



DEPARTAMENTO DE ESTATÍSTICA

20 abril 2023

Entrega 1 - Lista 2

Prof. Dr. George von Borries

Análise Multivariada 1

Aluno: Bruno Gondim Toledo | Matrícula: 15/0167636

13. Considere uma matriz de correlação ($r \times r$) 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 $\rho = 0.1, 0.3, 0.5, 0.7, 0.9$.

Seja $r = 2$ e $\rho = 0.1$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(2 \times 2)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.1, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.1 0.9
##
## $vectors
##      [,1]      [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

Seja $r = 3$ e $\rho = 0.1$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(3 \times 3)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.1, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.2 0.9 0.9
##
## $vectors
##      [,1]      [,2]      [,3]
## [1,] 0.5773503  0.0000000  0.8164966
## [2,] 0.5773503 -0.7071068 -0.4082483
## [3,] 0.5773503  0.7071068 -0.4082483
```

Seja $r = 4$ e $\rho = 0.1$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(4 \times 4)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.1, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.3 0.9 0.9 0.9
##
## $vectors
##      [,1]      [,2]      [,3]      [,4]
## [1,] -0.5  0.5099852  0.000000e+00  0.699939347
## [2,] -0.5  0.4899141  8.756053e-17 -0.714131779
## [3,] -0.5 -0.4999496 -7.071068e-01  0.007096216
## [4,] -0.5 -0.4999496  7.071068e-01  0.007096216
```

Seja $r = 2$ e $\rho = 0.3$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(2 \times 2)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.3, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.3 0.7
##
## $vectors
##      [,1]      [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

Seja $n = 3$ e $\rho = 0.3$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(3 \times 3)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.3, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.6 0.7 0.7
##
## $vectors
```

```
##           [,1]      [,2]      [,3]
## [1,] -0.5773503  0.8164966  0.0000000
## [2,] -0.5773503 -0.4082483 -0.7071068
## [3,] -0.5773503 -0.4082483  0.7071068
```

Seja $n = 4$ e $\rho = 0.3$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(4 \times 4)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.3, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.9 0.7 0.7 0.7
##
## $vectors
##           [,1]      [,2]      [,3]      [,4]
## [1,] -0.5  0.8660254  0.0000000  0.0000000
## [2,] -0.5 -0.2886751  0.0000000  0.8164966
## [3,] -0.5 -0.2886751 -0.7071068 -0.4082483
## [4,] -0.5 -0.2886751  0.7071068 -0.4082483
```

Seja $n = 2$ e $\rho = 0.5$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(2 \times 2)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.5, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.5 0.5
##
## $vectors
##           [,1]      [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

Seja $n = 3$ e $\rho = 0.5$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(3 \times 3)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.5, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 2.0 0.5 0.5
##
## $vectors
##           [,1]      [,2]      [,3]
## [1,] -0.5773503  0.0000000  0.8164966
## [2,] -0.5773503 -0.7071068 -0.4082483
## [3,] -0.5773503  0.7071068 -0.4082483
```

Seja $n = 4$ e $\rho = 0.5$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(4 \times 4)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.5, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 2.5 0.5 0.5 0.5
##
## $vectors
##           [,1]      [,2]      [,3]      [,4]
## [1,] -0.5  0.8660254  0.0000000  0.0000000
## [2,] -0.5 -0.2886751 -0.5773503 -0.5773503
## [3,] -0.5 -0.2886751 -0.2113249  0.7886751
## [4,] -0.5 -0.2886751  0.7886751 -0.2113249
```

Seja $n = 2$ e $\rho = 0.7$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(2 \times 2)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.7, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
```

```
## [1] 1.7 0.3
##
## $vectors
##      [,1]      [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

Seja $n = 3$ e $\rho = 0.7$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(3 \times 3)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.7, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 2.4 0.3 0.3
##
## $vectors
##      [,1]      [,2]      [,3]
## [1,] 0.5773503  0.3555207  0.73503175
## [2,] 0.5773503 -0.8143165 -0.05962589
## [3,] 0.5773503  0.4587958 -0.67540586
```

Seja $n = 4$ e $\rho = 0.7$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(4 \times 4)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.7, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 3.1 0.3 0.3 0.3
##
## $vectors
##      [,1]      [,2]      [,3]      [,4]
## [1,] -0.5  0.0000000  0.0000000  0.8660254
## [2,] -0.5 -0.5773503 -0.5773503 -0.2886751
## [3,] -0.5 -0.2113249  0.7886751 -0.2886751
## [4,] -0.5  0.7886751 -0.2113249 -0.2886751
```

