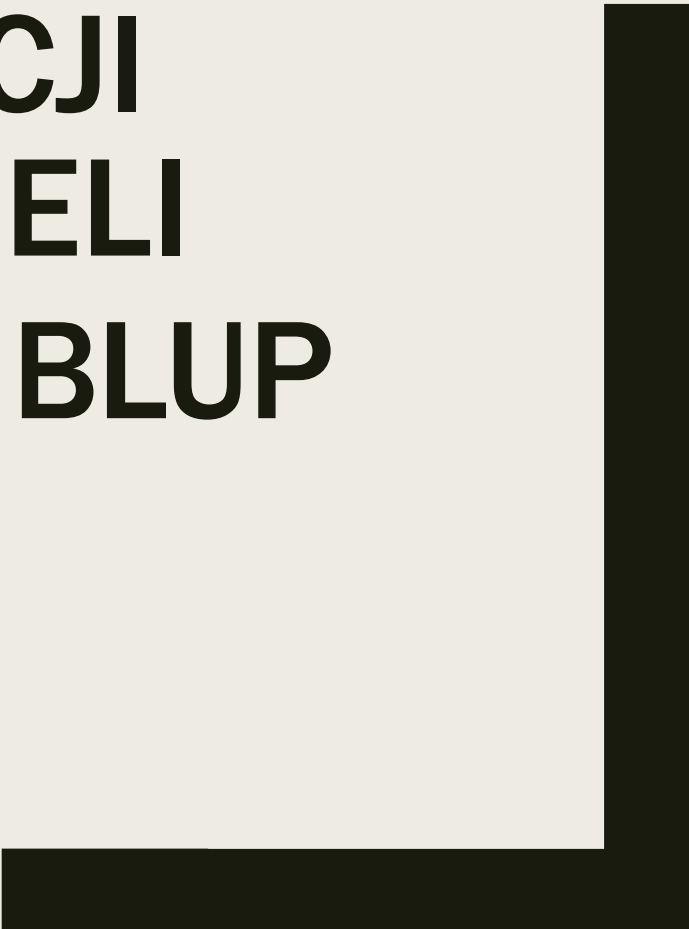




METODY REDUKCJI WYMIARÓW MODELI WIELOCECHOWYCH BLUP

Dawid Słomian

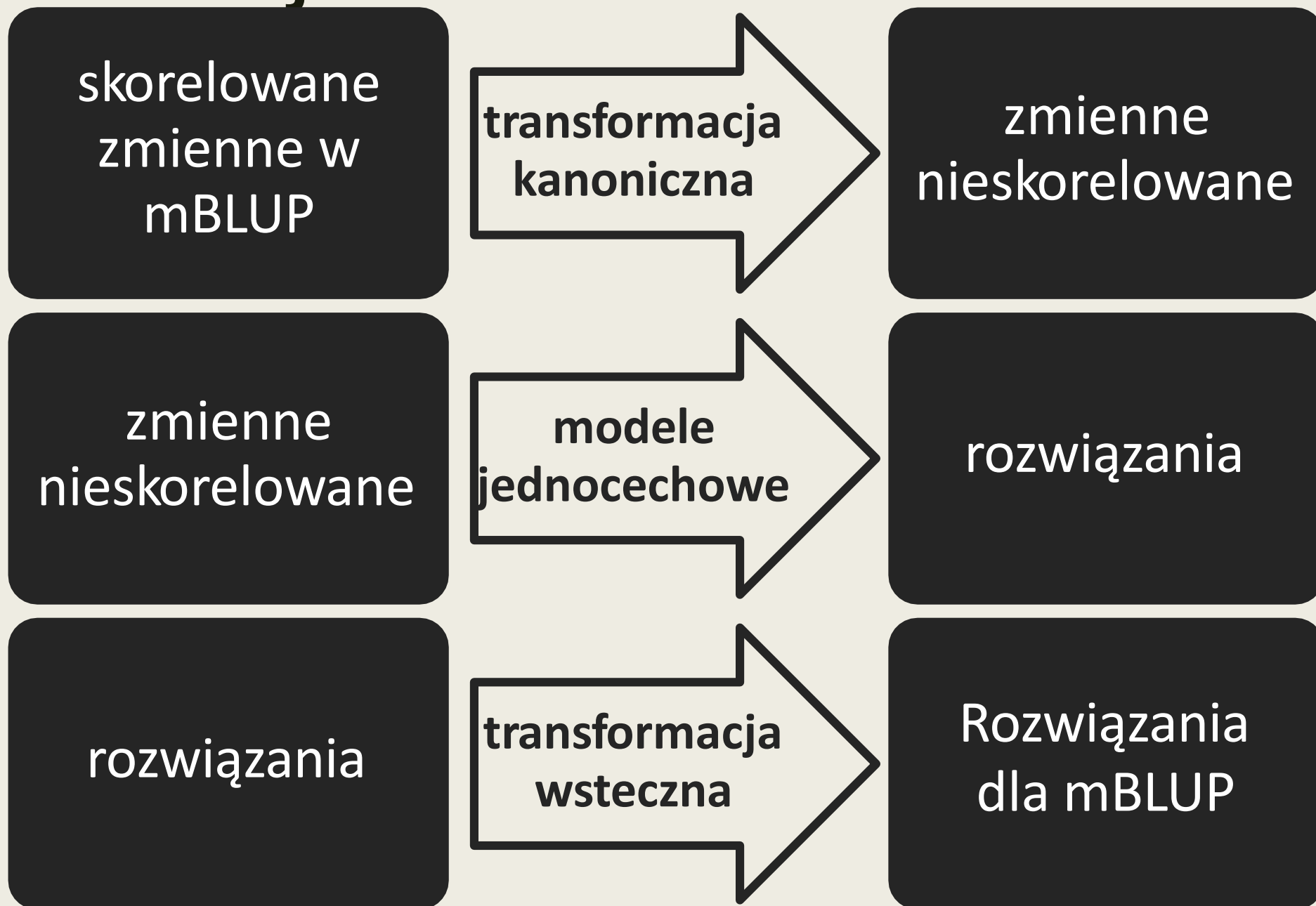


- Wielocechowy model BLUP
 - duże wymiary modelu
 - estymacja parametrów wymagająca obliczeniowo
- Metody redukcji wymiarów modelu
 - Transformacja kanoniczna ***canonical transformation*** → jednakowe macierze wystąpień, bez brakujących danych
 - Transformacja Choleskiego ***Cholesky transformation*** → sekwencyjnie brakujące dane

$$y = Xb + Za + e$$

- ***y*** obserwacje fenotypowe
31 331 172 x 1 31 998 730 x 1 34 000 684 x 1
- ***b*** estymatory efektów stałych
463 576 x 1 569 385 x 1 600 158 x 1
- ***a*** predyktory wartości hodowlanych
37 354 579 x 1 38 389 025 x 1 40 496 318 x 1
- ***e*** predyktory efektów błędu
31 331 172 x 1 31 998 730 x 1 34 000 684 x 1
- ***X*** 31 331 172 x 463 576 31 998 730 x 569 385
- ***Z*** 31 331 172 x 37 354 579 31 998 730 x 38 389 025

Transformacja kanoniczna



Transformacja kanoniczna

- Thompson (1977)
 - Identyczne macierze wystąpień
 - Nie ma brakujących danych
 - Dwa efekty losowe a i e
- Ducrocq & Besbes (1993)
 - Brakujące rekordy danych
- Misztal (1995)
 - Więcej niż dwa efekty losowe

Transformacja kanoniczna

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & \mathbf{0} \\ \mathbf{0} & X_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} Z_1 & \mathbf{0} \\ \mathbf{0} & Z_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$G_0 = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad R_0 = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

$$\text{var} \begin{bmatrix} a_1 \\ a_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} g_{11}A & g_{12}A & \mathbf{0} & \mathbf{0} \\ g_{21}A & g_{22}A & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & r_{11}I & r_{12}I \\ \mathbf{0} & \mathbf{0} & r_{21}I & r_{22}I \end{bmatrix}$$

