Análisis Multivariado

Producto académico 02

Kevin Heberth Haquehua Apaza

23 de julio del 2025

Table of Contents

# Examen AED multivariado, ACP y AFE

## Ejercicio 1:

Sea un vector aleatorio con matriz de varianza-covarianza dada por

1. Determine las componentes principales y .
2. Calcule la proporción de la varianza total explicada por la primera componente principal.
3. Calcule la matriz de correlaciones a partir de la matriz de covarianzas y determine las componentes principales y a partir de . Calcule la proporción de la varianza total explicada por
4. Calcule la correlación entre las variables y las componentes principales, es decir, calcule

### Solución

a y b)

1. **Calcule la matriz de correlaciones a partir de la matriz de covarianzas y determine las componentes principales y a partir de . Calcule la proporción de la varianza total explicada por**

S <- matrix(c(4, 1, 1, 3), 2, 2) ; S

## [,1] [,2]  
## [1,] 4 1  
## [2,] 1 3

Calcular la matriz de correlaciones

D\_inv <- diag(1 / sqrt(diag(S)))  
R <- D\_inv %\*% S %\*% D\_inv  
R

## [,1] [,2]  
## [1,] 1.0000000 0.2886751  
## [2,] 0.2886751 1.0000000

Ahora hallar los componentes principales

pca <- princomp(covmat = R, cor = TRUE)

Y las varianzas explicadas

explained\_var <- pca$sdev^2 / sum(pca$sdev^2)  
explained\_var

## Comp.1 Comp.2   
## 0.6443376 0.3556624

En el cálculo manual no hya tanta diferencia, debido al cálculo con decimales.

1. **Calcule la correlación entre las variables y las componentes principales, es decir, calcule**

L <- unclass(pca$loadings) ; L

## Comp.1 Comp.2  
## [1,] 0.7071068 0.7071068  
## [2,] 0.7071068 -0.7071068

# Autovalores  
lambda <- pca$sdev^2  
  
# Correlaciones  
cor <- L %\*% diag(sqrt(lambda))  
cor

## [,1] [,2]  
## [1,] 0.8027064 0.5963744  
## [2,] 0.8027064 -0.5963744

## Ejercicio 2:

La siguiente tabla muestra los datos sobre la longitud de huesos registrados de 20 jóvenes a los 8, 8.5, 9 y 9.5 años respectivamente; Verificar si alguno de los individuos es considerado un dato atípico multivariado. Realizar la comprobación paso a paso como se realizó en clase (matricialmente), además tienes que comprobarlo con la función directa en el R

### Solución

Aca la solución

## Ejercicio 3:

En el Excel *(pregunta3.xlsx)* se muestran los valores de cinco variables obtenidas en 20 alumnos que quieren entrar a alguna universidad del consejo de rectores. Las variables en estudio son la distancia en kilómetros al lugar del colegio en el que estudiaban (DIST), el promedio de horas que hacían actividad física a la semana (EF), índice de masa corporal (IMC), IQ (coeficiente intelectual) y NEM (promedio de notas con el cual postulan a las universidades). Se quiere determinar las relaciones existentes entre dichas variables intentando reducir la dimensionalidad del problema vía un análisis factorial exploratorio.

1. Sobre el conjunto de datos halle si se cumple o no la normalidad multivariada.
2. Verifique si existe presencia de valores outliers para la data.
3. Realice el análisis factorial exploratorio vía componentes principales e interprete y justifique sus conclusiones.

### Solución

Aca va la solución