Universidade de Brasília

DEPARTAMENTO DE ESTATÍSTICA

20 abril 2023

Lista 2

Prof. Dr. George von Borries

Análise Multivariada 1

Aluno: Bruno Gondim Toledo | Matrícula: 15/0167636

8. Provar o seguinte teorema: Sejam **A** e **B** matrizes idempotentes. Então,

(a) **A** + **B** é idempotente somente quando **AB** = **BA** = 0.

(b) **C** = **AB** é idempotente somente quando **AB** = **BA**.

(c) **I - A** é idempotente.

9. Provar o seguinte teorema: Seja $X_{(nÖk)}$ tal que rank(**X**) $= k < n$. Então, $P_X = X(X'X)^{-1}X'$ é idempotente e simétrica e consequentemente, uma matriz projeção ortogonal.

10. Utilizando o R verifique, através de exemplos, que uma matriz de projeção tem autovalores somente no conjunto {0, 1}. A demonstração pode ser feita utilizando a equação característica e lembrando que se **M** é uma matriz de projeção, então:

$$M = M^2 = M^T$$

11. Seja **X** uma matriz de dados $(nÖp)$ com matriz de covariância **S**. Sejam $\lambda_1, ..., \lambda_p$ os autovalores de **S**.

(a) Mostre que a soma das variâncias s_{ii} de **X** (variação amostral total) é dada por $\lambda_1 + ... + \lambda_p$.

(b) Mostre que a variância amostral generalizada é dada por $\lambda_1 \times ... \times \lambda_p$.

(c) Mostre que a variância amostral generalizada se anula se as colunas de **X** somarem zero.

12. Seja **A** uma matriz quadrada $(kÖk)$ positiva definida. Mostre que,

(a) $(A^{1/2})' = A^{1/2}$.

(b) $A^{1/2}A^{1/2} = A$.

(c) $(A^{1/2})^{-1} = CD^{-1/2}C'$, sendo $D^{-1/2} = diag(1/\sqrt{\lambda_1}, ..., 1/\sqrt{\lambda_k})$.

(d) $A^{1/2}A^{-1/2} = A^{-1/2}A^{1/2} = I$.

(e) $A^{-1/2}A^{-1/2} = A^{-1}$.

13. Considere uma matriz de correlação $(rÖr)$ com a mesma correlação (ρ) em todas as células fora da diagonal. Encontre os autovalores e autovetores desta matriz quando $r = 2, 3, 4$. Generalize seus resultados para qualquer número r de variáveis. Como exemplo, faça $\rho = 0.1, 0.3, 0.5, 0.7, 0.9$.

14. Considere a decomposição espectral de uma matrix $A_{p\times p}$ positiva definida, isto é, $A_{p\times p} = B\Delta B^T$. Seja

$$A = \begin{bmatrix} 3 & 2 & 3 & 2 \\ 2 & 5 & 1 & 1 \\ 3 & 1 & 8 & 2 \\ 2 & 1 & 2 & 3 \end{bmatrix}$$

(a) Obtenha **B** e Δ.

(b) Obtenha $A^{1/2}$ e mostre que $(A^{1/2})^2 = A$.

(c) Obtenha $(A^{1/2})^{-1}$ (descreva seus elementos) e mostre que $(A^{1/2})^{-1}A^{1/2} = I$.

15. Let $x' = [5, 1, 3]$ and $y' = [-1, 3, 1]$.

(a) Graph the two vectors.

(b) Find (i) the length of x, (ii) the angle between x and y, and (iii) the projection of y on x.

(c) Since $\bar{x} = 3$ and $\bar{y} = 1$, graph $[5-3, 1-3, 3-3] = [2, -2, 0]$ and $[-1-1, 3-1, 1-1] = [-2, 2, 0]$.

16. Given the matrices

$$A = \begin{bmatrix} -1 & 3 \\ 4 & 2 \end{bmatrix}$$

$$B = \begin{bmatrix} 4 & -3 \\ 1 & -2 \\ -2 & 0 \end{bmatrix}$$

and

$$C = \begin{bmatrix} 5 \\ -4 \\ 2 \end{bmatrix}$$

perform the indicated multiplications.

(a) $5A$

(b) BA

(c) $A'B'$

(d) $C'B$

(e) Is **AB** defined?

17. Let

$$A = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$$

(a) Is **A** symmetric?

(b) Show that **A** is positive definite.

18. Let **A** be as given in previous exercise.

(a) Determine the eigenvalues and eigenvectors of **A**.

(b) Write the spectral decomposition of **A**.

(c) Find A^{-1}.