Transformacja kanoniczna

- Wektor skorelowanych obserwacji: $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$$\text{var}(\mathbf{y}) \sim \mathbf{G}_0 + \mathbf{R}_0 \qquad \text{var}[\mathbf{y}] = \begin{bmatrix} g_{11}\mathbf{A} & g_{12}\mathbf{A} & \mathbf{0} & \mathbf{0} \\ g_{21}\mathbf{A} & g_{22}\mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & r_{11}\mathbf{I} & r_{12}\mathbf{I} \\ \mathbf{0} & \mathbf{0} & r_{21}\mathbf{I} & r_{22}\mathbf{I} \end{bmatrix}$$
- Zdefiniować macierz $\mathbf{Q}$: $\mathbf{Q}\mathbf{G}_0\mathbf{Q}^T = \mathbf{W}$ oraz $\mathbf{Q}\mathbf{R}_0\mathbf{Q}^T = \mathbf{I}$
 $\mathbf{W} \rightarrow$ macierz diagonalna
 $\mathbf{I} \rightarrow$ macierz jednostkowa
- Wektor nieskorelowanych obserwacji: $\mathbf{y}^* = \mathbf{Q}\mathbf{y}$

$$\text{var}(\mathbf{y}^*) \sim \mathbf{W} + \mathbf{I} \qquad \text{var}[\mathbf{y}^*] = \begin{bmatrix} w_{11}\mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & w_{22}\mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & i_{11}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & i_{22}\mathbf{I} \end{bmatrix}$$

 $\text{var}(\mathbf{y}^*) \rightarrow$ macierz diagonalna = brak korelacji $\mathbf{y}^*$

Transformacja kanoniczna

- Model **jednocechowy** dla transformowanej cechy i

$$\mathbf{y}_i^* = \mathbf{X}\mathbf{b}_i^* + \mathbf{Z}\mathbf{a}_i^* + \mathbf{e}_i^*$$

- MME

$$\begin{bmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z} \\ \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} + A^{-1}\alpha_i \end{bmatrix} \begin{bmatrix} \mathbf{b}_i^* \\ \mathbf{a}_i^* \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T\mathbf{y}_i^* \\ \mathbf{Z}^T\mathbf{y}_i^* \end{bmatrix}$$

$$\alpha = \frac{1}{w_{ii}} \quad \mathbf{W} = \begin{bmatrix} w_i & 0 \\ 0 & w_{ii} \end{bmatrix}$$

$$\text{var}(\mathbf{a}_i^*) = \mathbf{A}w_{ii} \quad \text{var}(\mathbf{e}_i^*) = \mathbf{I} \quad \text{var}(\mathbf{y}_i^*) = \mathbf{Z}\mathbf{A}\mathbf{Z}^T w_{ii} + \mathbf{I}$$

Transformacja kanoniczna

- Transformacja wsteczna

$$b_i = Q^{-1} b_i^*$$

$$a_i = Q^{-1} a_i^*$$

Transformacja kanoniczna - przykład

Dane

Calf	Sex	Sire	Dam	WWG	PWG
4	Male	1	—	4.5	6.8
5	Female	3	2	2.9	5.0
6	Female	1	2	3.9	6.8
7	Male	4	5	3.5	6.0
8	Male	3	6	5.0	7.5

WWG, pre-weaning gain; PWG, post-weaning gain.

$$\mathbf{G}_0 = \begin{bmatrix} 20 & 18 \\ 18 & 40 \end{bmatrix} \quad \mathbf{R}_0 = \begin{bmatrix} 40 & 11 \\ 11 & 30 \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} 0.17 & -0.08 \\ 0.02 & 0.18 \end{bmatrix}$$

Transformacja kanoniczna - przykład

Kroki wyznaczania Q

1. Obliczyć wartości i wektory własne macierzy R_0

www.youtube.com/watch?v=HH8pouRwphA

$$R_0 = UBU^T \quad R_0 = \begin{bmatrix} 40 & 11 \\ 11 & 30 \end{bmatrix}$$

$$\text{Wartości własne } \text{diag}(\mathbf{B}) = \mathbf{B} = \begin{bmatrix} 47.08 & 0 \\ 0 & 22.92 \end{bmatrix}$$

$$\text{Wektory własne } \mathbf{U} = \begin{bmatrix} 0.84 & -0.54 \\ 0.54 & 0.84 \end{bmatrix}$$

Transformacja kanoniczna - przykład

Kroki wyznaczania **Q**

$$B = \begin{bmatrix} 47.08 & 0 \\ 0 & 22.92 \end{bmatrix} \quad U = \begin{bmatrix} 0.84 & -0.54 \\ 0.54 & 0.84 \end{bmatrix}$$

2. Wyznaczyć macierz **P**

$$P = U\sqrt{B^{-1}}U^T = \begin{bmatrix} 0.16 & -0.03 \\ -0.03 & 0.19 \end{bmatrix}$$

