \frac{1}{n} \mathbf{X} (I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n') \mathbf{X}'$

$$\zeta_{UN \otimes}^{(k)} = \min_{\Psi, \Sigma} L_S^{(k)}(\Psi \otimes \Sigma), \quad \zeta_{CS \otimes}^{(k)} = \min_{\rho_{CS}, \Sigma} L_S^{(k)}(\Psi_{CS} \otimes \Sigma), \quad \zeta_{AR \otimes}^{(k)} = \min_{\Psi_{AR}, \Sigma} L_S^{(k)}(\Psi_{AR} \otimes \Sigma)$$
$$k \in \{E, Q\}$$

- $r=1000$

- $\Omega = \Psi_{CS} \otimes I_q \implies \zeta_{CS \otimes}^{(k)} \leq \zeta_{AR \otimes}^{(k)}$

$$\Omega = \Psi_{AR} \otimes I_q \implies \zeta_{AR \otimes}^{(k)} \leq \zeta_{CS \otimes}^{(k)}$$

- proporcja poprawnie zidentyfikowanych struktur $\pi^{(k)}$

		Entropijna funkcja straty			Kwadratowa funkcja straty		
p	ρ	$\zeta_{UN \otimes}^{(E)}$	$\zeta_{CS \otimes}^{(E)}$	$\zeta_{AR \otimes}^{(E)}$	$\zeta_{UN \otimes}^{(Q)}$	$\zeta_{CS \otimes}^{(Q)}$	$\zeta_{AR \otimes}^{(Q)}$
3	-0.1		0.4101	0.4468		0.7384	0.7916
3	0		0.4101	0.4104		0.7457	0.7455
3	0.1		0.4101	0.4352		0.7445	0.7823
3	0.25		0.4101	0.5328		0.7436	0.9293
3	0.5		0.4101	0.7680		0.7405	1.2324
3	0.75		0.4101	1.0243		0.7425	1.5095
3	0.9		0.4101	1.1689		0.7467	1.6168
10	-0.1		5.8713	14.6891		9.2058	16.0191
10	0		5.8713	5.8705		9.1562	9.1570
10	0.1		5.8713	6.3907		9.1748	9.5268
10	0.25		5.8713	7.6288		9.1647	10.0114
10	0.5		5.8713	9.8416		9.1722	10.3823
10	0.75		5.8713	12.6912		9.1837	19.7743
10	0.9		5.8713	15.8255		9.1755	20.6869
15	0		15.3668	15.3673		20.6584	20.6600
15	0.1		15.3668	16.2939		20.7239	21.0452
15	0.25		15.3668	17.9833		20.6057	21.2136
15	0.5		15.3668	20.5370		20.6695	21.9300
15	0.75		15.3668	23.5496		20.6584	35.2353
15	0.9		15.3668	26.7463		20.7239	35.7882

		Entropijna funkcja straty			Kwadratowa funkcja straty		
p	ρ	$\zeta_{UN\otimes}^{(E)}$	$\zeta_{CS\otimes}^{(E)}$	$\zeta_{AR\otimes}^{(E)}$	$\zeta_{UN\otimes}^{(Q)}$	$\zeta_{CS\otimes}^{(Q)}$	$\zeta_{AR\otimes}^{(Q)}$
3	-0.1	0.3673	0.4101	0.4468	0.6656	0.7384	0.7916
3	0	0.3673	0.4101	0.4104	0.6680	0.7457	0.7455
3	0.1	0.3673	0.4101	0.4352	0.6685	0.7445	0.7823
3	0.25	0.3673	0.4101	0.5328	0.6688	0.7436	0.9293
3	0.5	0.3673	0.4101	0.7680	0.6648	0.7405	1.2324
3	0.75	0.3673	0.4101	1.0243	0.6635	0.7425	1.5095
3	0.9	0.3673	0.4101	1.1689	0.6697	0.7467	1.6168
10	-0.1	5.1192	5.8713	14.6891	7.5544	9.2058	16.0191
10	0	5.1192	5.8713	5.8705	7.5186	9.1562	9.1570
10	0.1	5.1192	5.8713	6.3907	7.5397	9.1748	9.5268
10	0.25	5.1192	5.8713	7.6288	7.5262	9.1647	10.0114
10	0.5	5.1192	5.8713	9.8416	7.5396	9.1722	10.3823
10	0.75	5.1192	5.8713	12.6912	7.5441	9.1837	19.7743
10	0.9	5.1192	5.8713	15.8255	7.5224	9.1755	20.6869
15	0	13.2331	15.3668	15.3673	16.1694	20.6584	20.6600
15	0.1	13.2331	15.3668	16.2939	16.1766	20.7239	21.0452
15	0.25	13.2331	15.3668	17.9833	16.1458	20.6057	21.2136
15	0.5	13.2331	15.3668	20.5370	16.1810	20.6695	21.9300
15	0.75	13.2331	15.3668	23.5496	16.1694	20.6584	35.2353
15	0.9	13.2331	15.3668	26.7463	16.1766	20.7239	35.7882