Seja $n = 2$ e $\rho = 0.9$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(2 \times 2)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.9, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 1.9 0.1
##
## $vectors
##      [,1]      [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

Seja $n = 3$ e $\rho = 0.9$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(3 \times 3)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.9, \forall i \neq j$) são:

```
## eigen() decomposition
## $values
## [1] 2.8 0.1 0.1
##
## $vectors
##      [,1]      [,2]      [,3]
## [1,] 0.5773503  0.4586617  0.67549693
## [2,] 0.5773503 -0.8143284  0.05946423
## [3,] 0.5773503  0.3556666 -0.73496116
```

Seja $n = 4$ e $\rho = 0.9$; Os autovalores e autovetores calculados da matriz de correlação $\mathbf{X}_{(4 \times 4)}$ (em que $x_{ii} = 1$ e $x_{ij} = 0.9, \forall i \neq j$) são:

```
## eigen() decomposition
```

```
## $values
## [1] 3.7 0.1 0.1 0.1
##
## $vectors
##      [,1]      [,2]      [,3]      [,4]
## [1,] -0.5  0.8660254  0.0000000  0.000000e+00
## [2,] -0.5 -0.2886751 -0.8164966 -1.751211e-16
## [3,] -0.5 -0.2886751  0.4082483  7.071068e-01
## [4,] -0.5 -0.2886751  0.4082483 -7.071068e-01
```

Generalizando, seja a matriz de correlações $\Omega(r \times r)$, com correlação ρ constante fora da diagonal principal; temos uma matriz do tipo

$$\Omega_{(n \times n)} = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho & \rho \\ \rho & 1 & \rho & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \rho & \rho \\ \rho & \cdots & \rho & 1 & \rho & \vdots \\ \vdots & & & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & \cdots & \rho & 1 \end{bmatrix}$$

Para $\Omega \cdot v = \lambda v \quad \forall v \neq 0$, e sendo $u_{(1 \times n)} = [1, 1, \dots, 1]$ temos:

$$\lambda v = \Omega \cdot v = \rho \cdot u + (1 - \rho) \cdot v$$

Então teremos o autovalor $\rho(n - 1) + 1$ e os autovalores $\lambda = 1 - \rho$ com multiplicidade $r - 1$. Ou seja, $\lambda v = [(\rho(n - 1) + 1), (1 - \rho), (1 - \rho), \dots, (1 - \rho)] \square$

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,

$$\mathbf{A}\mu'_x = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 3 \end{bmatrix}$$

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

A matriz de variância-covariância $\Omega_{(3 \times 3)}$ de \mathbf{AX} é dada por:

$$\text{Cov}(\mathbf{AX}) = \mathbf{A}\Sigma_x\mathbf{A}^T = \Omega_{3 \times 3}$$

, que é:

$$\begin{array}{ccc} \hline 6 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 36 \\ \hline \end{array}$$

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

Todos os pares $\omega_{ij} \forall i \neq j$ apresentam covariância $\sigma = 0$

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.

Começando com as medidas descritivas, temos:

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:

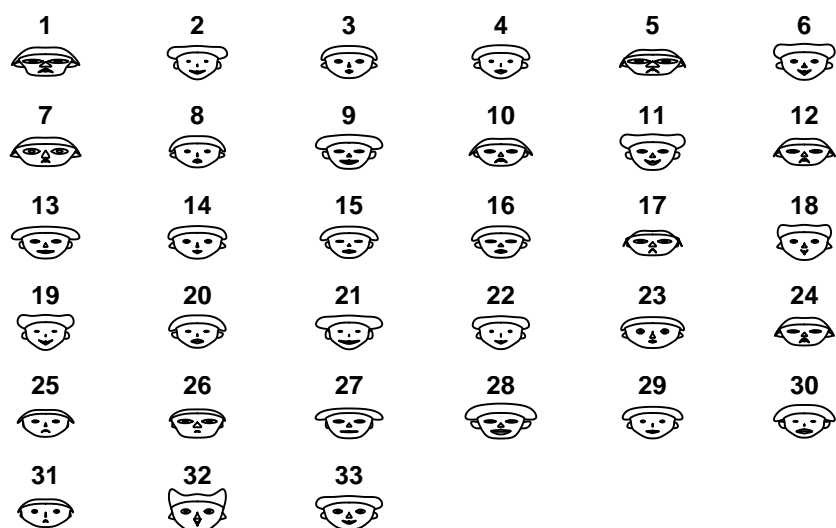
	Idade	Proteína	Energia	Albumina	IMC
Idade	174.9393939	-2.8203693	-3736.92562	0.5121212	19.3448864
Proteína	-2.8203693	0.2079267	152.26715	-0.0321449	-1.4953153
Energia	-3736.9256250	152.2671531	249960.54006	-11.0733438	-573.0939687