3. Wyznaczyć macierz **P****G**₀**P**^T

$$G_0 = \begin{bmatrix} 20 & 18 \\ 18 & 40 \end{bmatrix} \quad PG_0P^T = \begin{bmatrix} 0.40 & 0.26 \\ 0.26 & 1.27 \end{bmatrix}$$

Transformacja kanoniczna - przykład

Kroki wyznaczania Q

4. Obliczyć wartości i wektory własne macierzy PG_0P^T

$$PG_0P^T = LWL^T$$

$$W = \begin{bmatrix} 0.33 & 0 \\ 0 & 1.34 \end{bmatrix} \quad L = \begin{bmatrix} 0.96 & 0.27 \\ -0.27 & 0.96 \end{bmatrix}$$

5. Wyznaczyć macierz Q

$$P = \begin{bmatrix} 0.16 & -0.03 \\ -0.03 & 0.19 \end{bmatrix} \quad Q = L^T P = \begin{bmatrix} 0.17 & -0.08 \\ 0.02 & 0.18 \end{bmatrix}$$

Transformacja kanoniczna - przykład

Transformacja cech

Table 5.2. Pre-weaning gain and post-weaning gain for beef calves on the original and transformed scales.

Calf	Sex	Sire	Dam	Original scale		Transformed scale	
				WWG	PWG	VAR1	VAR2
4	Male	1	—	4.5	6.8	0.208	1.269
5	Female	3	2	2.9	5.0	0.085	0.926
6	Female	1	2	3.9	6.8	0.109	1.259
7	Male	4	5	3.5	6.0	0.106	1.112
8	Male	3	6	5.0	7.5	0.236	1.400

VAR1 and VAR2, transformed variables for WWG and PWG, respectively.

Np. dla osobnika 4

$$y_4^* = \mathbf{Q} y_4 = \begin{bmatrix} 0.17 & -0.08 \\ 0.02 & 0.18 \end{bmatrix} \begin{bmatrix} 4.5 \\ 6.8 \end{bmatrix} = \begin{bmatrix} 0.208 \\ 1.269 \end{bmatrix}$$

Transformacja kanoniczna - przykład

Rozwiązania MME dla WWG*

$$\begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hat{a}_6 \\ \hat{a}_7 \\ \hat{a}_8 \end{bmatrix} = \begin{bmatrix} 3.0 & 0.0 & 0.000 & 0.000 & 0.000 & 1.000 & 0.000 & 0.000 & 1.000 & 1.000 \\ 0.0 & 2.0 & 0.000 & 0.000 & 0.000 & 0.000 & 1.000 & 1.000 & 0.000 & 0.000 \\ 0.0 & 0.0 & 5.584 & 1.523 & 0.000 & -2.031 & 0.000 & -3.046 & 0.000 & 0.000 \\ 0.0 & 0.0 & 1.523 & 6.092 & 1.523 & 0.000 & -3.046 & -3.046 & 0.000 & 0.000 \\ 0.0 & 0.0 & 0.000 & 1.523 & 6.092 & 0.000 & -3.046 & 1.523 & 0.000 & -3.046 \\ 1.0 & 0.0 & -2.031 & 0.000 & 0.000 & 6.584 & 1.523 & 0.000 & -3.046 & 0.000 \\ 0.0 & 1.0 & 0.000 & -3.046 & -3.046 & 1.523 & 8.615 & 0.000 & -3.046 & 0.000 \\ 0.0 & 1.0 & -3.046 & -3.046 & 1.523 & 0.000 & 0.000 & 8.615 & 0.000 & -3.046 \\ 1.0 & 0.0 & 0.000 & 0.000 & 0.000 & -3.046 & -3.046 & 0.000 & 7.092 & 0.000 \\ 1.0 & 0.0 & 0.000 & 0.000 & -3.046 & 0.000 & 0.000 & -3.046 & 0.000 & 7.092 \end{bmatrix}^{-1} \begin{bmatrix} 0.549 \\ 0.194 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.208 \\ 0.085 \\ 0.108 \\ 0.105 \\ 0.235 \end{bmatrix}$$

estymatory dla cech transformowanych

Transformacja kanoniczna - przykład

Rozwiązania: mBLUB

Effects	Multivariate analysis traits	
	WWG	PWG
<i>Sex</i> *		
1	4.361	6.800
2	3.397	5.880
<i>Animal</i>		
1	0.151	0.280
2	-0.015	-0.008
3	-0.078	-0.170
4	-0.010	-0.013
5	-0.270	-0.478
6	0.276	0.517
7	-0.316	-0.479
8	0.244	0.392