		Entropijna funkcja straty			Kwadratowa funkcja straty		
p	ρ	$\zeta_{UN\otimes}^{(E)}$	$\zeta_{CS\otimes}^{(E)}$	$\zeta_{AR\otimes}^{(E)}$	$\zeta_{UN\otimes}^{(Q)}$	$\zeta_{CS\otimes}^{(Q)}$	$\zeta_{AR\otimes}^{(Q)}$
3	-0.9		5.6723	0.4102		2.9835	0.7385
3	-0.5		1.2159	0.4105		1.5982	0.7458
3	-0.1		0.4315	0.4105		0.7764	0.7444
3	0		0.4101	0.4104		0.7436	0.7441
3	0.1		0.4283	0.4104		0.7676	0.7406
3	0.5		0.7327	0.4101		1.1854	0.7424
3	0.9		1.1618	0.4099		1.6283	0.7465
10	-0.9		24.7209	5.8707		14.7257	9.2062
10	-0.5		11.5063	5.8705		12.5526	9.1550
10	-0.1		6.0907	5.8704		9.3764	9.1783
10	0		5.8713	5.8705		9.1647	9.1627
10	0.1		6.0799	5.8706		9.3499	9.1703
10	0.5		10.1883	5.8714		11.8149	9.1835
10	0.9		16.3827	5.8715		13.3154	9.1743
15	-0.9		43.2050	15.3688		25.9321	20.6622
15	-0.5		24.3583	15.3678		24.0033	20.7219
15	-0.1		15.7331	15.3673		20.8070	20.6064
15	0		15.3668	15.3673		20.6695	20.6709
15	0.1		15.7220	15.3674		20.7976	20.6146
15	0.5		22.8995	15.3679		23.3561	20.5539
15	0.9		34.2544	15.3669		30.6651	20.6651

		Entropijna funkcja straty			Kwadratowa funkcja straty		
p	ρ	$\zeta_{UN\otimes}^{(E)}$	$\zeta_{CS\otimes}^{(E)}$	$\zeta_{AR\otimes}^{(E)}$	$\zeta_{UN\otimes}^{(Q)}$	$\zeta_{CS\otimes}^{(Q)}$	$\zeta_{AR\otimes}^{(Q)}$
3	-0.9	0.3673	5.6723	0.4102	0.6656	2.9835	0.7385
3	-0.5	0.3673	1.2159	0.4105	0.6680	1.5982	0.7458
3	-0.1	0.3673	0.4315	0.4105	0.6685	0.7764	0.7444
3	0	0.3673	0.4101	0.4104	0.6688	0.7436	0.7441
3	0.1	0.3673	0.4283	0.4104	0.6648	0.7676	0.7406
3	0.5	0.3673	0.7327	0.4101	0.6635	1.1854	0.7424
3	0.9	0.3673	1.1618	0.4099	0.6697	1.6283	0.7465
10	-0.9	5.1192	24.7209	5.8707	7.5544	14.7257	9.2062
10	-0.5	5.1192	11.5063	5.8705	7.5186	12.5526	9.1550
10	-0.1	5.1192	6.0907	5.8704	7.5397	9.3764	9.1783
10	0	5.1192	5.8713	5.8705	7.5262	9.1647	9.1627
10	0.1	5.1192	6.0799	5.8706	7.5396	9.3499	9.1703
10	0.5	5.1192	10.1883	5.8714	7.5441	11.8149	9.1835
10	0.9	5.1192	16.3827	5.8715	7.5224	13.3154	9.1743
15	-0.9	13.2331	43.2050	15.3688	16.1694	25.9321	20.6622
15	-0.5	13.2331	24.3583	15.3678	16.1766	24.0033	20.7219
15	-0.1	13.2331	15.7331	15.3673	16.1458	20.8070	20.6064
15	0	13.2331	15.3668	15.3673	16.1810	20.6695	20.6709
15	0.1	13.2331	15.7220	15.3674	16.1384	20.7976	20.6146
15	0.5	13.2331	22.8995	15.3679	16.1350	23.3561	20.5539
15	0.9	13.2331	34.2544	15.3669	16.1544	25.9321	20.6622