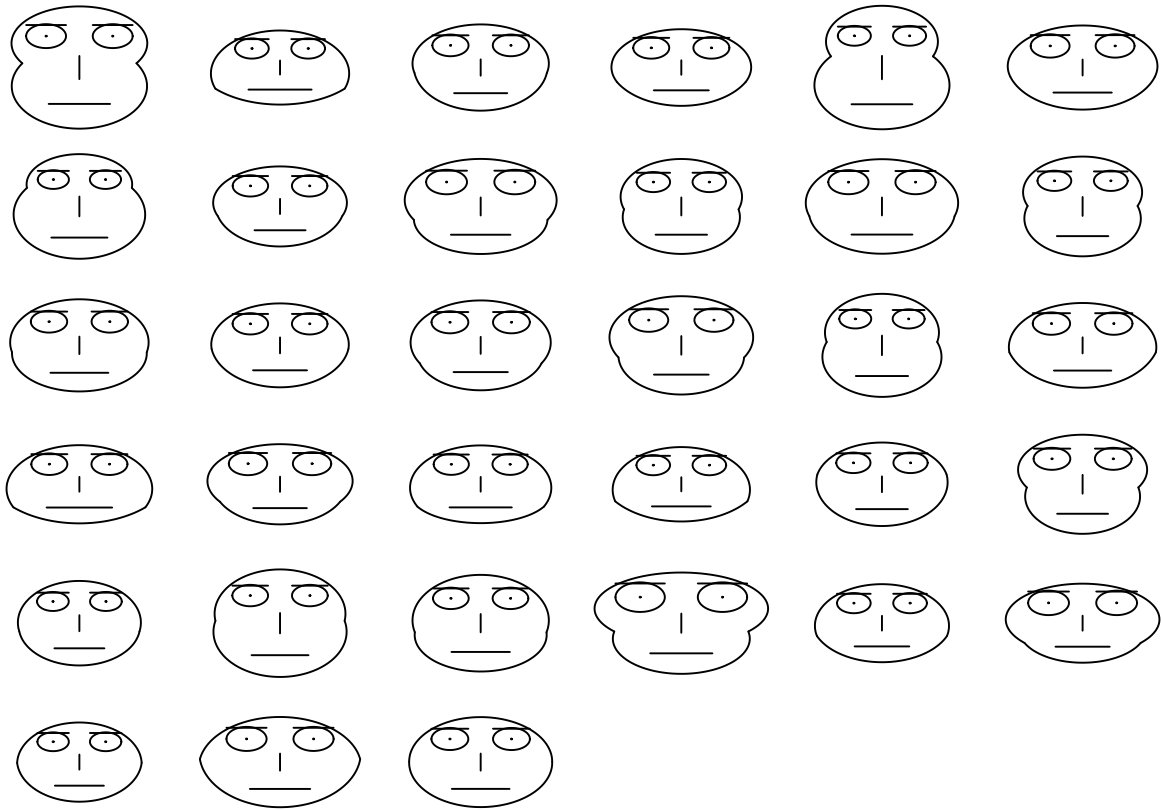
	Idade	Proteína	Energia	Albumina	IMC
Albumina	0.5121212	-0.0321449	-11.07334	0.0613258	0.4797727
IMC	19.3448864	-1.4953153	-573.09397	0.4797727	22.0800568

Matriz de correlações:

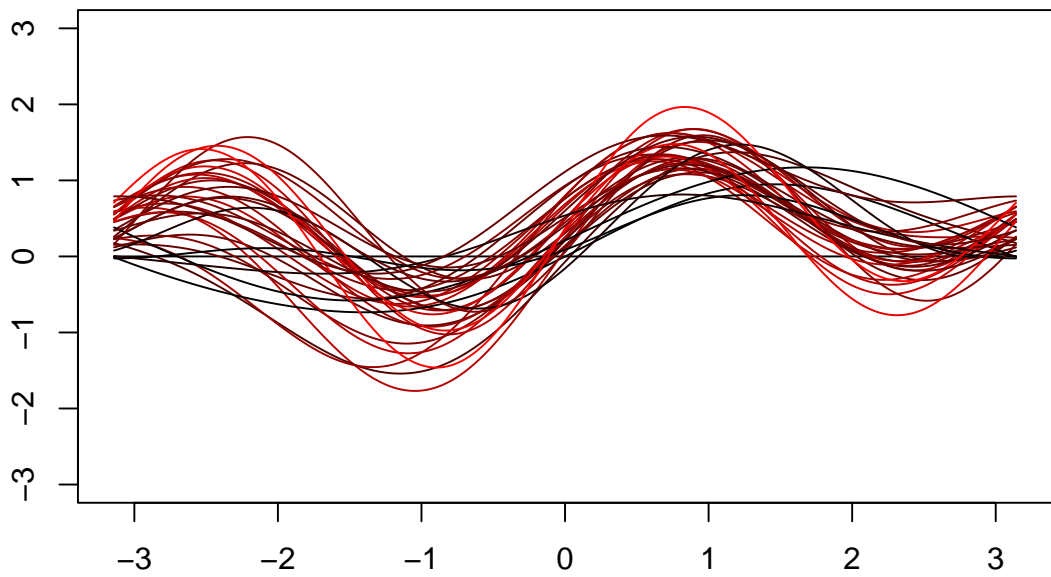
	Idade	Proteína	Energia	Albumina	IMC
Idade	1.0000000	-0.4676350	-0.5651125	0.1563535	0.3112593
Proteína	-0.4676350	1.0000000	0.6679060	-0.2846658	-0.6978749
Energia	-0.5651125	0.6679060	1.0000000	-0.0894379	-0.2439439
Albumina	0.1563535	-0.2846658	-0.0894379	1.0000000	0.4123006
IMC	0.3112593	-0.6978749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

Agora, representando graficamente:

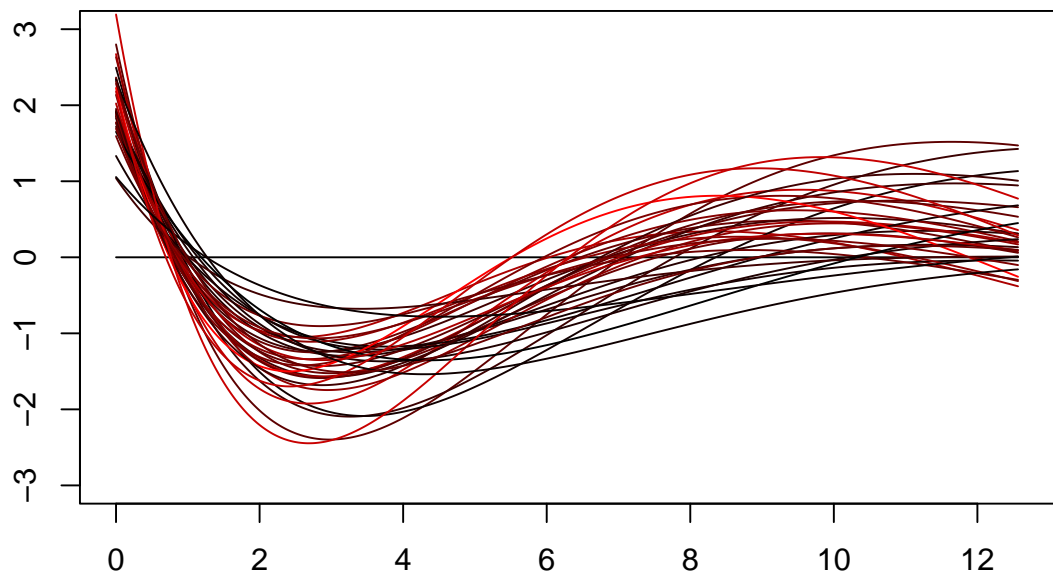




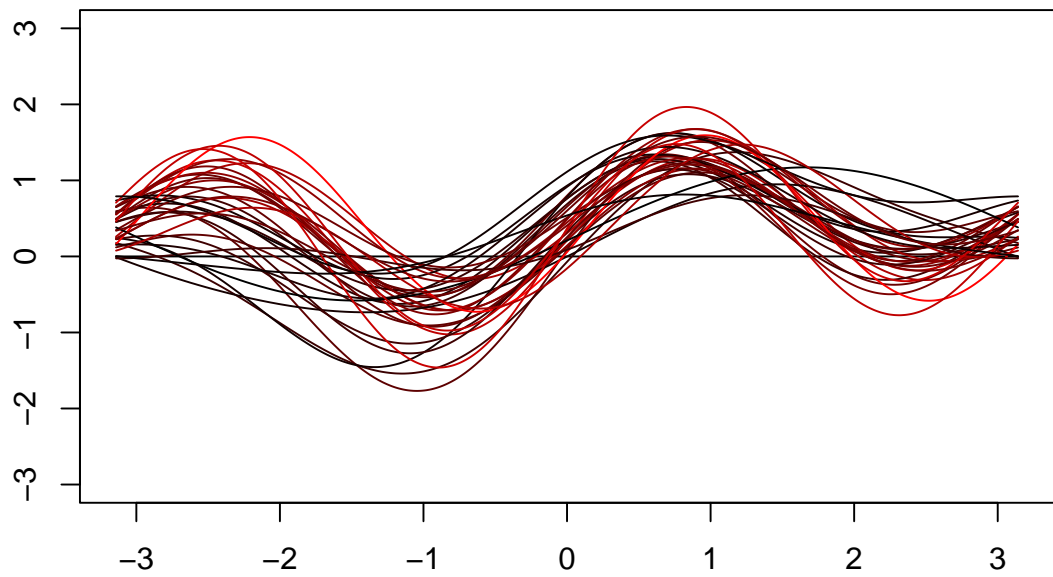
