Wielowymiarowe modele w analizie danych biologicznych

Monika Mokrzycka

Instytut Genetyki Roślin PAN w Poznaniu

Warsztaty, Politechnika Warszawska

Rodzaje zbóż

https://blog.fmcagro.pl/rodzaje-zboz/

Rodzaje zbóż

Pszenica zwyczajna (*Triticum aestivum* L.)

https://blog.fmcagro.pl/rodzaje-zboz/

- jedna z najstarszych roślin uprawnych
- pierwsze wzmianki pochodzą z 9-6 tysiącleci p.n.e.
- dziś zboże nr 1
- zawartość odżywcza:
  - 12g białka (80% gluten)
  - 2,5g tłuszczu
  - 58g węglowodanów
  - 11g błonnika
- zawiera:
  - mangan
  - krzem
  - kobalt
  - selen
  - miedź
  - fosfor
  - molibden
  - żelazo

https://farm-pl.desigusxpro.com

Rodzaje zbóż

Jęczmień zwyczajny (*Hordeum vulgare* L.)

https://blog.fmcagro.pl/rodzaje-zboz/

- uprawę rozpoczęto ok. 10 tysięcy lat temu
- dziś zajmuje 4 miejsce pod względem powierzchni zasiewów po pszenicy, kukurydzy i ryżu
- zawartość odżywcza:
  - 10,3g białka (23% gluten)
  - 2,4g tłuszczu
  - 56,4g węglowodanów
  - 14,5g błonnika
- zawiera:
  - krzem
  - kobalt
  - mangan
  - miedź
  - fosfor
  - żelazo
  - selen
  - magnez

https://farm-pl.desigusxpro.com

Rodzaje zbóż

Żyto (*Secale* L.)

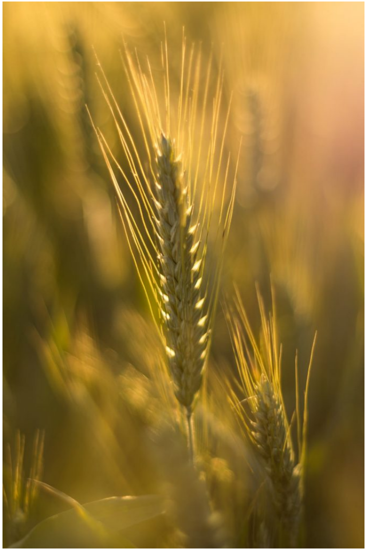
https://farm-pl.desigusxpro.com

- pierwotnie postrzegane jako chwast
- udomowienie nastąpiło 2 tysiąclecia p.n.e.
- dziś: Polska, Niemcy
- zawartość odżywcza:
  - 9,9g białka (16% gluten)
  - 2,2g tłuszczu
  - 58,8g węglowodanów
  - 16,4g błonnika
- zawiera:
  - krzem
  - mangan
  - kobalt
  - miedź
  - selen
  - fosfor
  - żelazo
  - magnez
  - molibden

https://farm-pl.desigusxpro.com

Rodzaje zbóż

Pszenżyto ($\times$ *Triticosecale*)



https://blog.fmcagro.pl/rodzaje-zboz/

- powstaje w wyniku zapylenia pszenicy żytem
- łączy cechy wyglądu obu gatunków
- zawartość odżywcza:
  - 13g białka (gluten)
  - 2g tłuszczu
  - 72g węglowodanów
  - 0g błonnika
- zawiera:
  - wapń
  - żelazo
  - magnez
  - fosfor
  - potas

https://zywienie.abczdrowie.pl/pszenzyto

Rodzaje zbóż

Owies siewny (*Avena sativa* L.)

https://farm-pl.desigusxpro.com

- pierwotnie postrzegane jako chwast
- udomowienie nastąpiło 2 tysiąclecia p.n.e.
- najwyższa kaloryczność z omawianych zbóż
- zawartość odżywcza:
  - 10g białka (0-21% gluten)
  - 6,2g tłuszczu
  - 55,1g węglowodanów
  - 12g błonnika
- zawiera:
  - mangan
  - kobalt
  - miedź
  - molibden
  - selen
  - fosfor
  - magnez
  - żelazo
  - cynk

https://farm-pl.desigusxpro.com

Dane metabolomiczne

Identyfiacja struktury macierzy kowariancji
na przykładzie danych metabolomicznych

Metabolomika - dziedzina nauki zajmująca się analizą niskocząsteczkowych produktów naturalnych endogennych metabolitów (pierwotnych i wtórnych), które tworzą metabolom.