(d) Find the eigenvalues and eigenvectors of A^{-1}

19. Given the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$$

find the eigenvalues λ_1 and λ_2 and the associated normalized eigenvectors e_1 and e_2. Determine the spectral decomposition (2-16) of **A**.

20. Consider an arbitrary $n \times p$ matrix **A**. Then $A'A$ is a symmetric $p \times p$ matrix. Show that $A'A$ is necessarily nonnegative definite. Hint: Set $y = A\ x$ so that $y'y = x'A'A\ x$

21. Consider the set of points (x_1, x_2) whose "distances" from the Origin are given by

$$c^2 = 4x_1^2 + 3x_2^2 - 2\sqrt{2}x_1x_2$$

for $c^2 = 1$ and for $c^2 = 4$. Determine the major and minor axes of the ellipses of constant distances and their associated lengths. Sketch the ellipses of constant distances and comment on their positions. What will happen as c^2 increases?

22. Let $A^{1/2}_{(m\times m)} = \sum_{i=1}^{m} \sqrt{\lambda_i} e_i e_i' = P\Lambda^{1/2}P'$, where $PP' = P'P = I$.(The $\lambda_i's$ and the $e_i's$ are the eigenvalues and associated normalized eigenvectors of the matrix A.) Show properties (1)-(4) of the square-root matrix in (2-22).

23. (See result 2A.15) Using the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 2 & -2 \\ 2 & 2 \end{bmatrix}$$

(a) Calculate $A'A$ and obtain its eigenvalues and eigenvectors.

(b) Calculate AA' and obtain its eigenvalues and eigenvectors. Check that the nonzero eigenvalues are the same as those in part a.

(c) Obtain the singular-value decomposition of **A**

24. Let **X** have covariance matrix

$$\Sigma = \begin{bmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{bmatrix}$$

(a) Determine ρ and $V^{1/2}$.

(b) Multiply your matrices to check the relation $V^{1/2}\rho V^{1/2} = \Sigma$.

Nota: ρ é a matriz de correlação populacional e **V** é a matriz diagonal de variâncias. Para calcular ρ é mais fácil fazer $\rho = (V^{1/2})^{-1} \sum (V^{1/2})^{-1}$, uma vez que **V** é uma matriz diagonal, i.e., $V = diag[\sqrt{\sigma_{11}}, ..., \sqrt{\sigma_{11}}]$.

25. Derive exoressions for the mean and variance of the following linear combinations in terms of the mean and covariances of the random variable $X_1, X_2,\ and\ X_3$.

(a) $X_1 - 2X_2$

(b) $-X_1 + 3X_2$

(c) $X_1 + X_2 + X_3$

(d) $X_1 + 2X_2 - X_3$

(e) $3X_1 - 4X_2$ if X_1 and X_2 are independent random variables.

26. You are given the random vector $X' = [X_1, X_2, ..., X_5]$ with mean vector $\mu'_x = [2, 4, -1, 3, 0]$ and variance-covariance matrix

$$\Sigma_x = \begin{bmatrix} 4 & -1 & \frac{1}{2} & -\frac{1}{2} & 0 \\ -1 & 3 & 1 & -1 & 0 \\ \frac{1}{2} & 1 & 6 & 1 & -1 \\ -\frac{1}{2} & -1 & 1 & 4 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{bmatrix}$$

Partition **X** as

$$X = \begin{bmatrix} X_1 \\ X_2 \\ ..... \\ X_3 \\ X_4 \\ X_5 \end{bmatrix}$$

=

$$\begin{bmatrix} X^{(1)} \\ ..... \\ X^{(2)} \end{bmatrix}$$

Let

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

and consider the linear combinations $AX^{(1)}$ *and* $BX^{(2)}$. Find

(a) $E(X^{(1)})$

(b) $E(AX^{(1)})$

(c) $Cov(X^{(1)})$

(d) $Cov(AX^{(1)})$

(e) $E(X^{(2)})$

(f) $E(BX^{(2)})$

(g) $Cov(X^{(2)})$

(h) $Cov(BX^{(2)})$

(i) $Cov(X^{(1)}, X^{(2)})$

(J) $Cov(AX^{(1)}, BX^{(2)})$

27. You are given the random vector $X' = [X_1, X_2, X_3, X_4]$ with mean vector $\mu_x' = [3, 2, -2, 0]$ and variance-covariance matrix

$$\Sigma_x = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Let

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

(a) Find E($\mathbf{AX}$), the mean of $\mathbf{AX}$.