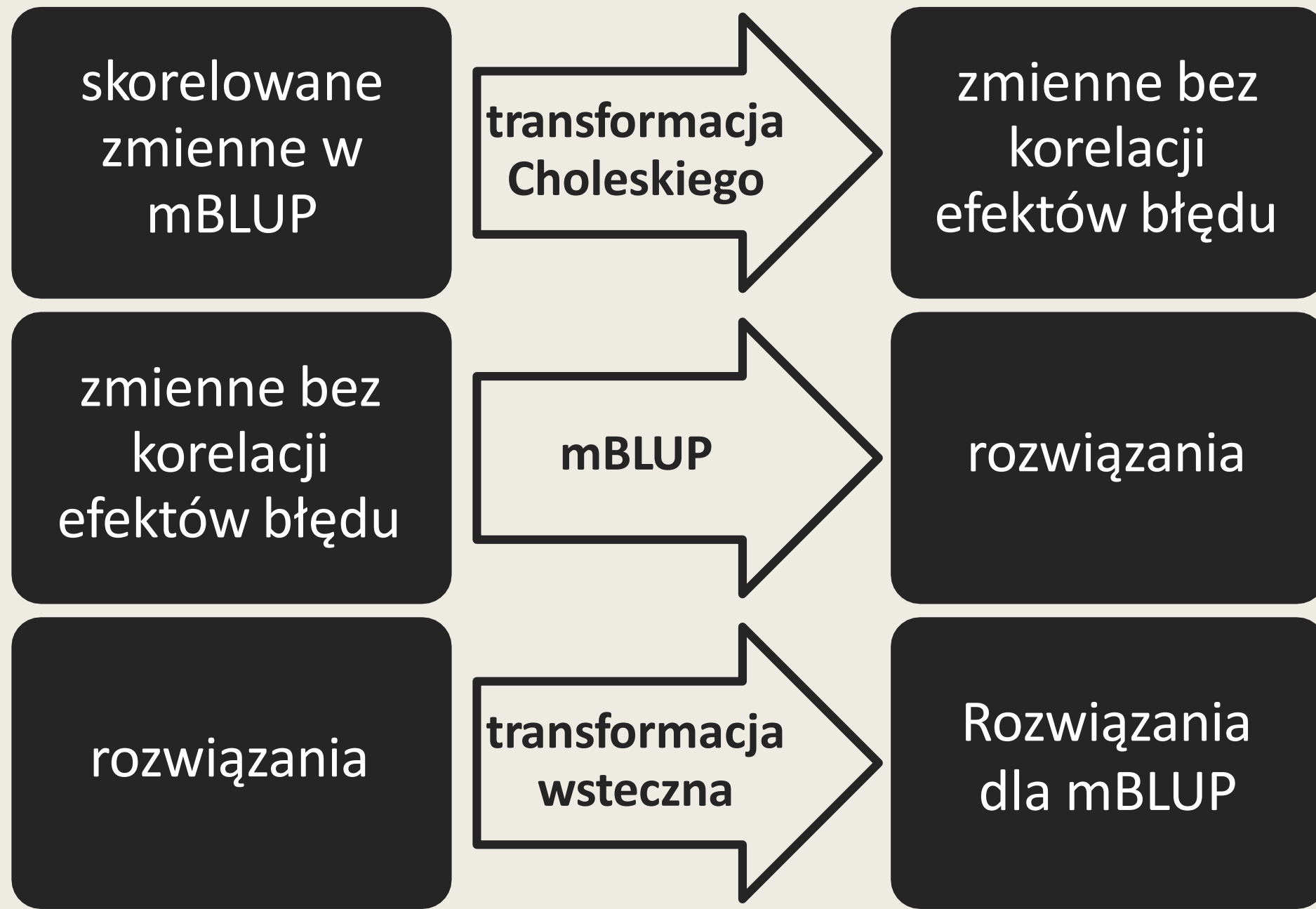
Rozwiązania – transformacja kanoniczna

Effects	Canonical scale		Original scale	
	VAR1	VAR2	WWG	PWG
<i>Sex</i>				
1	0.185	1.266	4.361	6.800
2	0.098	1.089	3.397	5.880
<i>Animal</i>				
1	0.003	0.052	0.151	0.280
2	-0.002	-0.002	-0.015	-0.008
3	0.000	-0.031	-0.078	-0.170
4	-0.001	-0.002	-0.010	-0.013
5	-0.007	-0.088	-0.270	-0.478
6	0.005	0.095	0.276	0.517
7	-0.015	-0.089	-0.316	-0.479
8	0.009	0.073	0.244	0.392