Andrews Plot 1

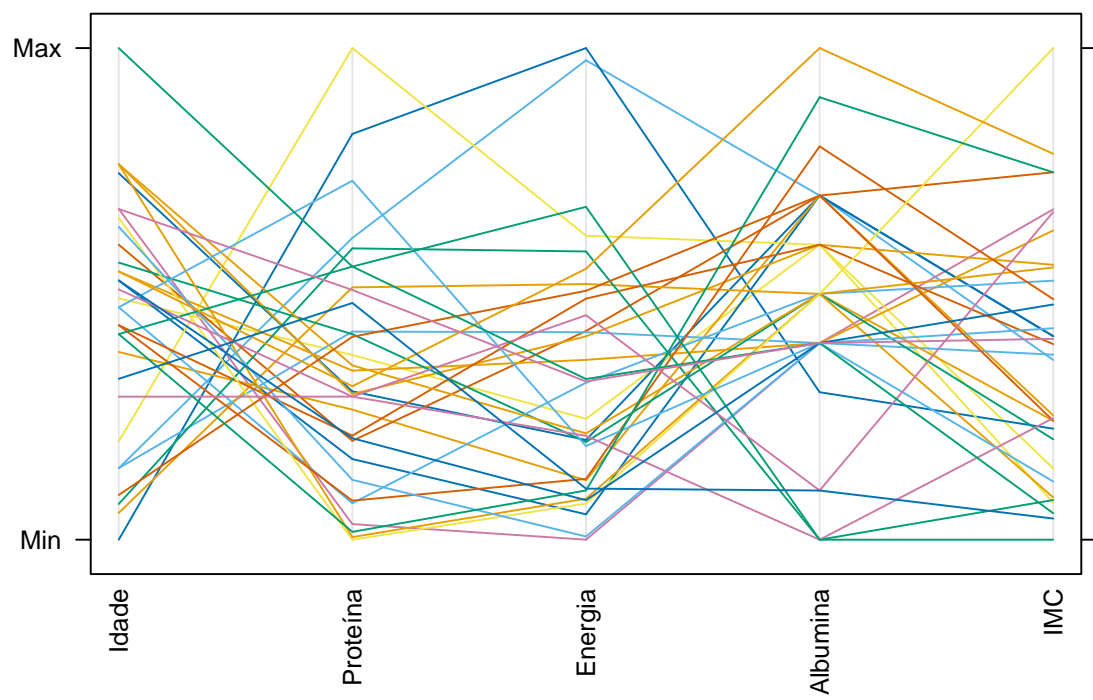
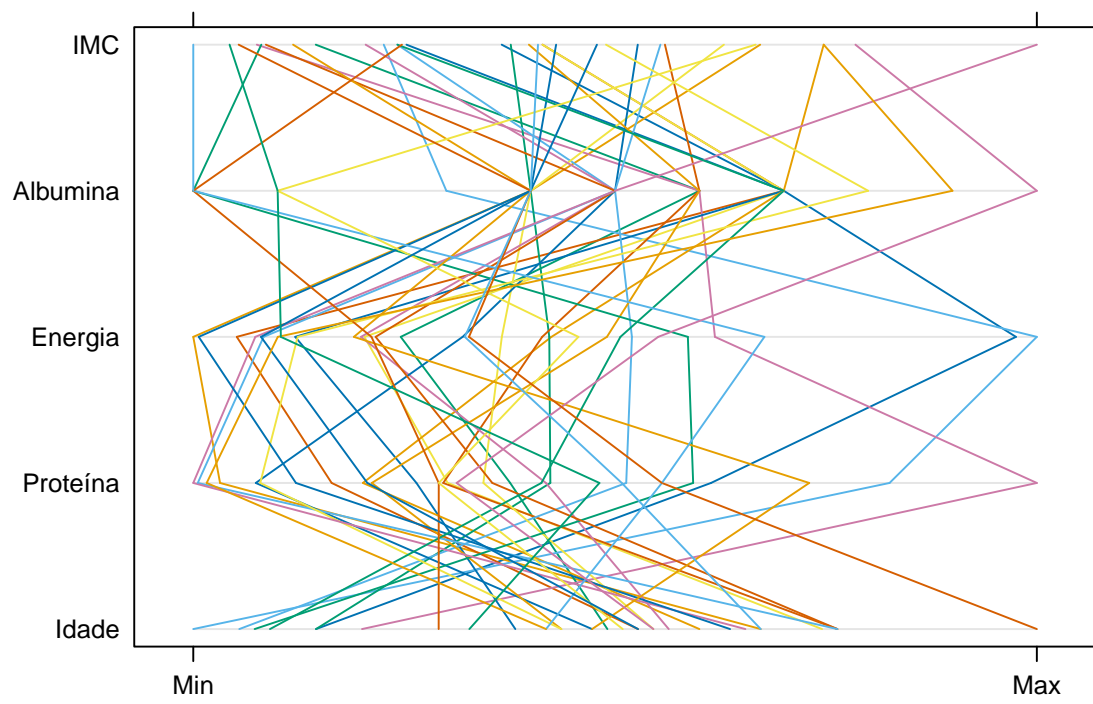