Utilizando a propriedade da linearidade da esperança, e pelo dado que $E(X) = \mu_x'$:

$$E(\mathbf{AX}) = \mathbf{A}E(\mathbf{X}) = \mathbf{A}\mu_x'$$

Então,

$$\mathbf{A}\mu_x' = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1(3) + (-1)(2) + 0(-2) + 0(0) \\ 1(3) + 1(2) + (-2)(-2) + 0(0) \\ 1(3) + 1(2) + 1(-2) + (-3)(0) \end{bmatrix} = \begin{bmatrix} -1 & 6 & 0 \end{bmatrix}$$

(b) Find Cov($\mathbf{AX}$), the variances and covariances of $\mathbf{AX}$

(c) Which pairs of linear combinations have zero covariances?

28. Considere o seguinte conjunto de dados de Pacientes em Tratamento de Hemodiálise.

Idade	Proteína	Energia	Albumina	IMC
32	1.59	2738.86	4.2	24.1
61	0.49	824.26	3.9	29.8
51	1.14	1307.03	4.1	20.0
53	0.74	925.47	4.2	25.0
24	1.99	2787.46	3.8	21.5
65	1.00	1222.51	4.2	25.0
35	2.32	2038.28	4.1	18.7
45	0.93	1061.53	4.2	22.0
57	0.81	1657.73	4.2	31.2
32	1.23	1652.76	3.9	24.3
66	0.99	1636.25	4.1	27.7
27	1.40	1845.07	4.0	21.8
54	1.08	1542.30	3.9	29.0
55	1.22	1214.53	4.0	21.1
50	0.57	1451.17	4.0	27.1
48	0.83	1786.95	4.1	24.7
28	1.55	1975.26	3.5	18.8
66	1.10	1248.64	4.0	18.9
66	0.44	987.86	4.0	27.6
48	0.58	1067.10	4.3	26.4
60	0.43	968.62	4.0	35.9
59	0.66	836.94	3.9	25.3
50	1.81	1197.99	3.9	19.5
29	1.21	1818.31	4.2	21.8
40	0.98	1238.91	3.5	21.9
47	1.48	2153.47	3.5	17.3
52	0.98	1720.60	3.6	29.7
54	1.02	1906.30	4.5	31.9
53	0.82	981.85	3.9	26.2
47	0.46	1020.95	4.4	31.2
42	1.34	1028.10	3.6	18.1
79	1.48	1465.91	3.9	18.3
61	1.39	1456.12	3.9	24.9

(a) Represente graficamente e através de medidas descritivas.

Matriz de covariâncias:

	Idade	Proteína	Energia	Albumina	IMC
Idade	174.9393939	-2.8203693	-3736.92562	0.5121212	19.3448864
Proteína	-2.8203693	0.2079267	152.26715	-0.0321449	-1.4953153
Energia	-3736.9256250	152.2671531	249960.54006	-11.0733438	-573.0939687
Albumina	0.5121212	-0.0321449	-11.07334	0.0613258	0.4797727
IMC	19.3448864	-1.4953153	-573.09397	0.4797727	22.0800568

Matriz de variâncias:

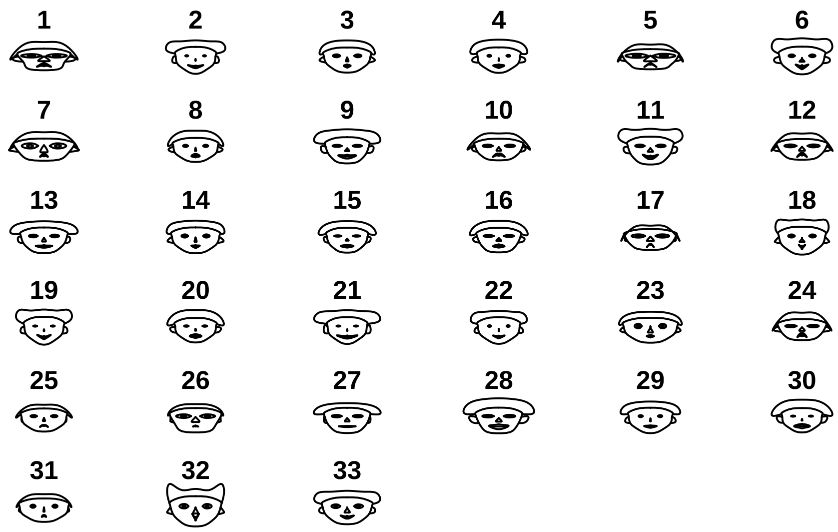
	Idade	Proteína	Energia	Albumina	IMC
Idade	174.9393939	-2.8203693	-3736.92562	0.5121212	19.3448864
Proteína	-2.8203693	0.2079267	152.26715	-0.0321449	-1.4953153
Energia	-3736.9256250	152.2671531	249960.54006	-11.0733438	-573.0939687
Albumina	0.5121212	-0.0321449	-11.07334	0.0613258	0.4797727