Metabolit – produkt metabolizmu (przemian chemicznych zachodzących w organizmach). Metabolity to związki organiczne i nieorganiczne produkowane przez komórki. Przyjęto jednak, że określenie to nie dotyczy białek i kwasów nukleinowych.

Metabolom – zestaw wszystkich metabolitów obecnych w organizmie, tkance, komórce lub przedziale komórkowym (np. pochodne aminokwasów, lipidów, węglowodanów, nukleotydów, drobnocząsteczkowych hormonów takich jak np. tyroksyna, testosteron, estradiol, cykliczny AMP itp.).

www.wikipedia.pl

Dane metabolomiczne

- dane jęczmienne (*Hordeum vulgare* L.)
- chromatografia gazowa ze spektrometrią mas (GC-MS) film
- odporność na suszę

- surowe dane: 51135 charakterystyk dla 422 próbek
- 211 prób biologicznych po uśrednieniu po powtórzeniach technicznych
- 781 charakterystyk po wstępnym przetworzeniu
- przekształcone przez logarytm o podstawie 1.2

Identyfikacja struktury

Struktury kowariancyjne $m \times m$:

- kompletnej symetrii (CS)
- trójdiagonalne macierze Toeplitza (T_1)
- pięciodiagonalne macierze Toeplitza (T_2)
- autoregresji stopnia pierwszego (AR)

Metody identyfikacji struktury:

- norma Frobeniusa
- entropijna funkcja straty

- kompletnej symetrii (CS)

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

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

Struktury kowariancyjne

- kompletnej symetrii (CS)
- wstęgowa macierz Toeplitza (T_p)

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

$\mathbf{H}_i$ - symetryczna macierz stopnia m z elementami na $i-$tych przekątnych równych 1, pozostałe równe 0
$\mathbf{H}_0 = \mathbf{I}_m$

$$T_1 \in \mathbb{R}^> : \quad \sigma^2 > 0 \text{ i } \rho_1 \in \left(-\frac{1}{2\cos(\pi/(m+1))}; \frac{1}{2\cos(\pi/(m+1))} \right)$$

Struktury kowariancyjne

- kompletnej symetrii (CS)
- wstęgowa macierz Toeplitza (T_p)
- autoregresji pierwszego rzędu (AR)

$$\mathbf{\Gamma}_{AR} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{m-1} \\ \rho & 1 & \rho & \dots & \rho^{m-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \dots & 1 \end{pmatrix} = \sigma^2 \left(\mathbf{I}_m + \sum_{i=1}^{m-1} \rho^i \mathbf{H}_i \right)$$

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

Struktury kowariancyjne

$\mathscr{S} = \{\mathscr{S}_{CS}, \mathscr{S}_{T_1}, \mathscr{S}_{T_2}, \mathscr{S}_{AR}\}.$

$$\zeta = \min_{\boldsymbol{\Gamma} \in \mathscr{S}} f(\boldsymbol{\Sigma}, \boldsymbol{\Gamma})$$

$\boldsymbol{\Sigma}$ - nieznana

MLE($\boldsymbol{\Sigma}$)

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

gdzie

- $\mathbf{X}$ - macierz obserwacji $\quad \mathbf{X} \sim N_{m,n}(\mu \mathbf{1}_n', \boldsymbol{\Sigma}_m, \mathbf{I}_n)$
- $\mathbf{Q}_{1_n} = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$