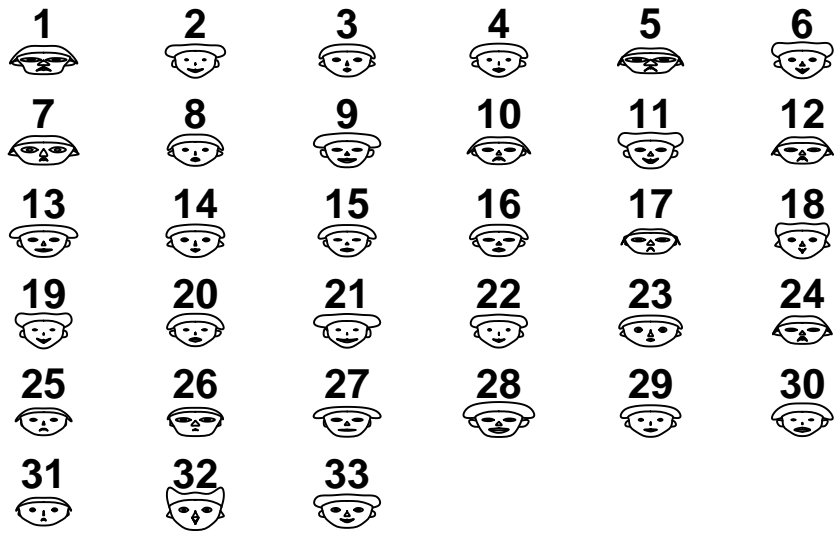
Andrews Plot 2



Andrews Plot 3



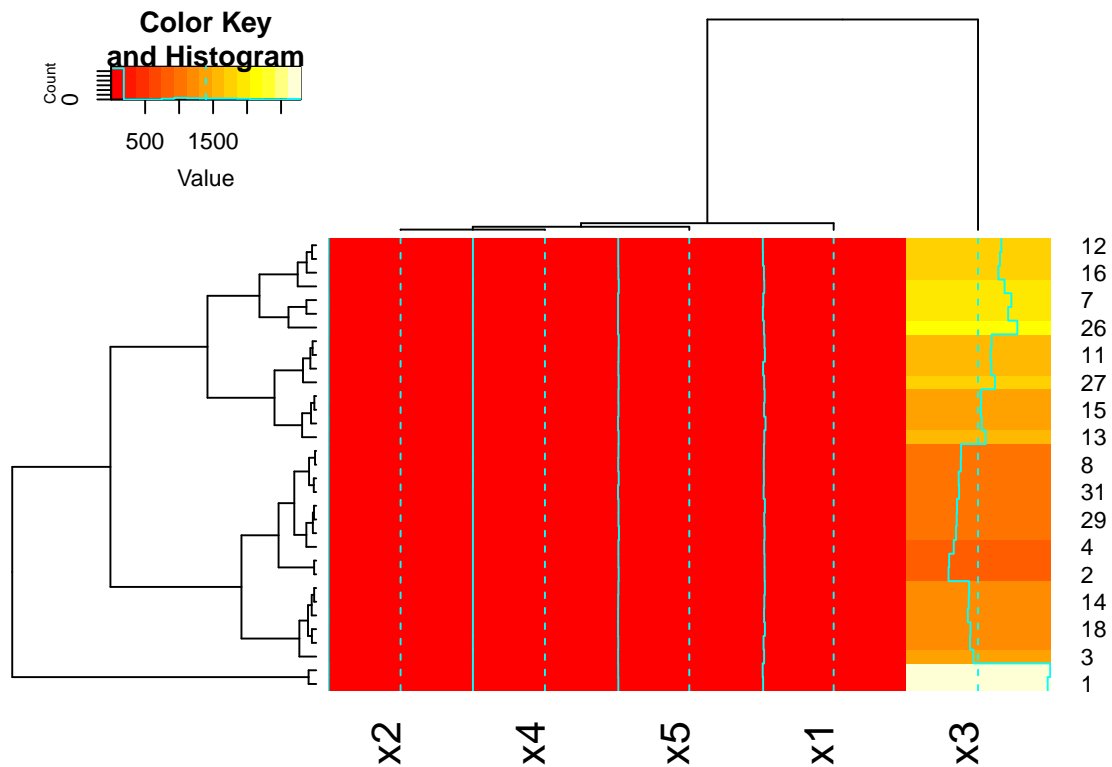




```
## effect of variables:
## modified item      Var
## "height of face"   "Idade"
## "width of face"    "Proteína"
## "structure of face" "Energia"
## "height of mouth"  "Albumina"
## "width of mouth"   "IMC"
## "smiling"          "Idade"
## "height of eyes"   "Proteína"
## "width of eyes"    "Energia"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

```
## $d
## [1] 8952.322440 152.222259 28.608980 2.767222 1.287161
##
## $u
##      [,1]      [,2]
## [1,] -0.30592194 -0.324267973
## [2,] -0.09226951 0.270009424
## [3,] -0.14612013 0.085552992
## [4,] -0.10353565 0.188927788
## [5,] -0.31131788 -0.389764025
## [6,] -0.13673702 0.200854095
## [7,] -0.22770684 -0.170074647
## [8,] -0.11869546 0.104305485
## [9,] -0.18531142 0.073147111
## [10,] -0.18466507 -0.095650959
## [11,] -0.18293641 0.125870242
## [12,] -0.20611534 -0.172401493
## [13,] -0.17241102 0.073998455
## [14,] -0.13580762 0.132118095
## [15,] -0.16222091 0.064305096
## [16,] -0.19969861 -0.023795158
## [17,] -0.22064851 -0.200170061
## [18,] -0.13964746 0.188375045
## [19,] -0.11054700 0.261901520
## [20,] -0.11933414 0.131218754
## [21,] -0.10839331 0.246625504
## [22,] -0.09367140 0.245270087
## [23,] -0.13394239 0.101143632
## [24,] -0.20313418 -0.154326317
```