Np. dla osobnika 1

$$a_1 = Q^{-1} a_1^* = \begin{bmatrix} 5.77 & 2.60 \\ -0.55 & 5.45 \end{bmatrix} \begin{bmatrix} 0.003 \\ 0.052 \end{bmatrix} = \begin{bmatrix} 0.151 \\ 0.280 \end{bmatrix}$$

Transformacja Choleskiego



Transformacja Choleskiego

- Quaas (1984)
- Brakujące wartości cech

Transformacja Choleskiego

- Wektor skorelowanych obserwacji: $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$$\text{var}(\mathbf{y}) \sim \mathbf{G}_0 + \mathbf{R}_0$$

- Przeprowadzić dekompozycję Choleskiego i zdefiniować macierz

$$\mathbf{T}: \mathbf{R}_0 = \mathbf{T}\mathbf{T}^T$$

$\mathbf{T}$: lower triangular $\mathbf{T}^T$: upper triangular

- Wektor obserwacji o nieskorelowanych błędach: $\mathbf{y}^* = \mathbf{T}^{-1}\mathbf{y}$

$$\text{var}(\mathbf{y}^*) \sim \mathbf{T}^{-1}\mathbf{G}_0\mathbf{T}^{-1T} + \mathbf{I} = \mathbf{M} + \mathbf{I}$$

w przypadku brakujących cech $y_i^* = 0$

Transformacja Choleskiego

- Model wielocechowy dla cech transformowanych

$$\begin{bmatrix} y_1^* \\ y_2^* \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} b_1^* \\ b_2^* \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} + \begin{bmatrix} e_1^* \\ e_2^* \end{bmatrix}$$

$$\text{var}(y^*) \sim M + I$$

$$M = \begin{bmatrix} m_{ii} & m_{ij} \\ m_{ij} & m_{ii} \end{bmatrix} \quad M^{-1} = \begin{bmatrix} m^{ii} & m^{ij} \\ m^{ij} & m^{ii} \end{bmatrix}$$

- MME

$$\begin{bmatrix} X_1^T X_1 & 0 & X_1^T Z_1 & 0 \\ 0 & X_2^T X_2 & 0 & X_2^T Z_2 \\ Z_1^T X_1 & 0 & Z_1^T Z_1 + A^{-1} m^{11} & Z_1^T Z_2 + A^{-1} m^{12} \\ 0 & Z_2^T X_2 & A^{-1} m^{21} & Z_2^T Z_2 + A^{-1} m^{22} \end{bmatrix} \begin{bmatrix} b_1^* \\ b_2^* \\ a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} X_1^T y_1^* \\ X_2^T y_2^* \\ Z_1^T y_1^* \\ Z_2^T y_2^* \end{bmatrix}$$

Transformacja Choleskiego

- Standardowy model wielocechowy

$$\begin{bmatrix} X_1^T R_{11} X_1 & X_1^T R_{12} X_2 & X_1^T R_{11} Z_1 & X_1^T R_{12} Z_2 \\ X_2^T R_{21} X_1 & X_2^T R_{22} X_2 & X_2^T R_{21} Z_1 & X_2^T R_{22} Z_2 \\ Z_1^T R_{11} X_1 & Z_1^T R_{12} X_2 & Z_1^T R_{22} Z_1 + A^{-1} g^{11} & Z_1^T R_{12} Z_2 + A^{-1} g^{12} \\ Z_2^T R_{21} X_1 & Z_2^T R_{22} X_2 & Z_2^T R_{22} Z_1 + A^{-1} g^{21} & Z_2^T R_{22} Z_2 + A^{-1} g^{22} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} X_1^T y_1 \\ X_2^T y_2 \\ Z_1^T y_1 \\ Z_2^T y_2 \end{bmatrix}$$