Metody identyfikacji struktury

$\mathbf{S}, \boldsymbol{\Gamma} \in \mathbb{R}^{>}_{m}, \; \mathrm{Sym(m)}$

Frobenius norm

$$f_F(\mathbf{S}, \boldsymbol{\Gamma}) = ||\mathbf{S} - \boldsymbol{\Gamma}||_F = \sqrt{\mathrm{tr}\,[(\mathbf{S} - \boldsymbol{\Gamma})^2]}$$

Entropy loss function

$$f(\mathbf{S}, \boldsymbol{\Gamma}) \quad = \quad \mathrm{tr}\left(\mathbf{S}^{-1}\boldsymbol{\Gamma}\right) - \ln\left(\det\left[\mathbf{S}^{-1}\boldsymbol{\Gamma}\right]\right) - m$$

$$\zeta_F = \min_{\boldsymbol{\Gamma} \in \mathscr{S}} f_F(\mathbf{S}, \boldsymbol{\Gamma}) \qquad \text{i} \qquad \zeta_E = \min_{\boldsymbol{\Gamma} \in \mathscr{S}} f_E(\mathbf{S}, \boldsymbol{\Gamma})$$

Aproksymacja za pomocą normy Frobeniusa [Cui et al. (2016)]

CS structure:

$$\begin{cases} \rho & = \frac{\alpha}{(m-1)\mathrm{tr}(\mathbf{S})} \\ \sigma^2 & = \frac{\mathrm{tr}(\mathbf{S})+\rho\alpha}{m+m(m-1)\rho^2} \end{cases}$$

$$\alpha = \mathrm{tr}\left[\mathbf{S}(\mathbf{1}_m\mathbf{1}_m' - \mathbf{I}_m)\right]$$

T_1 structure:

$$\begin{cases} \sigma^2 & = \frac{\mathrm{tr}(\mathbf{S})}{m} \\ \rho_1 & = \frac{m\mathrm{tr}(\mathbf{SH}_1)}{2(m-1)\mathrm{tr}(\mathbf{S})}, \end{cases}$$

$AR(1)$ structure:

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

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

Aproksymacja za pomocą normy Frobeniusa [Cui et al. (2016)]

CS structure:

$$\begin{cases} \rho &= \frac{\alpha}{(m-1)\mathrm{tr}(\mathbf{S})} \\ \sigma^2 &= \frac{\mathrm{tr}(\mathbf{S})+\rho\alpha}{m+m(m-1)\rho^2} \end{cases}$$

$$\alpha = \mathrm{tr}\left[\mathbf{S}(\mathbf{1}_m\mathbf{1}_m' - \mathbf{I}_m)\right]$$

T_1 structure:

$$\begin{cases} \sigma^2 &= \frac{\mathrm{tr}(\mathbf{S})}{m} \\ \rho_1 &= \frac{m\,\mathrm{tr}(\mathbf{SH}_1)}{2(m-1)\mathrm{tr}(\mathbf{S})} \end{cases}$$

Filipiak et al., 2018

$AR(1)$ structure:

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

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

Aproksymacja za pomocą entropijnej funkcji straty [Lin et al. (2014)]

CS structure:

$$\begin{cases} \rho & = & -\frac{\alpha}{(m-1)\mathrm{tr}(\mathbf{S}^{-1})+(m-2)\alpha} \\ \frac{m}{\sigma^2} & = & \mathrm{tr}(\mathbf{S}^{-1})+\rho\alpha \end{cases}$$

$$\alpha = \mathrm{tr}\left[\mathbf{S}^{-1}(\mathbf{1}_m\mathbf{1}_m' - \mathbf{I}_m)\right]$$

T_1 structure:

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

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

$AR(1)$ structure:

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

Zbiory danych

Spośród 781 charakterystyk wybrano 3 podzbiory

- cechy od 1 do 200
- cechy od 201 do 400
- cechy od 401 do 600

$\mathbf{X} : 200 \times 211$ $\mathbf{S} : 200 \times 200$

Skorygowane wartości:

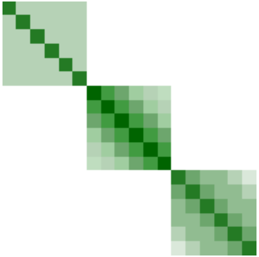
- $\xi_F = \zeta_F / ||\mathbf{S}||_F$
- $\xi_E = 1 - 1/(1 + \zeta_E)$

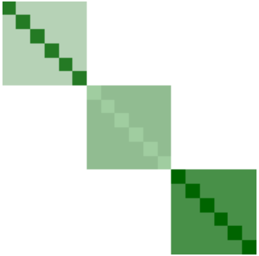
Rozbieżności

struktura	podzbiór	ζ_F	ζ_E	ξ_F	ξ_E
CS	1	1399.1336	655.5110	0.3271	0.7380
	2	1387.0465	588.9344	0.2992	0.7348
	3	1512.8434	619.9874	0.2852	0.7364
T_1	1	4235.6872	659.3827	0.9902	0.7382
	2	4594.1416	589.3869	0.9910	0.7348
	3	5257.9262	623.2153	0.9912	0.7365
T_2	1	4219.1889	659.3437	0.9864	0.7382
	2	4572.1144	589.0878	0.9862	0.7348
	3	5232.9662	623.1698	0.9865	0.7365
AR	1	1404.7870	659.3951	0.3284	0.7382
	2	1379.1035	589.0559	0.2975	0.7348
	3	1477.2560	623.2280	0.2785	0.7365

Estymatory

struktura	podzbiór	ζ_F	ζ_E	ξ_F	ξ_E
CS	1	1399.1336	655.5110	0.3271	0.7380
	2	1387.0465	588.9344	0.2992	0.7348
	3	1512.8434	619.9874	0.2852	0.7364
T_1	1	4235.6872	659.3827	0.9902	0.7382
	2	4594.1416	589.3869	0.9910	0.7348
	3	5257.9262	623.2153	0.9912	0.7365
T_2	1	4219.1889	659.3437	0.9864	0.7382
	2	4572.1144	589.0878	0.9862	0.7348
	3	5232.9662	623.1698	0.9865	0.7365
AR	1	1404.7870	659.3951	0.3284	0.7382
	2	1379.1035	589.0559	0.2975	0.7348
	3	1477.2560	623.2280	0.2785	0.7365

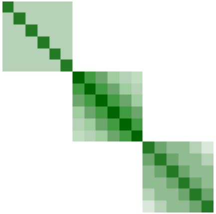




Estymatory

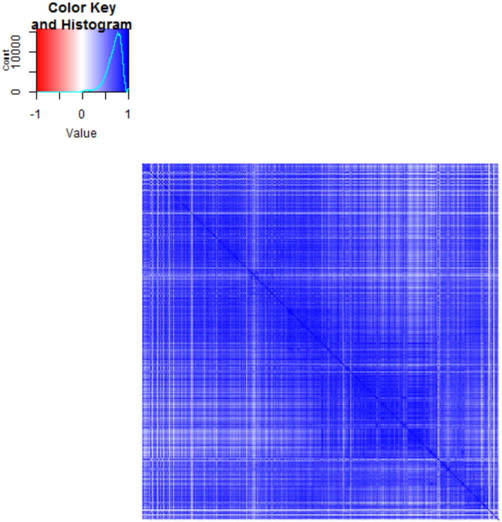
Estymatory uzyskane za pomocą:

- normy Frobeniusa, $\widehat{\Gamma}$
- entropijnej funkcji straty, $\widetilde{\Gamma}$
- funkcji wiarogodności, $\check{\Gamma}$