$\Psi_{CS} \otimes I_3$

	p=3		p=10		p=15	
ρ	$\pi^{(E)}$	$\pi^{(Q)}$	$\pi^{(E)}$	$\pi^{(Q)}$	$\pi^{(E)}$	$\pi^{(Q)}$
-0.1	83.3	82.3	100.0	100.0	-	-
0	49.5	49.4	51.3	50.4	48.1	50.9
0.1	79.0	77.3	100.0	98.4	100.0	95.4
0.25	96.0	93.9	100.0	100.0	100.0	99.2
0.5	99.7	99.7	100.0	100.0	100.0	99.7
0.75	100.0	100.0	100.0	100.0	100.0	100.0
0.9	100.0	100.0	100.0	100.0	100.0	100.0

 $\Psi_{AR} \otimes I_3$

	p=3		p=10		p=15	
ρ	$\pi^{(E)}$	$\pi^{(Q)}$	$\pi^{(E)}$	$\pi^{(Q)}$	$\pi^{(E)}$	$\pi^{(Q)}$
-0.9	100.0	100.0	100.0	100.0	100.0	100.0
-0.5	100.0	100.0	100.0	100.0	100.0	100.0
-0.1	77.1	75.9	98.8	91.7	99.2	86.9
0	50.5	48.6	48.7	50.5	51.9	49.5
0.1	73.1	71.3	98.0	89.7	99.3	86.4
0.5	99.8	99.4	100.0	100.0	100.0	100.0
0.9	100.0	99.9	100.0	100.0	100.0	100.0

Poprawne wykrywanie struktury kowariancyjnej symulowanych danych

- dla $k \in \{F, E\}$

- $\bullet \quad \Omega = \Gamma_{CS} \implies \min\{\zeta_{CS}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{AR}^{(k)}\} = \zeta_{CS}^{(k)}$

- $\bullet \quad \Omega = \Gamma_{T_1} \implies \min\{\zeta_{CS}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{AR}^{(k)}\} = \zeta_{T_1}^{(k)}$

- $\bullet \quad \Omega = \Gamma_{T_2} \implies \min\{\zeta_{CS}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{AR}^{(k)}\} = \zeta_{T_2}^{(k)}$

- $\bullet \quad \Omega = \Gamma_{AR} \implies \min\{\zeta_{CS}^{(k)}, \zeta_{T_1}^{(k)}, \zeta_{T_2}^{(k)}, \zeta_{AR}^{(k)}\} = \zeta_{AR}^{(k)}$

- dla $k \in \{E, Q\}$

$$\Omega \in \mathcal{S}_{CS \otimes} \implies \zeta_{CS \otimes}^{(k)} \leq \zeta_{AR \otimes}^{(k)}$$

$$\Omega \in \mathcal{S}_{AR \otimes} \implies \zeta_{AR \otimes}^{(k)} \leq \zeta_{CS \otimes}^{(k)}$$

- wartości $\zeta_{UN \otimes}^{(k)}$ są zawsze najmniejsze

- ❶ T.W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, 2003.
- ❷ 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.
- ❸ P.L. Fackler. Notes on matrix calculus. <http://www4.ncsu.edu/~pfackler/MatCalc.pdf>, 2005.
- ❹ 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.
- ❺ 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.
- ❻ 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.
- ❼ 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.
- ❽ K. Filipiak, A. Markiewicz, A. Mieldzioc, A. Sawikowska. On projection of a positive definite Toeplitz matrices. *Electronic Journal of Linear Algebra*, 33, 74–82, 2018.
- ❾ L. Lin, N. Higham, J. Pan. Covariance structure regularization via entropy loss function. *Comput. Statist. Data Anal.* 72, 315–327, 2014.
- ❿ J. Magnus, H. Neudecker. Symmetry, 0-1 matrices and Jacobians, a review. *Econom. Theory* 2, 157–190, 1986.
- ⓫ C.F. van Loan, N. Pitsianis. Approximation with Kronecker products. In: B.L.R. De Moor, M.S. Moonen, G.H. Golub (Eds.), *Linear Algebra for Large Scale and Real-Time Applications* (pp. 293–314), Kluwer Publications, Dordrecht, The Netherlands, 1992.