```

## [25,] -0.13848258 0.035620244
## [26,] -0.24060312 -0.123122442
## [27,] -0.19231188 0.025547528
## [28,] -0.21305452 0.003909098
## [29,] -0.10983205 0.179571733
## [30,] -0.11418641 0.145078879
## [31,] -0.11494705 0.084176526
## [32,] -0.16394515 0.221944065
## [33,] -0.16280525 0.126666699
##
## $v
##           [,1]           [,2]
## [1,] -0.0286968729 0.9432313346
## [2,] -0.0007260725 0.0008550346
## [3,] -0.9994775101 -0.0319941722
## [4,] -0.0024219141 0.0448994730
## [5,] -0.0146561766 0.3275275997
##
##           [,1] [,2]
## [1,] -0.3 -0.3
## [2,] -0.1 0.3
## [3,] -0.1 0.1
## [4,] -0.1 0.2
## [5,] -0.3 -0.4
## [6,] -0.1 0.2
## [7,] -0.2 -0.2
## [8,] -0.1 0.1
## [9,] -0.2 0.1
## [10,] -0.2 -0.1
## [11,] -0.2 0.1
## [12,] -0.2 -0.2
## [13,] -0.2 0.1
## [14,] -0.1 0.1
## [15,] -0.2 0.1
## [16,] -0.2 0.0
## [17,] -0.2 -0.2
## [18,] -0.1 0.2
## [19,] -0.1 0.3
## [20,] -0.1 0.1
## [21,] -0.1 0.2
## [22,] -0.1 0.2
## [23,] -0.1 0.1
## [24,] -0.2 -0.2
## [25,] -0.1 0.0
## [26,] -0.2 -0.1
## [27,] -0.2 0.0
## [28,] -0.2 0.0
## [29,] -0.1 0.2
## [30,] -0.1 0.1
## [31,] -0.1 0.1
## [32,] -0.2 0.2
## [33,] -0.2 0.1
##
##           [,1] [,2]
## [1,] 0 0.9
## [2,] 0 0.0
## [3,] -1 0.0
## [4,] 0 0.0
## [5,] 0 0.3

```

```
##      [,1] [,2]
## [1,] 8952.3 0.0
## [2,] 0.0 152.2

##      Idade Proteína Energia Albumina IMC
## [1,] 73.1 1.6 53.2 4.2 37.8
## [2,] 19.9 0.5 -71.0 3.9 16.1
## [3,] 37.3 1.1 411.8 4.1 15.4
## [4,] 25.6 0.7 30.2 4.2 15.9
## [5,] 78.8 2.0 101.8 3.8 39.8
## [6,] 37.6 1.0 327.3 4.2 15.9
## [7,] 62.4 2.3 247.8 4.1 27.8
## [8,] 31.3 0.9 166.3 4.2 17.4
## [9,] 43.3 0.8 -132.7 4.2 26.6
## [10,] 45.7 1.2 -137.7 3.9 28.9
## [11,] 52.3 1.0 -154.2 4.1 23.1
## [12,] 54.4 1.4 54.6 4.0 30.9
## [13,] 40.3 1.1 -248.2 3.9 24.4
## [14,] 41.3 1.2 319.3 4.0 16.5
## [15,] 36.3 0.6 -339.3 4.0 22.5
## [16,] 48.0 0.8 -3.5 4.1 24.7
## [17,] 55.4 1.6 184.8 3.5 27.9
## [18,] 38.6 1.1 353.4 4.0 9.8
## [19,] 24.9 0.4 92.6 4.0 13.9
## [20,] 34.3 0.6 171.9 4.3 21.8
## [21,] 32.6 0.4 73.4 4.0 26.8
## [22,] 31.6 0.7 -58.3 3.9 16.2
## [23,] 36.3 1.8 302.8 3.9 14.9
## [24,] 56.4 1.2 27.8 4.2 30.9
## [25,] 40.0 1.0 343.7 3.5 21.9
## [26,] 60.7 1.5 363.0 3.5 21.9
## [27,] 52.0 1.0 -69.9 3.6 29.7
## [28,] 54.0 1.0 115.8 4.5 31.9
## [29,] 25.6 0.8 86.6 3.9 17.1
## [30,] 33.3 0.5 125.7 4.4 26.6
## [31,] 28.3 1.3 132.9 3.6 13.5
## [32,] 51.6 1.5 -324.5 3.9 9.2
## [33,] 47.3 1.4 -334.3 3.9 20.3

## [1] 99.64203
```

Com este resultado, concluímos que a redução de dimensionalidade neste caso não só é possível como retêm a absoluta maior parte da informação (>99%). Logo, a redução se justifica pela retenção da informação dado o ganho de eficiência computacional para a matriz reduzida à dois vetores singulares.