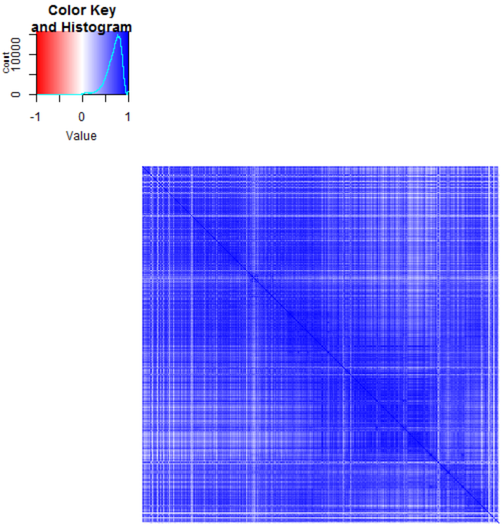
estymator	podzbiór	σ^2	ρ	ρ_1	ρ_2
$\widehat{\Gamma}_{CS}$	1	30.7322	0.6555	-	-
	2	29.5329	0.7475	-	-
	3	33.2362	0.7635	-	-
$\widetilde{\Gamma}_{CS}$	1	0.0632	0.5145	-	-
	2	0.0599	0.4558	-	-
	3	0.0443	0.3737	-	-
$\check{\Gamma}_{CS}$	1	30.7322	0.6555	-	-
	2	29.5329	0.7475	-	-
	3	33.2362	0.7635	-	-
$\widehat{\Gamma}_{T_1}$	1	34.4749	-	0.5001	-
	2	35.8953	-	0.5001	-
	3	40.5434	-	0.5001	-
$\widetilde{\Gamma}_{T_1}$	1	0.0310	-	0.0495	-
	2	0.0340	-	0.1329	-
	3	0.0281	-	0.0015	-
$\widehat{\Gamma}_{T_2}$	1	36.0277	-	0.5351	0.4136
	2	39.1610	-	0.5405	0.4114
	3	44.3696	-	0.5398	0.4117
$\widetilde{\Gamma}_{T_2}$	1	0.0310	-	0.0015	-0.0004
	2	0.0341	-	0.0049	0.0013
	3	0.0281	-	0.0015	-0.0004
$\widehat{\Gamma}_{AR}$	1	20.7750	0.9996	-	-
	2	23.4344	0.9991	-	-
	3	27.9402	0.9985	-	-
$\widetilde{\Gamma}_{AR}$	1	0.0310	0.0485	-	-
	2	0.0342	0.1494	-	-
	3	0.0281	0.0525	-	-
$\check{\Gamma}_{AR}$	1	30.6336	0.6808	-	-
	2	29.4393	0.8230	-	-
	3	32.9569	0.8277	-	-

All data

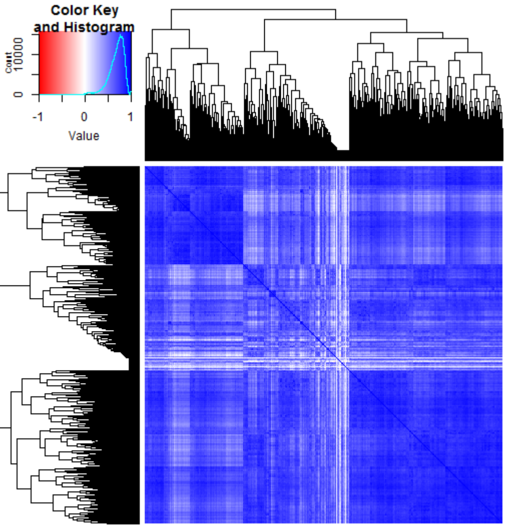


- $\mathbf{X}_{781\times 211} \sim N_{m,n}(\mu \mathbf{1}'_n, \mathbf{\Sigma}, \mathbf{I}_n)$
- $\mathbf{S}_{781\times 781}$ - MLE of $\mathbf{\Sigma}$
- $\det(\mathbf{S}) = 0$
- Macierz korelacji $\mathbf{R} = \mathbf{D}^{-1}\mathbf{S}\mathbf{D}^{-1}$
- $\mathbf{D}^{-1} = diag(\frac{1}{\sqrt{S_{11}}}, \frac{1}{\sqrt{S_{22}}}, \ldots, \frac{1}{\sqrt{S_{mm}}})$

