**Universidad del Valle - Escuela de Estadística**

**Curso: Análisis Multivariante. Profesor: Andrés Felipe Ochoa Muñoz**

**Laboratorio 1: Construcción de la Inercia**

Los siguientes datos se refieren a la altura de una planta X1 (en metros), su longitud radicular x2 (en cm), su área foliar x3 (en ), su peso en pulpa x4 (en grs), de una variedad de manzano. Los datos son los siguientes:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Planta** | **X1** | **X2** | **X3** | **X4** |
| 1 | 1.89 | 53 | 4.5 | 120 |
| 2 | 1.92 | 62 | 5.2 | 136 |
| 3 | 1.95 | 71 | 5.4 | 145 |
| 4 | 2.11 | 75 | 6.1 | 156 |
| 5 | 1.78 | 58 | 5.0 | 127 |
| 6 | 2.12 | 77 | 6.5 | 160 |
| 7 | 2.06 | 82 | 7.6 | 169 |
| 8 | 2.19 | 78 | 7.3 | 164 |
| 9 | 1.93 | 60 | 5.5 | 142 |
| 10 | 1.84 | 72 | 5.7 | 147 |

Teniendo en cuenta lo anterior, realice los siguientes puntos:

1. Para la matriz de datos original calcule la sumatoria de las distancias al cuadrado de cada individuo al centro de gravedad (concepto inercia); sumatoria de las distancias divida sobre el número de árboles. Realice el mismo cálculo para datos centrados. (Tenga presente las métricas)
2. Para la matriz de datos estandarizada realice el cálculo anterior. Que diferencias aprecia?
3. Calcule la matriz de varianzas y covarianzas
4. Calcule la matriz de correlaciones
5. Realice la descomposición en valores y vectores propios de las matrices del punto 3 y 4.
6. Las descomposiciones anteriores son idénticas? En caso contrario a que se debe su diferencia? Las dos descomposiciones representan la misma cantidad de información contenida en el conjunto de variables?
7. De acuerdo al tipo de datos original, a partir de cual matriz (la del punto 3 o 4) es mejor realizar un Análisis Multivariado?
8. Usando los vectores propios encontrados en los puntos 3 y 4. Calcule la proyección de la matriz de datos originales y centrados asociada a los vectores propios del punto 3. Y con la matriz estandarizada calcule la proyección ahora con los vectores propios del punto 4. Que significa dicha proyección? (Sugerencia: Utilice un individuo y un vector propio)
9. Seleccione 3 variables y realice un gráfico en tres dimensiones para la matriz original y la matriz estandarizada.
10. Con las proyecciones encontradas en el punto 8, realice un gráfico en 2 dimensiones (usando las dos primeras componentes). Además realice el gráfico en tres dimensiones.
11. Conforme un informe con los hallazgos más importantes del conjunto de datos. Utilice los análisis que considere necesarios. Realice los Análisis usando el software R, con un informe máximo de 8 páginas.