Universidade Federal do Amazonas
Instituto de Ciências Exatas
Curso de Bacharelado em Estatística
IEE062: Estatística Multivariada II
Exercício Escolar 1
Entrega: 04/06/2024

1. Faça um relatório para a questão abaixo, a Tabela 7.2 está no fim desse documento.

Carry out a principal component analysis on the temperature data of Table 7.2. Use both **S** and **R**. Which do you think is more appropriate here? Show the percent of variance explained. Based on the average eigenvalue or a scree plot, decide how many components to retain. Can you interpret the components of either **S** or **R**?

The data in Table 7.2 relate temperature, humidity, and evaporation (courtesy of R. J. Freund). The variables are

y_1 = maximum daily air temperature,

y_2 = minimum daily air temperature,

y_3 = integrated area under daily air temperature curve, that is, a measure of average air temperature,

y_4 = maximum daily soil temperature,

y_5 = minimum daily soil temperature,

y_6 = integrated area under soil temperature curve,

y_7 = maximum daily relative humidity,

y_8 = minimum daily relative humidity,

y_9 = integrated area under daily humidity curve,

y_{10} = total wind, measured in miles per day,

y_{11} = evaporation.

2. Faça um relatório para as questões abaixo, 8.18 – 8.21. Faça também o biplot. As Tabelas 1.9 e 8.18 são dados de Johnson & Wichern e estão disponíveis no Classroom.

8.18. The data on national track records for women are listed in Table 1.9.

(a) Obtain the sample correlation matrix **R** for these data, and determine its eigenvalues and eigenvectors.

(b) Determine the first two principal components for the standardized variables. Prepare a table showing the correlations of the standardized variables with the components, and the cumulative percentage of the total (standardized) sample variance explained by the two components.

(c) Interpret the two principal components obtained in Part b. (Note that the first component is essentially a normalized unit vector and might measure the athletic excellence of a given nation. The second component might measure the relative strength of a nation at the various running distances.)

(d) Rank the nations based on their score on the first principal component. Does this ranking correspond with your inituitive notion of athletic excellence for the various countries?

8.19. Refer to Exercise 8.18. Convert the national track records for women in Table 1.9 to speeds measured in meters per second. Notice that the records for 800 m, 1500 m, 3000 m, and the marathon are given in minutes. The marathon is 26.2 miles, or 42,195 meters, long. Perform a principal components analysis using the covariance matrix **S** of the speed data. Compare the results with the results in Exercise 8.18. Do your interpretations of the components differ? If the nations are ranked on the basis of their score on the first principal component, does the subsequent ranking differ from that in Exercise 8.18? Which analysis do you prefer? Why?

8.20. The data on national track records for men are listed in Table 8.6. (See also the data on national track records for men on the website www.prenhall.com/statistics) Repeat the principal component analysis outlined in Exercise 8.18 for the men. Are the results consistent with those obtained from the women's data?

8.21. Refer to Exercise 8.20. Convert the national track records for men in Table 8.6 to speeds measured in meters per second. Notice that the records for 800 m, 1500 m, 5000 m, 10,000 m and the marathon are given in minutes. The marathon is 26.2 miles, or 42,195 meters, long. Perform a principal component analysis using the covariance matrix **S** of the speed data. Compare the results with the results in Exercise 8.20. Which analysis do you prefer? Why?

Tabela 1: Tabela 7.2 Temperature, Humidity, and Evaporation

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}
84	65	147	85	59	151	95	40	398	273	30
84	65	149	86	61	159	94	28	345	140	34
79	66	142	83	64	152	94	41	368	318	33
81	67	147	83	65	158	94	50	406	282	26
84	68	167	88	69	180	93	46	379	311	41
74	66	131	77	67	147	96	73	478	446	4
73	66	131	78	69	159	96	72	462	294	5
75	67	134	84	68	159	95	70	464	313	20
84	68	161	89	71	195	95	63	430	455	31
86	72	169	91	76	206	93	56	406	604	36
88	73	176	91	76	206	94	55	393	610	43
90	74	187	94	76	211	94	51	385	520	47
88	72	171	94	75	211	96	54	405	663	45
58	72	171	92	70	201	95	51	392	467	45
81	69	154	87	68	167	95	61	448	184	11
79	68	149	83	68	162	95	59	436	177	10
84	69	160	87	66	173	95	42	392	173	30
84	70	160	87	68	177	94	44	392	76	29
84	70	168	88	70	169	95	48	396	72	23
77	67	147	83	66	170	97	60	431	183	16
87	67	166	92	67	196	96	44	379	76	37
89	69	171	92	72	199	94	48	393	230	50
89	72	180	94	72	204	95	48	394	193	36
93	72	186	92	73	201	94	47	386	400	54
93	74	188	93	72	206	95	47	389	339	44
94	75	199	94	72	208	96	45	370	172	41
93	74	193	95	73	214	95	50	396	238	45
93	74	196	95	70	210	96	45	380	118	42
96	75	198	95	71	207	93	40	365	93	50
95	76	202	95	69	202	93	39	357	269	48
84	73	173	96	69	173	94	58	418	128	17
91	71	170	91	69	168	94	44	420	423	20
88	72	179	89	70	189	93	50	399	415	15
89	72	179	95	71	210	98	46	389	300	42
91	72	182	96	73	208	95	43	384	193	44
92	74	196	97	75	215	96	46	389	195	41
94	75	192	96	69	198	95	36	380	215	49
96	75	195	95	67	196	97	24	354	185	53
93	76	198	94	75	211	93	43	364	466	53
88	74	188	92	73	198	95	52	405	399	21
88	74	178	90	74	197	95	61	447	232	1
91	72	175	94	70	205	94	42	380	275	44
92	72	190	95	71	209	96	44	379	166	44
92	73	189	96	72	208	93	42	372	189	46
94	75	194	95	71	208	93	43	373	164	47
96	76	202	96	71	208	94	40	368	139	50