- Model wielocechowy dla cech transformowanych

$$\begin{bmatrix} X_1^T X_1 & 0 & X_1^T Z_1 & 0 \\ 0 & X_2^T X_2 & 0 & X_2^T Z_2 \\ Z_1^T X_1 & 0 & Z_1^T Z_1 + A^{-1} m^{11} & Z_1^T Z_2 + A^{-1} m^{12} \\ 0 & Z_2^T X_2 & A^{-1} m^{21} & Z_2^T Z_2 + A^{-1} m^{22} \end{bmatrix} \begin{bmatrix} b_1^* \\ b_2^* \\ a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} X_1^T y_1^* \\ X_2^T y_2^* \\ Z_1^T y_1^* \\ Z_2^T y_2^* \end{bmatrix}$$

Transformacja Choleskiego

- Transformacja wsteczna

$$\mathbf{b}_i = \mathbf{T} \mathbf{b}_i^*$$

$$\mathbf{a}_i = \mathbf{T} \mathbf{a}_i^*$$

Transformacja Choleskiego - przykład

Dane

Calf	Sex	Sire	Dam	WWG (kg)	PWG (kg)
4	Male	1	–	4.5	–
5	Female	3	2	2.9	5.0
6	Female	1	2	3.9	6.8
7	Male	4	5	3.5	6.0
8	Male	3	6	5.0	7.5
9	Female	7	8	4.0	–

$$R_m = 40 \quad R_o = \begin{bmatrix} 40 & 11 \\ 11 & 30 \end{bmatrix} \quad G_0 = \begin{bmatrix} 20 & 18 \\ 18 & 40 \end{bmatrix} \quad T = \begin{bmatrix} 6.3 & 0 \\ 1.7 & 5.2 \end{bmatrix}$$

Transformacja Choleskiego - przykład

Wyznaczanie **T** → Cholesky decomposition

www.youtube.com/watch?v=TprfUB3nI8Y

$$R_0 = \mathbf{T}\mathbf{T}^T$$

- Elementy diagonalne: $t_{ii} = \sqrt{r_{ii} - \sum_{j=1}^{i-1} t_{ij}^2}$
- Elementy dolnej poza-przekątnej:

$$t_{ik} = \frac{1}{t_k} \left(r_{ik} - \sum_{j=1}^{k-1} t_{ij} t_{kj} \right)$$

Transformacja Choleskiego - przykład

Transformacja cech

Calf	Sex	Sire	Dam	Original traits		Transformed traits	
				WWG	PWG	y_1^*	y_2^*
4	Male	1	—	4.5	—	0.712	—
5	Female	3	2	2.9	5.0	0.456	0.809
6	Female	1	2	3.9	6.8	0.617	1.103
7	Male	4	5	3.5	6.0	0.553	0.970
8	Male	3	6	5.0	7.5	0.791	1.179
9	Female	7	8	4.0	—	0.632	—

Np. dla osobnika 4, cecha $i=1$

$$y_{14}^* = t^{11} y_{14}$$

Np. dla osobnika 5, cechy $i=1$ oraz $j=2$

$$y_{15}^* = t^{11} y_{15}$$

$$y_{25}^* = t^{22} y_{25} + t^{12} y_{15}$$

Transformacja Choleskiego - przykład

Rozwiązania: mBLUB

Effects	Multivariate analysis	
	WWG	PWG
<i>Sex</i>		
1	4.327	6.793
2	3.598	5.966
<i>Animal</i>		
1	0.154	0.288
2	-0.059	-0.053
3	-0.061	-0.163
4	0.027	0.036
5	-0.307	-0.521
6	0.235	0.477
7	-0.280	-0.452
8	0.272	0.407
9	0.077	0.051

Rozwiązania – transformacja Choleskiego

Effects	Transformed scale		Original scale	
	WWG	PWG	WWG	PWG
<i>Sex</i>				
1	0.684	1.079	4.327	6.793
2	0.569	0.958	3.598	5.966
<i>Animal</i>				
1	0.024	0.047	0.154	0.288
2	-0.009	-0.007	-0.059	-0.053
3	-0.010	-0.028	-0.061	-0.163
4	0.004	0.005	0.027	0.036
5	-0.048	-0.084	-0.307	-0.521
6	0.037	0.079	0.235	0.477
7	-0.044	-0.072	-0.280	-0.452
8	0.043	0.064	0.272	0.407
9	0.012	0.006	0.077	0.051

Dziękuję za uwagę