

PRÁCTICO 3 - ANÁLISIS DE COMPONENTES PRINCIPALES

1. Dada la matriz de covarianzas

$$S = \begin{pmatrix} 1+d & 1 & 1 \\ 1 & 1+d & 1 \\ 1 & 1 & 1+d \end{pmatrix}$$

encuentre las componentes principales. Calcule la proporción de variabilidad explicada por cada uno y las correlaciones entre las componentes y las variables.

2. Se considera la matriz de datos estandarizada X y la matriz de factores

	x_1	x_2	x_3
\mathbf{x}_1	-1.095	1.426	-0.475
\mathbf{x}_2	-0.548	-0.475	1.426
\mathbf{x}_3	0.548	-0.095	-0.856
\mathbf{x}_4	1.095	-0.856	-0.095

columnas w_1, w_2, w_3

	Factor 1	Factor 2	Factor 3
1	0.894	-0.446	0.048
2	-0.972	C	0.051
3	0.178	0.983	0.034

- Encontrar, justificando, el valor de C .
- Hallar los valores propios y verificar que la suma es igual a 3.
- Hallar los vectores a_1, a_2 y a_3 .
- Calcular los porcentajes de las inercias definidas por los tres ejes
- Hallar las coordenadas del individuo 2 sobre cada uno de los ejes.

3. Consideramos la siguiente matriz de datos

	x_1	x_2	x_3
\mathbf{x}_1	1	0	0
\mathbf{x}_2	1	2	0
\mathbf{x}_3	2	2	2
\mathbf{x}_4	0	0	2

- Centrar y reducir la matriz
- Calcular la matriz de varianzas covarianzas, sus valores propios y vectores propios.

- c) Completar la tabla siguiente:

	Z_1	Z_2	Z_3	cal Z_1	cal Z_2	cal Z_3	ctr Z_1	ctr Z_2	ctr Z_3
\mathbf{x}_1									
\mathbf{x}_2									
\mathbf{x}_3									
\mathbf{x}_4									

donde cal Z_i es la calidad sobre el eje i (en%) y ctr Z_i es la contribución a la construcción del eje i (en%). Se justificarán los cálculos y se verificarán con R.

- Hacer la representación gráfica sobre el plano dado por Z_1 y Z_2 .

e) Completar la tabla siguiente:

	Z_1	Z_2	Z_3	cal Z_1	cal Z_2	cal Z_3	ctr Z_1	ctr Z_2	ctr Z_3
x_1									
x_2									
x_3									

f) Hacer la representación gráfica sobre el plano dado por Z_1 y Z_2 de individuos y variables.

g) Individuos y variables suplementarias. ¿Qué son? Construir la representación gráfica del individuo de coordenadas $(0, 2, 0)$ y de la variable de coordenadas $(1, -1, 1, -1, 0)$

4. Leer el laboratorio 2.

5. Consultar el laboratorio en <http://factominer.free.fr/factomethods/principal-components-analysis.html>

6. Se considera la siguiente tabla de notas obtenidas por los estudiantes de un liceo en cinco parciales (sobre 40 puntos)

	Matemática	Física	Música	Dibujo	Id. Español
Juan	12	12	10	11	16
Alina	16	16	16	16	18
Ana	12	14	22	19	22
Mónica	29	29	31	30	16
Daniel	28	28	24	24	20
Andrés	22	20	11	14	26
Pedro	11	14	28	23	20
Valentina	26	25	17	19	24
Sandra	18	19	25	24	36

Realizar un análisis de componentes principales completo a partir de la matriz de datos reducidas. Deberán ser incluidas las calidades y contribuciones de cada variable et individuo.

7. Kernel PCA

- Buscar información sobre Kernel PCA. Se deberá realizar un informe donde esté claramente explicada la técnica, su relación con el Análisis de Componentes Principales (PCA) tradicional y en qué lo mejora. Se sugiere consultar el video *Part of the Course "Statistical Machine Learning", Summer Term 2020, Ulrike von Luxburg, University of Tbingen* <https://www.youtube.com/watch?v=7kfkGyhPDU0>
- Buscar un conjunto de datos (por ejemplo el swiss roll dataset o los tres circulos concéntricos) donde la técnica PCA falla pero KernelPCA no. Implementar el estudio en R.