All data



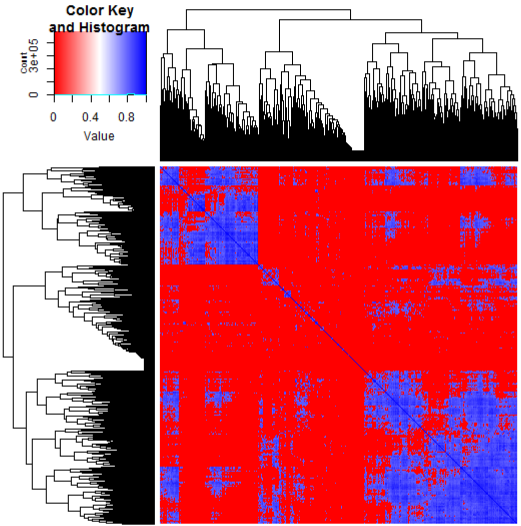
Color Key
and Histogram
Count
0
10000
-1
0
1
Value



Color Key
and Histogram
Count
0
10000
-1
0
1
Value

All data

- if $\mathbf{R}_{ij} < 0.8$ then $\mathbf{R}_{ij} = 0$



Załóżmy:

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Sigma}_{33} \end{pmatrix}$$

gdzie

$$\begin{aligned} \boldsymbol{\Sigma}_{11} &= \sigma_1^2(\rho_1 \mathbf{J}_{215} + (1-\rho_1))\mathbf{I}_{215} \\ \boldsymbol{\Sigma}_{22} &= \sigma_2^2 \mathbf{I}_{231} \\ \boldsymbol{\Sigma}_{33} &= \sigma_3^2(\rho_3 \mathbf{J}_{335} + (1-\rho_3))\mathbf{I}_{335} \end{aligned}$$

Wyniki

$$\mathbf{S} \longrightarrow \mathbf{PSP}' = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{33} \end{pmatrix}$$

$$\min \|\mathbf{S}_{ii} - \mathbf{\Sigma}_{ii}\|_F, \quad i = 1, 2, 3$$

$\widetilde{\mathbf{\Sigma}}_{11}$

$$\begin{aligned} \sigma^2 &= 31.4724 \\ \rho &= 0.5746 \end{aligned}$$

$\widetilde{\mathbf{\Sigma}}_{22}$

$$\sigma^2 = 29.14777$$

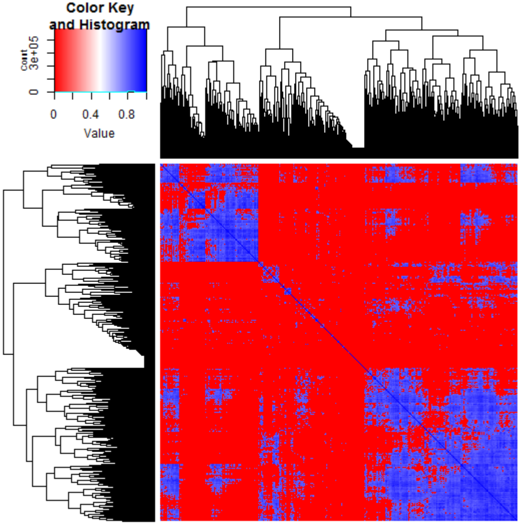
$\widetilde{\mathbf{\Sigma}}_{33}$

$$\begin{aligned} \sigma^2 &= 26.4708 \\ \rho &= 0.5868 \end{aligned}$$

$$\zeta_F = \min \|\mathbf{PSP}' - \widetilde{\mathbf{\Sigma}}_1\|_F = 12817.32$$

$$\xi_F = \zeta_F / ||\mathbf{PSP}'||_F = 0.8325$$

All data



Załóżmy:

$$\widetilde{\mathbf{S}}_2 = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{13} \\ \mathbf{\Sigma}_{12}' & \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{23} \\ \mathbf{\Sigma}_{13}' & \mathbf{\Sigma}_{23}' & \mathbf{\Sigma}_{33} \end{pmatrix}$$