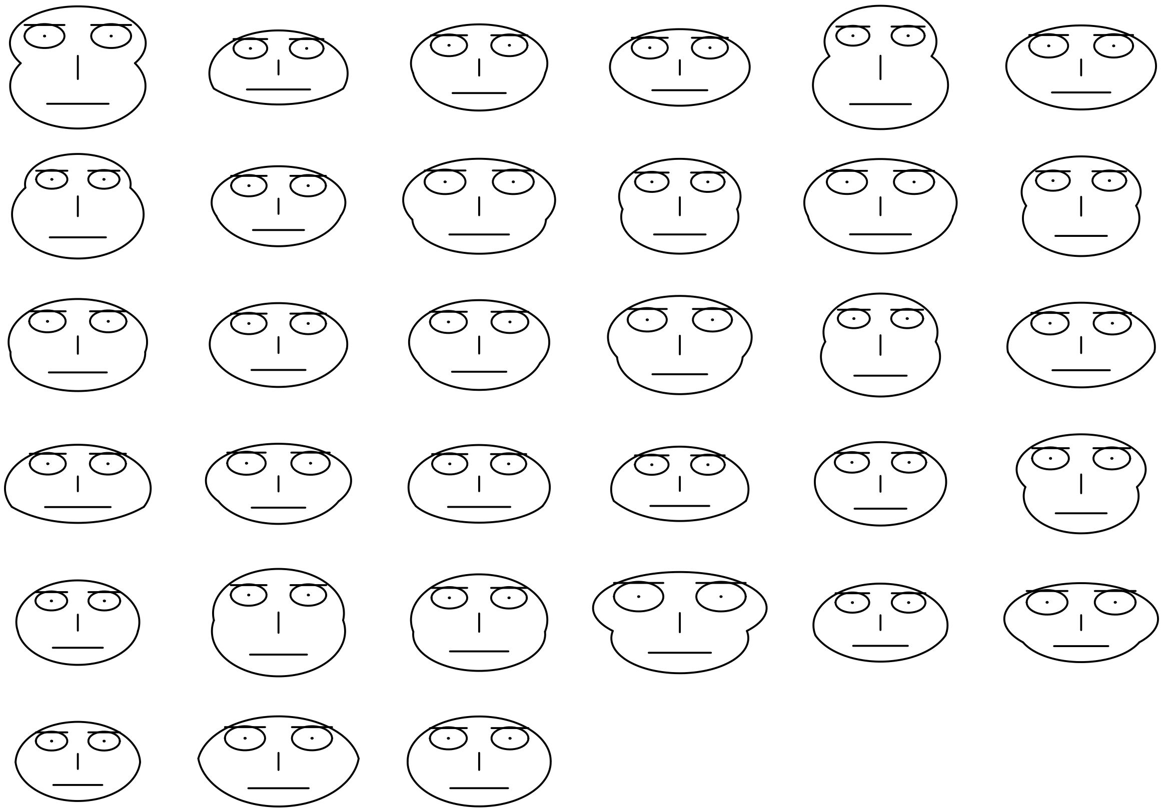
	Idade	Proteína	Energia	Albumina	IMC
IMC	19.3448864	-1.4953153	-573.09397	0.4797727	22.0800568

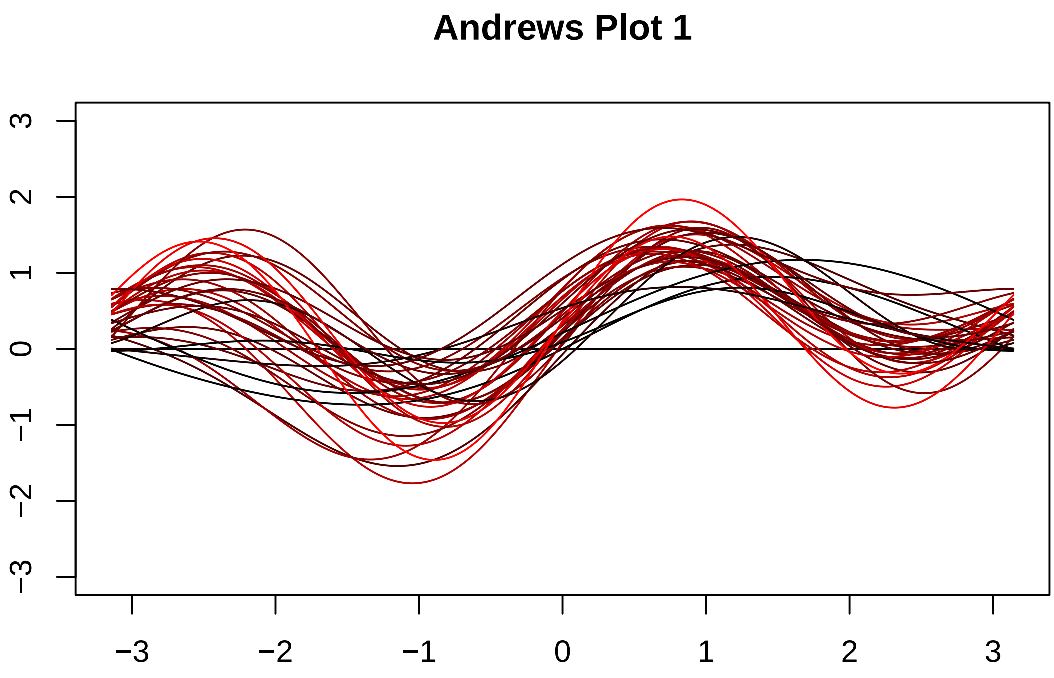
Matriz de correlações:

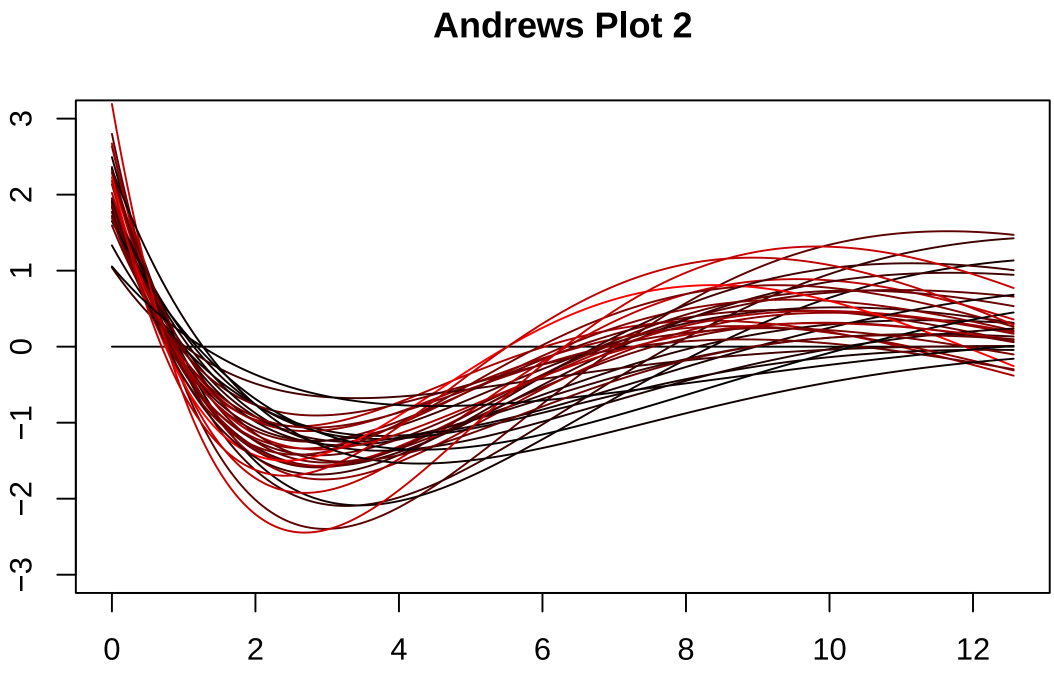
	Idade	Proteína	Energia	Albumina	IMC
Idade	1.0000000	-0.4676350	-0.5651125	0.1563535	0.3112593
Proteína	-0.4676350	1.0000000	0.6679060	-0.2846658	-0.6978749
Energia	-0.5651125	0.6679060	1.0000000	-0.0894379	-0.2439439
Albumina	0.1563535	-0.2846658	-0.0894379	1.0000000	0.4123006
IMC	0.3112593	-0.6978749	-0.2439439	0.4123006	1.0000000

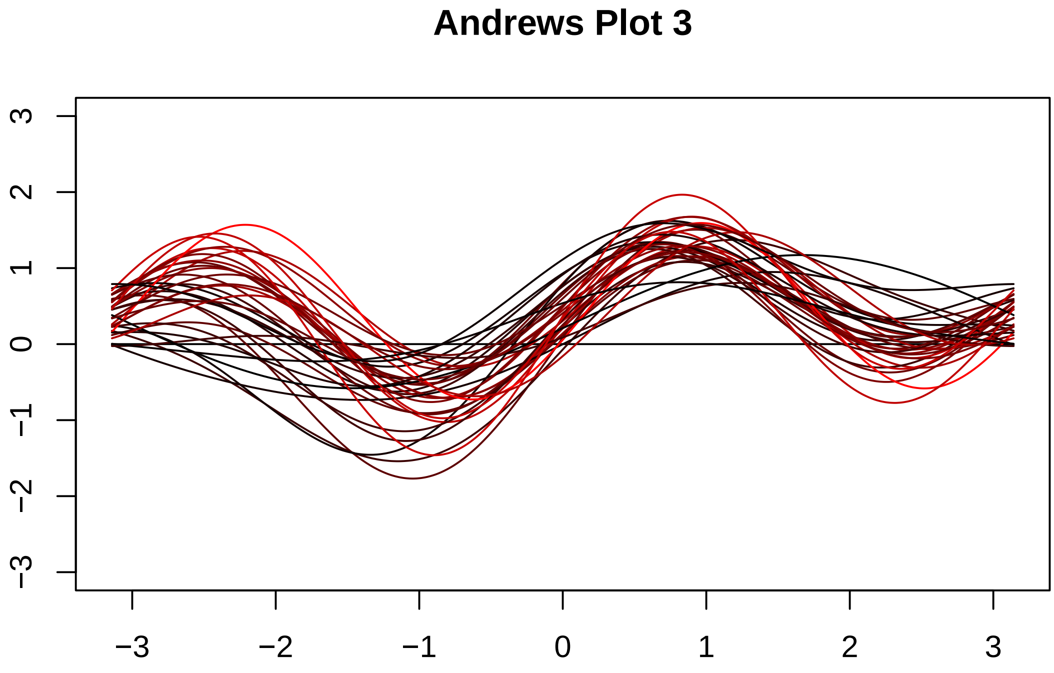
Agora, representando graficamente:

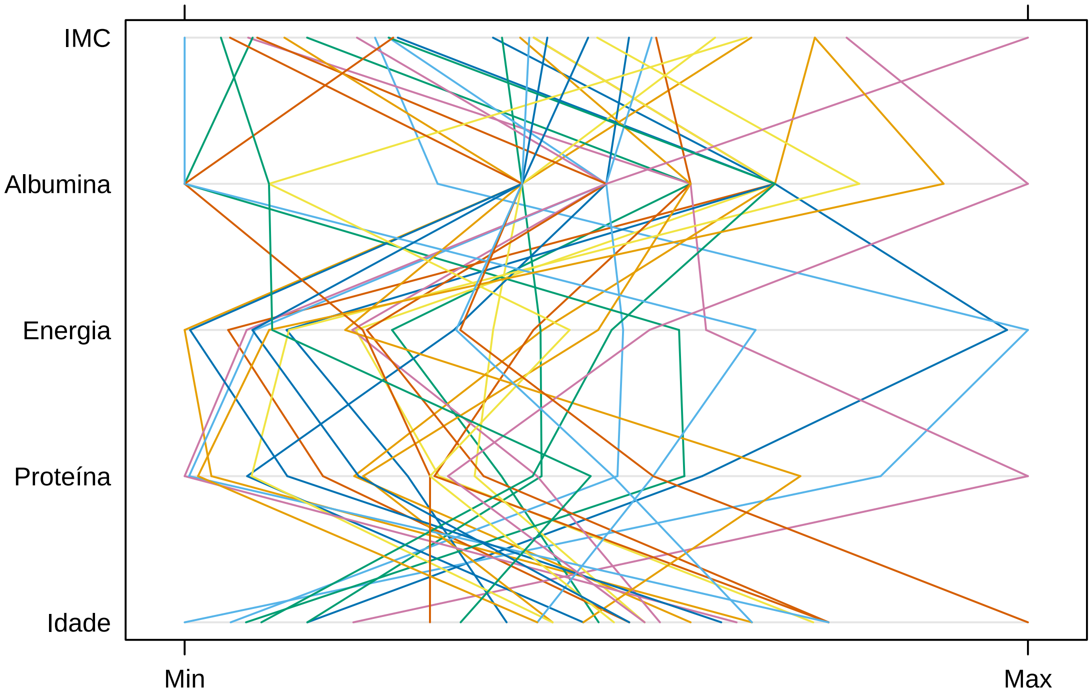


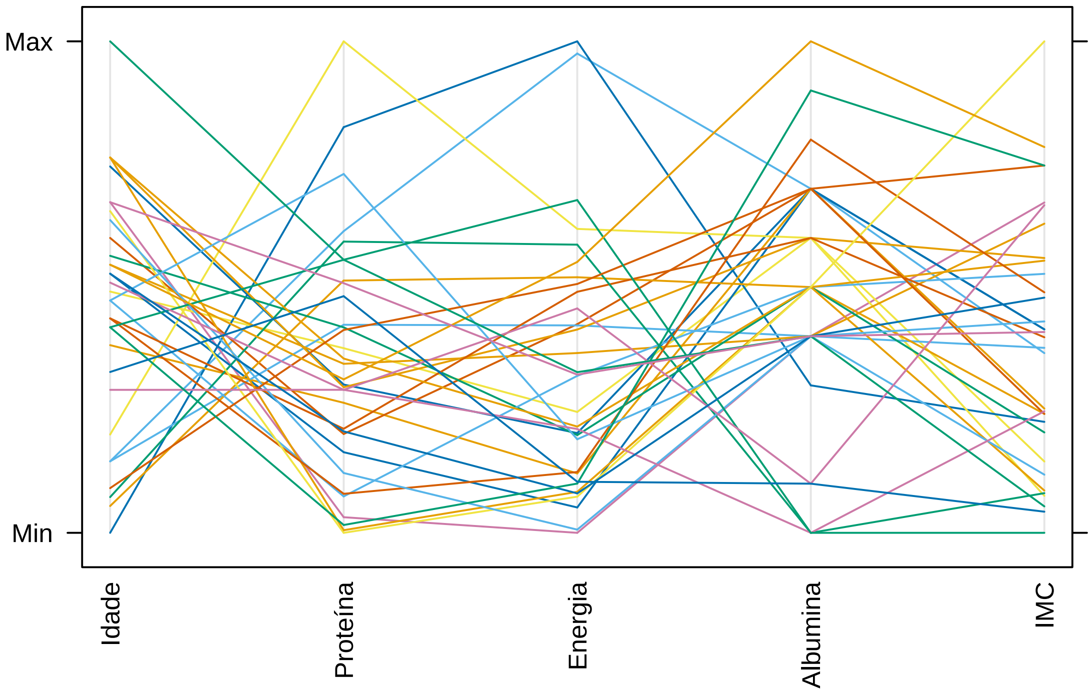


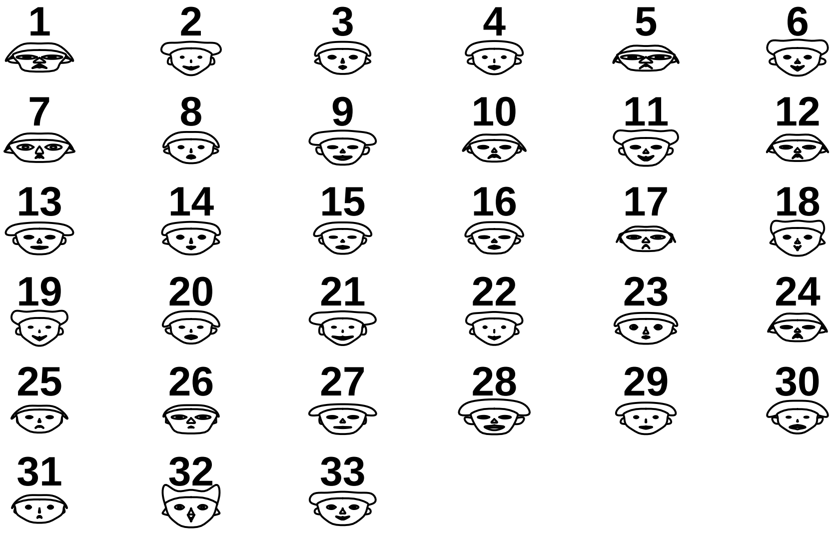

Andrews Plot 1
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3


Andrews Plot 2
3
2
1
0
−1
−2
−3
0
2
4
6
8
10
12

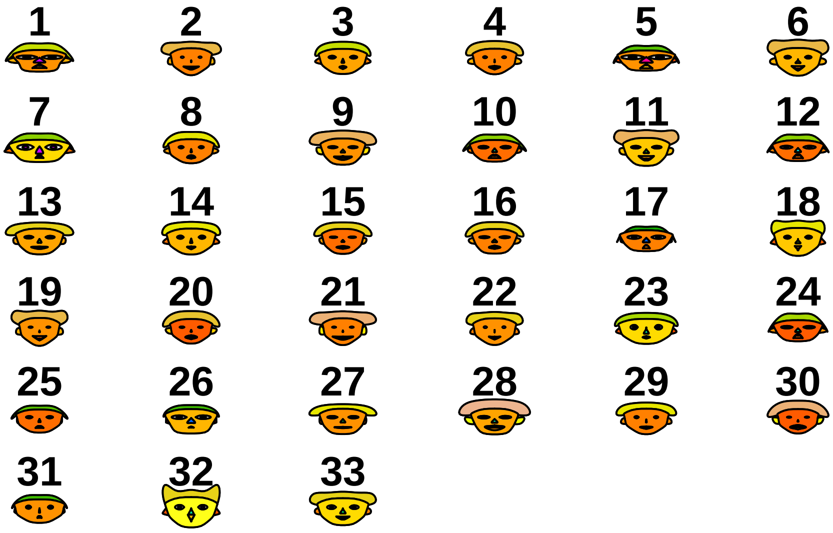

Andrews Plot 3
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3


IMC
Albumina
Energia
Proteína
Idade
Min
Max

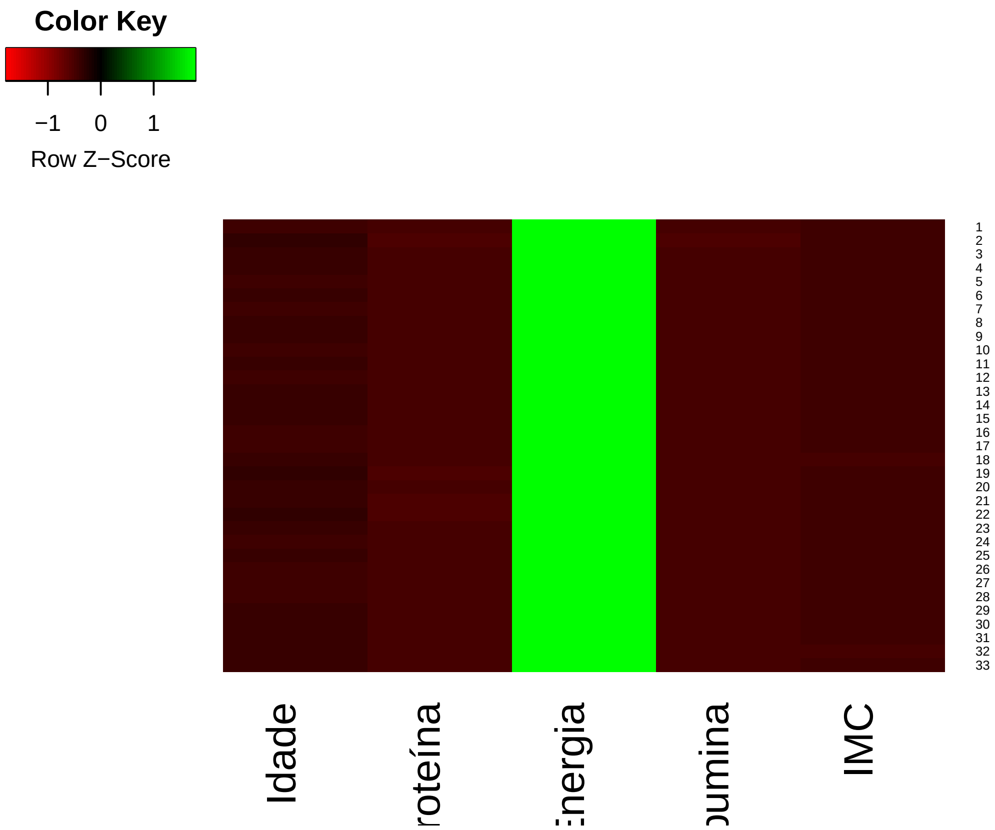

Max
Min
Idade
Proteína
Energia
Albumina
IMC




```
## effect of variables:
##  modified item       Var
##  "height of face   " "Idade"
##  "width of face    " "Proteína"
##  "structure of face" "Energia"
##  "height of mouth  " "Albumina"
##  "width of mouth   " "IMC"
##  "smiling          " "Idade"
##  "height of eyes   " "Proteína"
##  "width of eyes    " "Energia"
##  "height of hair   " "Albumina"
##  "width of hair   "  "IMC"
##  "style of hair   "  "Idade"
##  "height of nose  "  "Proteína"
##  "width of nose   "  "Energia"
##  "width of ear    "  "Albumina"
##  "height of ear   "  "IMC"
```




```
## effect of variables:
##  modified item       Var
##  "height of face   " "Idade"
##  "width of face    " "Proteína"
##  "structure of face" "Energia"
##  "height of mouth  " "Albumina"
##  "width of mouth   " "IMC"
##  "smiling          " "Idade"
##  "height of eyes   " "Proteína"
##  "width of eyes    " "Energia"
##  "height of hair   " "Albumina"
##  "width of hair   "  "IMC"
##  "style of hair   "  "Idade"
##  "height of nose  "  "Proteína"
##  "width of nose   "  "Energia"
##  "width of ear    "  "Albumina"
##  "height of ear   "  "IMC"
```



(b) Obtenha a decomposição espectral e verifique se existe indicação de uma possível redução da dimensão do estudo em questão. Justifique.