gdzie

$$\begin{aligned} \mathbf{\Sigma}_{11} &= \sigma_1^2(\rho_1 \mathbf{J}_{215} + (1-\rho_1))\mathbf{I}_{215} \\ \mathbf{\Sigma}_{22} &= \sigma_2^2 \mathbf{I}_{231} \\ \mathbf{\Sigma}_{33} &= \sigma_3^2(\rho_3 \mathbf{J}_{335} + (1-\rho_3))\mathbf{I}_{335} \\ \\ \mathbf{\Sigma}_{ij} &= \delta_k \mathbf{J} \\ & \quad i = 1,2 < j = 2,3, \quad k = 1,2,3 \end{aligned}$$

Results

$$\mathbf{S} \longrightarrow \mathbf{PSP'} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{S}'_{12} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{S}'_{13} & \mathbf{S}'_{23} & \mathbf{S}_{33} \end{pmatrix}$$

$$\min \|\mathbf{S}_{ij} - \boldsymbol{\Sigma}_{ij}\|_F, \quad i = 1,2 < j = 2,3$$

$\widetilde{\boldsymbol{\Sigma}}_{11}$

$$\begin{aligned} \sigma_1^2 &= 31.4724 \\ \rho_1 &= 0.5746 \end{aligned}$$

$\widetilde{\boldsymbol{\Sigma}}_{22}$

$$\sigma_2^2 = 29.14777$$

$\widetilde{\boldsymbol{\Sigma}}_{33}$

$$\begin{aligned} \sigma_3^2 &= 26.4708 \\ \rho_3 &= 0.5868 \end{aligned}$$

$\widetilde{\boldsymbol{\Sigma}}_{ij}$

$$\begin{aligned} \delta_1 &= 14.71865 \\ \delta_2 &= 18.57934 \\ \delta_3 &= 16.74399 \end{aligned}$$

$$\zeta_F = \min \|\mathbf{PSP'} - \widetilde{\boldsymbol{\Sigma}}_2\|_F = 7046.1770$$

$$\xi_F = \zeta_F / \|\mathbf{PSP'}\|_F = 0.4577$$

References

1. Cui X., Li X., Zhao J., Zeng L., Zhang D., Pan J. Covariance structure regularization via Frobenius norm discrepancy. Linear Algebra Appl. 510:124–145, 2016.
2. Dey D.K., Srinivasan C. Estimation of a covariance matrix under Stein's loss. Ann. Statist. 13(4):1581-1591, 1985.
3. Filipiak K., Klein D. Approximation with Kronecker product structure with one component as compound symmetry or autoregression. *Linear Algebra Appl.* 559, 11–33, 2018.
4. Filipiak K., Klein D., Markiewicz A., Mokrzycka M. Kronecker product approximation via entropy loss function. *Linear Algebra Appl.* 610, 625–646, 2021.
5. Filipiak K., Klein D., Mokrzycka M. Estimators comparison of separable covariance structure with one component as compound symmetry matrix. *Electronic J. Linear Algebra* 33, 83–98, 2018.
6. Filipiak K., Markiewicz A., Mieldzioc A., Sawikowska A. On projection of a positive definite matrix on a cone of nonnegative definite Toeplitz matrices. *Electronic Journal of Linear Algebra*, 33, 74–82, 2018.
7. Mieldzioc A., Mokrzycka M., Sawikowska A. Covariance regularization for metabolomic data on the drought resistance of barley. *Biometrical Letters*, 56:2, 165–181, 2020.
8. Mieldzioc A., Mokrzycka M., Sawikowska A. Identification of Block-Structured Covariance Matrix on an Example of Metabolomic Data. Separations. 8(11), 205, 2021.
9. Lin L., Higham N. J., Pan J. Covariance structure regularization via entropy loss function. Computational Statistics and Data Analysis 72, 315–327, 2014.